Assessing the Usability of Urdu Learning Mobile Apps for Children

Noor Hussain, Zahid Hussain[✉], and Baqar Ali

Department of Information Technology,
Quaid-e-Awam University of Engineering, Science and Technology,
Nawabshah, Pakistan
`noorhussainl02@gmail.com`,
`{zhussain,bazardari34}@quest.edu.pk`

Abstract. This study evaluated the usability of the children's Urdu learning mobile Apps for Tablet-PCs. The pre-test questionnaire based on demographic information of the participants, post-test questionnaire using the Fun Toolkit and the usability tasks were developed for collecting data from the children. The usability test was conducted by 232 participating children having age between 5 and 10 years of the primary schools. After filling the demographic questionnaire, the usability tasks were given to the children where time was recorded for each of the tasks. Afterwards, the post-test questionnaire was administered among the children in order to gather their opinion about the apps. The study result illustrated that the App3 was the best App in terms of efficiency and effectiveness. According to the Smileyometer results, the children reported more fun in using the App3. The Again-Again table results also show that the children wanted to use the App3 again and again. This study proposed the best Urdu learning mobile App from learning perspective for the children which is recommended to be complemented with the existing traditional classroom learning in the schools of the country for enhancing their learning outcomes.

Keywords: Child-computer interaction · Usability testing
Urdu Learning Mobile Apps · Fun Toolkit

1 Introduction

Information and communication technology (ICT) has been heavily utilized in education sector of the developed world. However in the developing countries, the utilization of ICT in education, specifically in primary education, is still evolving. Usability plays a crucial role for the success of an ICT system [2]. Measuring usability is difficult when it involves children [3]. However, Read and MacFarlane state that usability in this age group can be measured in terms of fun [4]. Read et al. have worked on evaluating usability of computer application, school trip and website designing [3]. They have introduced Fun Toolkit scale for measuring the children's self-reported experience of fun [3]. Read et al. have explored the Fun Toolkit as measuring tool specially designed for children aged from 5 to 10 years. They have discussed about the relationship between fun and usability [5].

This research paper is based on the master's thesis of the first author [1].

Several researchers have worked on child-computer interaction and have assessed usability or fun for various technologies, i.e., fun games for children [6], mobile applications for educational purpose [7], educational software for learning English [8], the educational app for improving teaching-learning methods [9], measuring fun for Sindhi to English mobile learning App using Fun Toolkit [10], Interactive Learning System [11], measuring entertainment products [12], Child e-Learning Programs [13], and computer applications [14].

There is not much work carried out in the developing countries, specifically, in Pakistan, on evaluating the usability of the learning mobile Apps. To the best of our knowledge, no work has been carried out on the usability assessment of the Urdu learning mobile Apps for children. Therefore, this study tries to fill this gap. In this paper, the results of the usability study regarding the Urdu learning mobile Apps for children are presented.

1.1 Selection of the Urdu Learning Mobile Apps

From the Urdu learning mobile Apps available at Google Play store, four Apps were selected on the basis of their rankings, number of downloads, ratings and the reviews. Moreover, these Apps had similar activities for Urdu learning. These selected four Apps are shown in Fig. 1.

App1	Learn Urdu Free
App2	Urdu Qaida Activity Book Lite
App3	Urdu Qaida Kids Alif Bay Pay
App4	Best Urdu Qaida for Kids

Fig. 1. Selected Urdu learning mobile Apps [1, 15]

2 Methodology

The usability study was conducted at the primary schools of District Shaheed Benazirabad, Pakistan. A total of 232 students participated in the study. The demographic questionnaire, Fun Toolkit based post-test questionnaire and the usability tasks were developed. The questionnaires were filled by the participating children with the help of their teachers and the researchers. The following three usability tasks, which were performed by the children on each of the Apps on Tablet-PC, are shown in Fig. 2.

Task 1 Find out the Urdu alphabet letter "Pay" "پ".

Task 2 In the game, recognize the specific Urdu alphabet letter.

Task 3 Write the Urdu alphabet letter "Seen" "س".

Fig. 2. Usability tasks [1]

In the current study, two instruments from the Fun Toolkit have been used: Smileyometer and Again-Again table. In Smileyometer, five faces with meaningful labels have been given as choices of different expressions to be selected by children. These faces express children's feeling from "Awful" to "Brilliant" as shown in Fig. 3 [3].
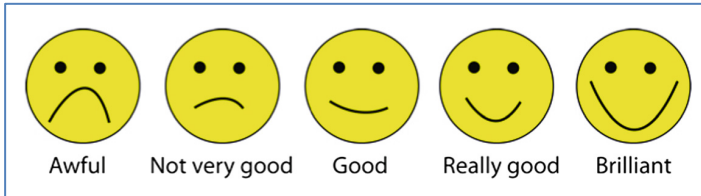

Fig. 3. The Smileyometer [3]

The Again–Again table measures engagement of children [3]. This measure is based on the knowledge that people like to do fun things again. Children use the Again-Again table to answer the question if they would like to do the activities again. Figure 4 shows the Again-Again table for all the Apps used in the current study.

	Yes	Maybe	No
App1		✓	
App2		✓	
App3	✓		
App4			✓

Fig. 4. The Again-Again table [1, 3]

3 Results and Discussions

3.1 Demographic

In this study, from the test participating children, 57.33% were boys and 42.67% were girls. Among these participating children, 46.60% students were in 5–6 years age group, 23.72% students were in 7–8 years age group and 29.68% students were in 9 years or above age group.

3.2 Usability Test

In this study, the efficiency of the Apps was measured through the task completion time. The three tasks were given to the participants one by one and time was recorded

for each task by the stop watch. The effectiveness of the Apps was measured by recording the tasks completion rate. Immediately after the usability test, the participating children's experience of fun was measured through the Fun Toolkit by getting their opinions about the Apps. Two instruments from the Fun Toolkit were used: Smileyometer and Again-Again table.

3.3   Tasks Completion Time

The average tasks completion time in seconds for all the Urdu learning mobile Apps was recorded. On average, the participants had completed all the tasks of the App1 in 11.95 seconds, of the App2 in 12.11 seconds, of the App3 in 9.41 seconds and they took 16.58 seconds for completing all the tasks of App4. The participants have taken minimum time in the App3 which was 9.41 seconds, as shown in Fig. 5. Hence, the App3 was more efficient App in terms of tasks completion time.
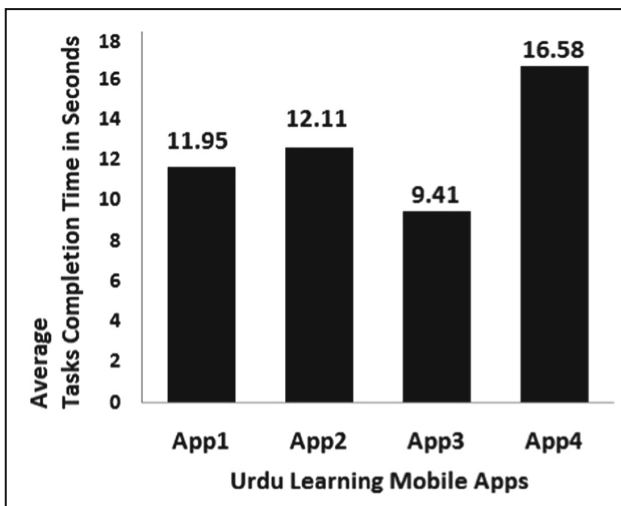

Fig. 5. Average tasks completion time for each App

Table 1 shows group differences of the Apps regarding tasks completion time. The data were analyzed using SPSS software.

Table 1. Descriptive statistics regarding tasks completion time

Urdu learning mobile Apps	Mean	Standard deviation
App1	11.95	3.25
App2	12.11	5.03
App3	9.41	3.94
App4	16.58	4.96

Table 2. One-way ANOVA

Method	Variable	Values	Significant
One-way ANOVA	Tasks_completion_time	F (3,180) = 23.201	P = 0.000

To determine whether there was a difference in the Apps in terms of tasks completion time, one-way ANOVA was used, F (3,180) = 23.201, p < 0.001 as shown in Table 2. The results show that there is statistical significant difference among the Apps in terms of tasks completion time. In Table 3, tukey post hoc pairwise statistical differences among the Apps show that there is statistically significant difference between App3 and App1 (P < 0.05), between App3 and App2 (P < 0.05), between App3 and App4 (P < 0.05). Since mean tasks completion time of the App3 was less comparing to all other Apps and the results were statistically significant, thus it can be concluded that the App3 was the most efficient App as the participants completed all the tasks in lesser time than the other Apps.

Table 3. Multiple comparisons

Tukey HSD dependent variable: Tasks_completion_time

(I) Apps	(J) Apps	Mean difference (I − J)	Std. error	Sig.	95% confidence interval	
					Lower bound	Upper bound
1	2	.00718	.90228	1.000	−2.3326	2.3470
	3	2.53902*	.88538	.024	.2431	4.8350
	4	−4.59723*	.85693	.000	−6.8194	−2.3750
2	1	−.00718	.90228	1.000	−2.3470	2.3326
	3	2.53184*	.92534	.034	.1323	4.9314
	4	−4.60440*	.89816	.000	−6.9335	−2.2753
3	1	−2.53902*	.88538	.024	−4.8350	−.2431
	2	−2.53184*	.92534	.034	−4.9314	−.1323
	4	−7.13625*	.88118	.000	−9.4213	−4.8512
4	1	4.59723*	.85693	.000	2.3750	6.8194
	2	4.60440*	.89816	.000	2.2753	6.9335
	3	7.13625*	.88118	.000	4.8512	9.4213

*The mean difference is significant at the 0.05 level.

3.4 Tasks Completion Rate

Figure 6 shows number of the participants in percentage in two categories: "Tasks Completed" and "Tasks Not Completed". Only 78.40% children have completed all the tasks for the App1 while 74.27% children have completed all the tasks for the App2. However, all the children have completed all the tasks for the App3 while only 71.59% children have completed all the tasks for the App4. Hence, the most effective App was the App3 in terms of tasks completion rate.
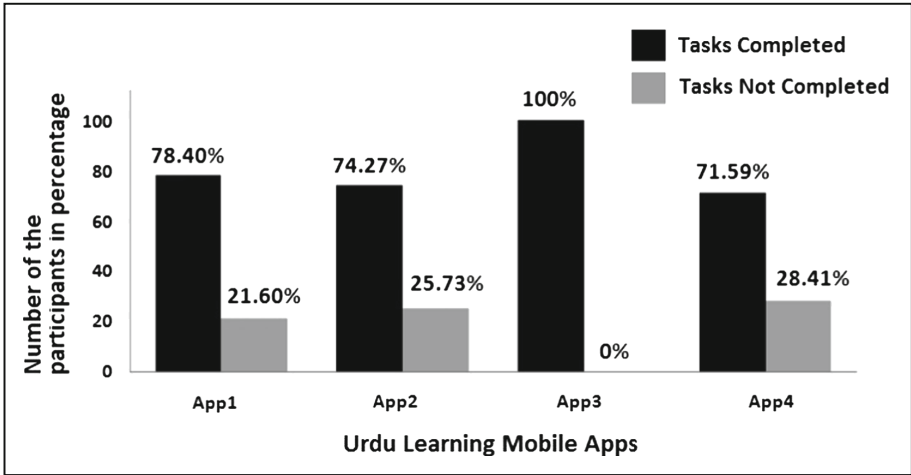
Fig. 6. Tasks completion rate of all the tasks for each App

3.5   Again-Again Table

Using Again-Again table, the participants were asked to report regarding each of the Apps: "Would you like to use this App again?" Figure 7 shows responses of the participants. For the App1, 78.22% participants' reported response was "Yes". For the App2, 81.49% participants responded as "Yes". The App3 was most rated as 95.22% of the participants rated it as "Yes". For the App4, 73.66% of the participants' response was "Yes". According to the results of the Again-Again table, most of the children wanted to use the App3 again and again.
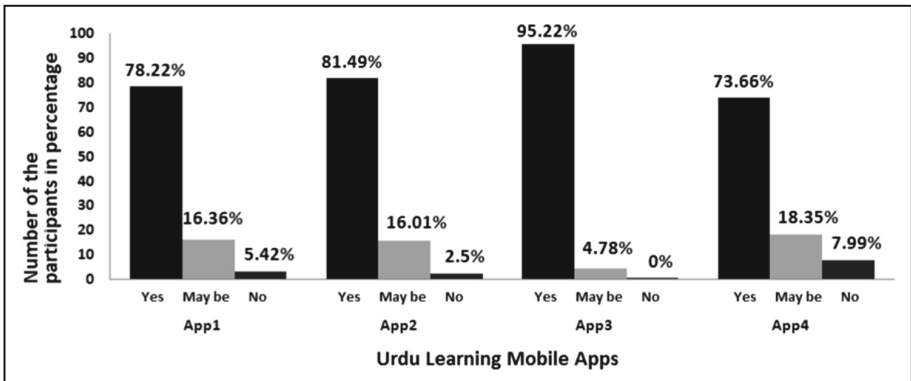

Fig. 7. Results of Again-Again table for all the Apps

The Again-Again results were coded as 3 for "Yes", 2 for "Maybe" and 1 for "No". Table 4 shows descriptive statistics of the Again-Again table responses.

Table 4. Descriptive statistics of Again-Again table

Urdu learning mobile Apps	Mean	Standard Deviation
App1	2.72	0.560
App2	2.79	0.468
App3	2.95	0.213
App4	2.66	0.618

Table 5. One-way ANOVA

Method	Variable	Values	Significant
One-way ANOVA	Again-Again table	F (3,924) = 15.319	P = 0.000

To determine whether there was a difference in the Apps in terms of the responses of Again-Again table, one-way ANOVA was used, $F_{(3,924)} = 15.319$, $p < 0.001$ as shown in Table 5. The results show that there is statistical significant difference among the Apps in terms of Again-Again table responses. Tukey post hoc pairwise statistical differences among the Apps show that there is statistically significant difference between App3 and App1 ($P < 0.05$), between App3 and App2 ($P < 0.05$), between App3 and App4 ($P < 0.05$). Since mean Again-Again table responses of the App3 was less comparing to all other Apps and the results were statistically significant; hence the App3 was the App which most of the children wanted to use again and again.

3.6 Responses of Smileyometer

Using Smileyometer, the participants were asked to report regarding each of the Apps: "How much fun was it to use the App?" Figure 8 shows Smileyometer responses of the participants. For the App1, 57.56% of the participants' response was "Brilliant" and 30.56% participants' response was "Really Good". For the App2, 54.31% of the participants rated "Brilliant" and 34% rated "Really Good". 71.45% of the participants responded as "Brilliant" and 24.44% responded as "Really Good" for the App3. Finally for the App4, 46.77% of the participants' response was "Brilliant" and 25.47% of the participants' response was "Really Good". According to the Smileyometer results, the maximum positive response was reported for the App3.
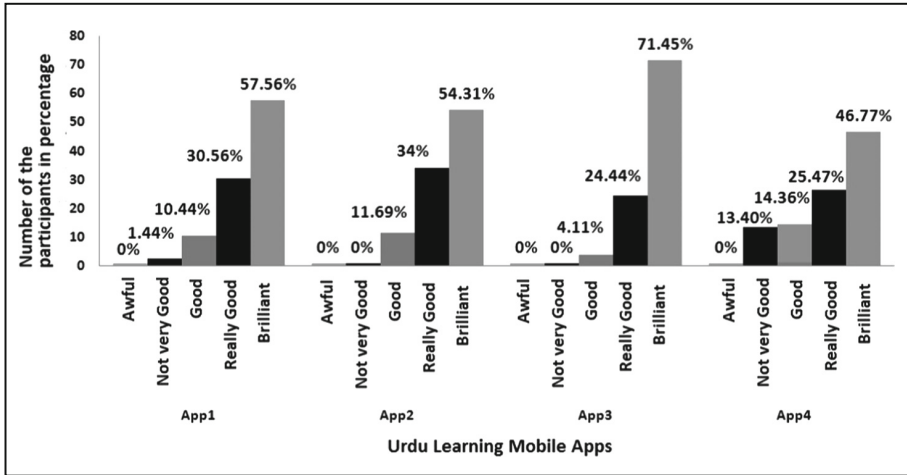
Fig. 8. Smileyometer results for each App

The Smileyometer responses were coded from 1 to 5 where 1 was used for "Awful" and 5 was used for "Brilliant". Table 6 shows descriptive statistics of the Smileyometer responses.

Table 6. Descriptive statistics of Smileyometer

Urdu learning mobile Apps	Mean	Standard deviation
App1	4.24	1.228
App2	4.43	0.692
App3	4.68	0.546
App4	4.06	1.071

Table 7. One-way ANOVA

Method	Variable	Values	Significant
One-way ANOVA	Smileyometer	F (3,924) = 18.769	P = 0.000

To determine whether there was a difference in the Apps in terms of the responses of Smileyometer, one-way ANOVA was used, F (3,924) = 18.769, $p < 0.001$ as shown in Table 7. The results show that there is statistical significant difference among the Apps in terms of Smileyometer responses. Tukey post hoc pairwise statistical differences among the Apps show that there is statistically significant difference between App3 and App1 ($P < 0.05$), between App3 and App2 ($P < 0.05$), between App3 and App4 ($P < 0.05$). Since mean Smileyometer responses of the App3 was less comparing to all other Apps and the results were statistically significant; hence the App3 was the App which most of the children reported as having most fun.

4   Conclusions

In this research four Urdu learning mobile Apps were compared from the usability perspective. According to the tasks completion time results, the App3 was the most efficient App as the participants completed all the tasks in lesser time than the other Apps. The results were statistically significant. The most effective App was the App3 in terms of tasks completion rate and the results were also statistically significant. Since mean Again-Again table responses of the App3 was less comparing to all other Apps and the results were statistically significant; hence the App3 was the App which most of the children wanted to use again and again. Since mean Smileyometer responses of the App3 was less comparing to all other Apps and the results were statistically significant; hence the App3 was the App which most of the children reported as having most fun. According to the usability perspective, the App3 is the best App. It is recommended that the App3 should be complemented with the existing traditional classroom learning in the primary schools of District Shaheed Benazirabad and other cities of the country for enhancing the learning outcomes of children. In future, more usability studies would be conducted in various schools throughout the country for getting more feedback. An experiment would also be conducted to measure the learning outcomes of the students in the primary school by introducing the App3.

References

1. Hussain, N.: Evaluating useful Urdu Qaida mobile Apps for kids, Master's thesis, Department of Information Technology, Quaid-e-Awam University of Engineering, Science and Technology, Nawabshah, Pakistan (2017)
2. Costabile, M.F., De Marsico, M., Lanzilotti, R., Plantamura, V.L., Roselli, T.: On the usability evaluation of e-learning applications. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences, pp. 1–10. IEEE (2005)
3. Read, J.C., MacFarlane, S.J., Casey, C.: Endurability, engagement and expectations: Measuring children's fun. In: Interaction Design and Children, vol. 2, pp. 1–23 (2002)
4. Read, J.C., MacFarlane, S.J.: Measuring Fun-Usability Testing for Children: Comput Fun 3. BCS HCI Group, York (2000)
5. Read, J.C., MacFarlane, S.: Using the fun toolkit and other survey methods to gather opinions in child computer interaction. In: Proceedings of the 2006 Conference on Interaction Design and Children, pp. 81–88. ACM (2006)
6. Sim, G., Read, J.C., Gregory, P., Xu, D.: From England to Uganda: children designing and evaluating serious games. Hum. Comput. Interact. **30**(3–4), 263–293 (2015)
7. Masood, M., Thigambaram, M.: The usability of mobile applications for pre-schoolers. Procedia Soc. Behav. Sciences **197**, 1818–1826 (2015). Chicago
8. Sim, G., MacFarlane, S., Horton, M.: Evaluating usability, fun and learning in educational software for children. In: Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications, pp. 1180–1187 (2005)
9. Kirci, P., Kahraman, M.O.: Game based education with android mobile devices. In: 2015 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO), pp. 1–4. IEEE (2015)

10. Hussain, Z., Slany, W., Rizvi, W.H., Riaz, A., Ramzan, U.: Measuring usability of the mobile learning app for the children. In: Zaphiris, P., Ioannou, A. (eds.) LCT 2017, Part I. LNCS, vol. 10295, pp. 353–363. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58509-3_28

11. Chandio, A.A., Hussain, Z., Vighio, M.S., Leghari, M.: Interactive learning system for primary schools using Tablet PC. In: Handbook of Research on Applied E-Learning in Engineering and Architecture Education, p. 446. IGI Global USA (2015)

12. Zaman, B., Abeele, V.V., De Grooff, D.: Measuring product liking in preschool children: an evaluation of the Smileyometer and This or That methods. Int. J. Child Comput. Interact. **1** (2), 61–70 (2013)

13. Alsumait, A., Al-Osaimi, A., AlFedaghi, H.: Use of survey techniques as usability evaluation for child e-learning programs. In: Proceedings of Conference ICL 2008, vol. 1(3), pp. 1–3 (2008)

14. Yusoff, Y.M., Landoni, M., Ruthven, I.: Assessing Fun: Young Children as Evaluators of Interactive Systems, pp. 11–18. University of Strathclyde, Glasgow (2010)

15. Google Play Store. https://play.google.com/store. Accessed 15 Aug 2017