A Measurement Model Based on Usability Metrics for Mobile Learning User Interface for Children

Rabail Tahir and Fahim Arif

Department of Computer Software Engineering, Military College of Signals (MCS),

National University of Sciences and Technology (NUST),

Islamabad, Pakistan

rabailtahir.mscs19@students.mcs.edu.pk, fahim@mcs.edu.pk

ABSTRACT

Mobile technology is becoming an integral learning tool for children. The Interface of mobile educational applications (apps) should be usable and compatible with the cognitive skills of children in order to provide an effective learning experience. Usability is a key quality attributes to measure the usefulness of application; therefore evaluating usability is a vital task. With the rapid advancement of mobile technology, usability of educational apps for children gains attention of modern researchers. This paper focuses on providing a measurement model for evaluating the interface of mobile educational apps designed for children .The paper attempts to review the existing interface design guidelines and consequently develop a measurement model. The model serves as basis for comprehensive usability evaluation consisting of guidelines, usability characteristics, goals (interface design criteria), questions, usability metrics (objective and subjective) and two evaluation instruments (task list and satisfaction questionnaire).To ensure the effectiveness and reliability of the model, it was validated by applying the proposed metrics and evaluation instruments in a usability study conducted on two android educational apps for children. Results gathered from usability testing proved that the Model is applicable for evaluation of mobile educational apps for children.

KEYWORDS

Usability evaluation; Interface design; Mobile educational applications; Goal Question metric (GQM); Guidelines

1. INTRODUCTION

Mobile technologies can provide a key support in education and help children develop new important skills. Results indicate that children are always excited to use mobile devices [1].A number of design challenges are involved in developing educational apps for children. The interface should be child friendly and also compatible with cognitive skills of children [1].

The term "mobile learning" is increasingly gaining attention of modern researchers with a new trend focusing on k-12 education. Trend suggest that preschoolers and elementary school age kids (k-5) would be using the mobile devices flawlessly first at their homes and then in the classrooms of 2015 as a normal part of growing up in this digital age. The results indicate that by the end of 2015, around 80% of word population will have access to mobile devices[2].The well designed educational apps are very effective for children learning. Current research shows that in mobile learning, interface design and attention to usability will lead to better mobile learning [3]. Therefore mobile educational applications (such as apps for math skills, vocabulary, memory games, drawing etc) are extremely useful for learning experiences of children and the user interface (UI) design of these applications is a key concern for their success and usefulness. Touch screen mobile technologies create new usability problems such as small screen, different interaction styles, navigation, etc [4] [5]. Same implies to children interaction. The cognitive and motor skills of children are different from those of adults as they are growing up [6]. Usability

and user experience are the important factors in creating successful applications. Novelty of mobile applications and their unique aspects become the key challenges in measuring the usability of mobile devices [7]. A number of evaluation methods and Model are available to evaluate the software usability. However the measurement models and frameworks specifically intended for the mobile applications are very limited such as [7] but they are not focused on educational apps for children. They lack one or the other criteria important for this area. Therefore existing Models and measurement models may not be appropriate to apply to educational apps designed for children because they may not be effective for this specific area. A number of limitations of current measures used to evaluate the mobile applications are as follows [10]:

- Do not have the ability to extend to the other domains.
- Not designed to evaluate mobile educational applications that use novel features specific to education and learning.
- Limited and inadequate usability measures for evaluation of unique aspects of educational apps for children such as pedagogic aspect, educational value, cognitive load, interaction, gestures, etc.

Unfortunately, very few clear guidelines are available on how the various definitions of usability characteristics and design criteria are related and how to evaluate the usability of educational apps for children. This paper aims to review the existing work to synthesize a set of usability guidelines for mobile educational apps for children and consequently develop a Model consisting of metrics and measurement instruments for evaluation. The next section presents a review of existing usability models and guidelines. Section 3 describes the evaluation Model followed by section 4 for usability evaluation and section 5 for results and

discussions. Finally section 6 concludes the paper.

2. RELATED WORK

2.1 Measurement Models

A number of usability measurement models are available [8] such as Metrics for Usability Standards in Computing (MUSiC; Bevan, 1995; Macleod et al., 1997) that include performance measures, Software Usability Measurement Inventory (SUMI; Kirakowski and Corbett, 1993) that provide measures of global satisfaction of five specific usability areas and McCall's model that is divided into three criteria, training, operability and effectiveness. In addition the semi-Automated Interface Designer and Evaluator (AIDE, Sears, 1995) is used for evaluating static HTML pages based on predetermined guidelines for webpage design, Goals Operators Methods and Selection rules model (GOMS; John and Kieras, 1996) describe series of methods needed to achieve specified goals for a task, and Quality in Use Integrated Measurement (QUIM; by Ahmed et al., 2006) is a consolidated model used for measuring actual use of software and identifying problems. Goal Question Metric model (GQM; by Basili et al.) is also used for developing measurement model for many areas including mobile phone apps [7] [8] [9].These models and many other however have their own limitations [7] [8].

2.2 Mobile Learning and Usability

Mobile learning is the form of learning that happens anytime and anywhere. It is supported by mobile devices and involves the mobility of learner and content, in the sense that it can be accessed from anywhere and anytime [2]. A large number of educational applications are available in market targeting young children and the increasing popularity of mobile has prompted a new wave of mobile learning in children education. According to a survey 88% of public schools in United States have policies on

acceptable student use of mobile phones [2].However this percentage is quite less in developing countries. Studies indicate that students achieved a higher percentage in math's and reading skills at suitable grade after teachers started incorporating the touch screen devices in the classrooms [2].

Traditional approaches of usability tend to be limited to metrics involving time to complete task, throughput, effort to complete task and the user's satisfaction. However for educational applications and mobile learning researchers now suggest to go beyond this by combining the specialized usability criteria (such as efficiency, reliability, consistency etc) with the pedagogical usability components including motivation, learner control, feedback and learner activity. The usability needs should be comprehended differently when it is being measured in the context of education and learning .The concept of pedagogical usability can be very helpful as a means of focusing on the relationship between usability and pedagogical design [3].

2.3 Usability Model

The definition of usability has evolved over a period of time and usability concept has been defined in multiple ways [8]. Some existing usability models include Nielsen (1993), Shneiderman (1992), Preece et al. (1994), Shackel (1991) and Constantine & Lockwood (1999). Likewise the International Organization for Standardization (ISO) has also developed a number of usability models but no one model covers all aspects of usability. According to [8] there are three major ISO standards which include: ISO 9241-11 (1998) which is most extensively used model also for mobile usability [10]. It identifies efficiency, effectiveness, and satisfaction as key attributes. ISO/IEC 9126-1 (2001) define usability as software quality attribute decomposed in five factors understandability, learnability, operability, attractiveness, and Usability compliance. ISO/IEC 9126-4 (2001) defines the concept of

quality in use and ISO/IEC 14598-1 (1999) is for measuring quality in use from the perspective of internal software quality attributes. Upon close review of literature the core concept of usability always appear to be ISO 9241-11 and rest usability characteristics depend on the type of interface being considered [10].
Therefore this paper adopts ISO 9241-11 as a baseline for this study. Along with this, quality attributes of ISO 9216-1 are selected as sub characteristics along with this baseline model [10].Therefore this study will focus on the following usability characteristics: efficiency, effectiveness, satisfaction, understandability, learnability, operability and attractiveness. For interface design both ISO 9241 and ISO 9126 are widely used. These two models are considered as the complementary definitions of usability and can be collectively used for usability evaluation of UI design [11].

2.4 Review of Guidelines from Literature

Literature review is a way of evaluating and identifying the related studies and current practices relevant to the area of interest. Many researchers such as Hornbæk et al [12] have employed literature review as basis for their research work.
The literature review for guidelines is done by reviewing research papers based on keywords "interface design", "mobile interface", "usability", "interface design for children", and "educational apps". A total of 27 research papers were selected and studied for synthesizing the guidelines for interface design of mobile educational apps for children. Analysis has been made to select only the relevant guidelines, identify and combine the duplicate guidelines, resolve conflicting guidelines and rephrase the unclear guidelines. For a research paper to be selected for review, the study should be related to either of the following areas: the usability goals for interface design of touch screen mobile devices in general, interfaces for children or usability goals for educational applications, instead of a broad concept of usability. This

criterion was set in order to obtain UI design criteria focused on mobile educational apps for children.

Some of these studies include the research of Tafresh et al [1] who explored the design requirements in order to develop a user friendly interface for children and proposed some design techniques that can be employed to meet the requirements. Florence et al [4] research focused to evaluate different UI designs and input methods for touch screen mobile phones. Aziz, N.A.A. et al [6] investigated the interaction of children aged two to twelve years with gestures such as tap, slide, drag-and-drop , spin/rotate ,pinch, flick and spread on a range of applications on tablet or iPad and the interface design of touch screen applications. The paper focuses on the point that for applications to be effective they need to be age-and-gesture-appropriate .Mary Ann et al [10] proposed a list of practical guidelines based on usability concerns for interface design of mobile device, which should be considered carefully when designing a mobile interface. Nilsson et al [13] in his paper presented a collection of user interface design patterns for the mobile applications. Tsai et al [14] carried out face-to-face interviews to explore the perceptions about smartphone interfaces. Sharma et al [15] strongly supported the need for an age based user interface. According to the research kid's user interfaces should only contain educational widgets, games and music. According to Heather Nam [16] suggested that new user interface design conventions should not be developed only because the audience includes children. Instead, interactions to standard design conventions should be limited, only using the ones that are easiest for children usage. Gilutz and Nielsen [17] investigated a variety of user experiences on the websites and discovered the similarities and differences between the response of adults and children.

According to [18] a usability study of kids found that children view ads as content, and tend to click them accordingly. Children prefer colorful designs yet demand simple navigation and text.

Asmaa Alsumait et al [19] introduced the Heuristic Evaluation for Child E-learning applications (HECE).HECE is a set of heuristics for child e-learning applications. Petri Nokelainen [20] in his paper presented pedagogical usability criteria for evaluating digital learning material. The following components, Learner control, Feedback, Added value, Learner activity, Motivation, Cooperative/Collaborative learning, Applicability and Goal orientation were focused in his work. Gavin Sim et al [21] reported the findings of an analysis of the relationship between fun, usability and learning in the educational software designed for the children. Walayat Hussain et al [22] has emphasized on how to make a webpage more usable in terms of readability for different age groups. The paper has focused on eight readability factors that are white space, graphics, line spacing, font style, text width, color contrast, headings, font size, and animation. Lisa Meloncon et al [23] presented a set of guidelines to aid the design process to develop educational websites for children.

3 MEASURMENT MODEL FOR MOBILE LEARNING USER INTERFACE DESIGN FOR CHILDREN

This section presents the evaluation model for measuring the usability of interface design of mobile educational applications for children. Overall process and model is shown in figure.1.
The Model consists mainly of three phases. The first phase explores the literature review and presents the usability characteristics and guidelines for interface design of educational apps for children.
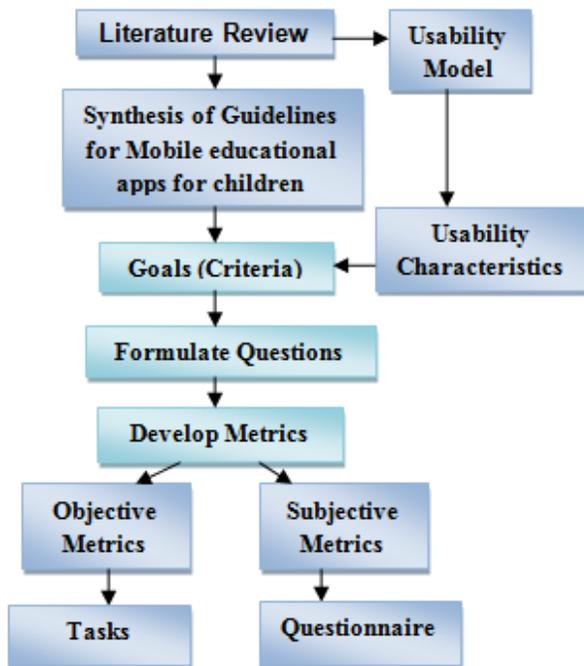
Figure 1. Measurement Model based on Usability Metrics

In the second phase Goal Question Metric (GQM) approach developed by Basil et al. [9] is used to develop metrics for usability evaluation of mobile educational apps for children. Originally GQM was employed to define and evaluate goals for a particular project or environment but its use has now been extended to larger perspectives and is adaptable to different environments and organizations, as confirmed by companies such as (NASA, Siemens, and Philips). It has now become a de facto standard for defining measurement models [7].GQM has three levels. The first level is the conceptual level in which goals are identified based on the guidelines created in the last section. Next is operational level in which questions are formulated to assess each goal and the final level is quantitative level in which a set of metrics are developed to provide information in order to answer the questions formulated in the previous level[9]. The guidelines from first phase will be used for metric development in GQM .The

developed metrics can be both objective and subjective. Interface design is evaluated using both metrics i.e. task-sensitive metric which are objective and make sure that user tasks direct the semantics of user interface design and task-independent metric which are subjective measures and make sure that interface design is visually pleasing for users [8].

In the last phase the metrics are separated into objective and subjective metrics which are then used to develop two measurement instrument task list and questionnaire respectively. These instruments can be used for usability evaluation of mobile educational apps for children in order to obtain quantitative and qualitative data.

3.1 Synthesized Guidelines for User Interface Design of Mobile Learning for Children

The previous section covered the review of usability models also stating the models chosen as a foundation for the model in this research study. The usability characteristics related to the employed usability models were also presented. Moreover review of guidelines from literature was also discussed which was the basis for obtaining the guidelines for UI design of mobile educational apps for children. These were the prerequisite for understanding and defining the evaluation Model presented in this section.

The next step was to synthesize guidelines for mobile educational apps for children which were obtained from the review of guidelines from literature as described earlier. The synthesized guidelines focus on the interface design of mobile educational apps designed for children, therefore the quality characteristics that do not focus on interface design were not included such as the network throughput, memory load, Display load, application size and battery usage etc. The selected guidelines resulted in 17 UI design criteria and 27 sub criteria based on review of guidelines from literature. Table 1 describes the guidelines along with criteria and sub criteria

Table 1. Synthesized guidelines for user interface design of mobile learning for children

UI Design Criteria	Sub criteria	Guidelines
Input/output		Ease to input
		Ease to understand output
Cognitive Load	Recognition	Recognition rather than recall
	Terminology	Use of appropriate language
	Content/ concept	Use of appropriate content
		Familiar concepts
Multimedia usage	Sound/Audio	Use sound/audio where appropriate
	Animation/ images	Use of images and animations match with children skills
	Text	Understandable text
Customization/ Personalization		Allow for customization
		Allow for personalization
Screen design	Aesthetic	Simple,attractive and organized design
	Colors	Use bright colors for children
	Font style/size	Use appropriate font style and size
	Menu	Provide proper menu for touch screen
	Buttons	Provide colorful and animated buttons
	Icons	Icons must be relevant to information they present
Layout		Clear and consistent screen layout
Learning Potential	Ease to learn	Ease of learning
	Educational value	Appropriate educational content
	Suitability	Suitability for all users and learner control.
	Learning activities	Learning approach/opportunities
Feedback		Provide appropriate feedback
	Pedagogic feedback	Provide pedagogic feedback for answers.
		Responsive to input
		Audio instructions
User Control		Provide appropriate controls e.g. save, reset, exit etc
Navigation/ Orientation		Ease of orientation
	Ease to navigate	Ease of navigation
		Clear and consistent navigation
	Navigation keys	Provide clear navigation buttons.
	Main menu/ start screen	Provide main menu for navigation
		Clear main menu/start page link
	Hierarchal menus	Hierarchal menu for easy navigation
	Scrolling	Scrolling may be difficult for children
		Allow scroll and view ,when a lot of information is present
Help/support		Provide sufficient help
	Tutorials	Provide tutorials
	Hints/clues	Provide task related clues and hint
Error		Provide short error messages
Interactivity	Gestures	Use of appropriate Gestures
	Interaction	Ease of interaction
Effort		Amount of task effort
Time required		Loading application
		Time to respond
		Time to complete task
Engagement	Motivation to learn	Endeavor engagement
		Provide interesting rewards.
Readability		Ease of readability
		Provide appropriate text size, spacing etc

3.2 Goal, Question, Metric (GQM) for Measurement Model

The UI design criteria developed from synthesized guidelines act as goals of GQM model. The usability characteristics from ISO 9241-11 and ISO 9126 are related to goals using [8] and [10] [11] [12].The goal represent the overall aim of evaluation presenting the UI design criteria of educational apps for children.
The goals and guidelines were used to formulate a list of questions to assess each one of them. We ensured that the questions we created can be answered.
Next step is to develop a set of metrics in order to collect data to answer each question in a quantitative way. The developed metrics contains both objective and subjective metrics as not all the created questions could be objectively answered, providing only objective metrics. Therefore some questions will be answered subjectively using a questionnaire to assess user satisfaction. These metrics can be useful for evaluating both objective and subjective usability of mobile educational apps for children. The resulted goals, questions and metrics for usability evaluation of mobile educational applications for children are shown in Table 2. The objective metrics are highlighted in Table 2 (blue in metric column) and the remaining are subjective metrics. These objective and subjective metric are used in the next phase to develop evaluation instruments that are tasks and questionnaire respectively.

Table 2. Usability characteristics, Goals, Questions and Metrics

Usability Characteristics	Goals (UI design criteria)	Questions	Metrics
Effectiveness	Interactivity	Is it easy to interact with the UI? Does UI provide interaction like collaboration or sharing?	**Number of mistakes during interaction** **Number of collaboration /sharing options**
		Are gestures easy to use for children?	**Number of mistakes in using gestures**
	Navigation /Orientation	Is it easy for children to navigate across the UI? Does UI provide clear and understandable navigation keys? Does UI indicate easy scrolling if a lot of information is present?	**Number of mistakes during navigation** Rating scale for navigation
		Does UI provide easy main menu for navigation? Is the main menu /home page icon effective for children?	Rating scale for main menu **Success/Failure rate to use main menu**
		Is the screen orientation of UI effective for children?	Rating scale for screen orientation
	Multimedia usage Feedback	Is multimedia usage of UI appropriate for children? Does UI provide appropriate feedback? Is Interface of application responsive to input? Does UI provide pedagogic feedback for self assessment?	Rating for multimedia usage Rating scale for appropriate feedback **Number of pedagogic feedback** **Success rate for understanding pedagogic feedback** Rating scale for pedagogic feedback
		Does UI provide a visual display to show the loading process? Does application provide audio instructions?	Rating scale for loading application **Number of times voice/audio instructions provided in a task** Rating scale for voice instruction
Understand	Input/output	Does Interface provide easy ways of input for children?	**Number of mistakes to enter**

-ability		Does UI provide easy to use keypad? Is it easy to understand the output for children?	**/give input** Ease to use virtual keypad Rating scale for ease to understand output
Efficiency	Time required	How much time is taken by the application to load? Is time taken by the UI to respond appropriate? Does UI provide appropriate time for user to respond? How much time is taken for completing a given task?	**Time taken to load application** Rating scale for time to respond **Time taken to complete task**
	Effort	How much effort is required for task completion?	Rating scale for task effort
Learnability	Help/ support	Does UI provide appropriate and sufficient help? Does UI provide clear and understandable help icon for finding help? Does UI provide brief and useful tutorial to understand task/activity? Does UI provide useful task related clues/ hints for children?	Rating for usefulness of help Rating scale for finding help **Success/failure for finding help** Rating scale for tutorials **Number of task related clues/hints**
	Cognitive load	Are children capable of recognizing the functions and their actions? Can children easily recognize an icon/link/button? Is the terminology/language used appropriate for children? Does UI provide appropriate content/information for children? Does UI use familiar concept matching children mental model?	**Number of icons/buttons not recognized in first attempt** Rating scale for appropriate language Rating scale for appropriate content
	Learning potential	Is application easy to learn for children?	Rating scale for ease of learning **Number of mistakes before learning to use** **Time to learn a task**
		Does educational content fit with age and curriculum of children? Is appropriate learner control provided to users? Does UI provide different difficulty levels for equal learning experiences for all users' novice or expert? Does app provides useful and interesting learning activities for children? Does application provide appropriate progress report/evaluation result for assessment of performance in a given activity?	Rating for educational value Rating scale for suitability for all users **Number of difficulty levels for practice** Rating scale for learning activities **Success/failure rate for performance assessment** Rating scale for performance assessment
Operability	Customiz-ability/ Personaliza-tion	Does UI allow for personalization? Does UI allow for customization?	**Number of options for personalization/ customization** **Success/failure for using personalization/ customization options**
	Error tolerance User control Readability	Does Interface provide short errors messages? Does application provide appropriate controls? Does application provide easy readability for children? Is the text size appropriate for child?	Rating scale for error messages **Success rate for using controls** Ease of readability Satisfaction with text
Satisfaction	Engagement	Is the Interface engaging for children? Does UI provide exciting rewards to engage children?	Rating scale for engagement
Attractiveness	Screen layout Screen Design	Is the screen layout clear and consistent? Is screen design attractive for children? Are children happy with the interface of application?	Rating scale for screen layout Rating scale for attractive screen design Rating scale for interface color Rating scale for icons/ buttons

3.3 Measurement Instruments

The subjective and objective metrics from the previous phase are used to develop two evaluation instruments user satisfaction questionnaire and task list respectively as shown in table 3 and table 4.

Table 3. Questionnaire

User Satisfaction Questionnaire

1. I found it easy to understand this application.
2. The app provides easy to use touch screen input or virtual keypad.
3. The application is too slow I had to wait for response to continue
4. The app took a lot of time for loading.
5. The app provides a visual display to show the loading process.
6. The app gives feedback on whether my answer is correct or wrong
7. The application provides useful voice instructions
8. The application does not provide appropriate feedback for my actions.
9. I was comfortable with the screen orientation of application.
10. The main menu of application is confusing
11. The app provide clear and understandable navigation keys such as back/next buttons to move to previous/ next screen
12. The application provides useful help information
13. It was difficult to find help
14. The application provides useful tutorials that explain how to perform a task/activity
15. It was difficult to understand the language used in the application
16. The topics/concept and information was understandable
17. I need to remember a lot of information throughout several actions to perform a task.
18. It was easy to complete the tasks without much effort.
19. It is difficult to learn to use the application.
20. The educational content matches with my course curriculum
21. The application provides different difficulty levels that I could easily change according to my choice.
22. The learning approach and activities in app were interesting and I learned from them
23. The application provides a progress report/result for my performance in every activity
24. The application gives error messages that clearly tell me how to fix problems
25. It was easy to read the text in this application
26. The text size used in this application is too small
27. I like the animation and images used in this application
28. The music and sound effects used in the app were disturbing
29. The organization of information on the app screens is not clear and consistent
30. It is easy to find the information I needed
31. I find the design of application attractive
32. The colors used in this application are not attractive
33. The icons and buttons used are attractive and recognizable
34. The application gives interesting rewards on my performance
35. Overall I enjoyed using the app.

Table 4. Task list

Task List

- Check for interactivity
  a. Check of user interaction with application
  b. Check of availability of communication tools
  c. Check of usage of gestures
- Navigation activity
  a. Check of main menu presence
  b. Check for scrolling
  c. Check for hierarchal menu
  d. Check for navigation keys
- Check for adequacy of feedback
  a. Response to input
  b. Audio instructions
- Check for time
  a. Loading application
  b. Task
- Check input/output availability
  a. Virtual keyboard
- Check for adequacy of Help
  a. Task related clues
  b. Tutorials
  c. Help icon
- Check for cognitive load
  a. Identify a link or icon usage
  b. Check for suitability of language
  c. Check for suitability of content
- Check for learning potential
  a. Check for presence of alternative learning options
  b. Check of assessment / result availability
- Check for personalization/customization
  a. Check for availability of settings option
- Check for short error messages
- Check for user controls
- Check suitability of reading

Tasks are developed using objective metrics in order to collect objective data and questionnaire is developed using subjective metrics to obtain results for subjective measures assessing satisfaction ratings with interface design. These developed instruments can be used in usability evaluation of educational apps for children by implementing tasks performance for objective measures and using questionnaire to assess the subjective measures.

When the satisfaction questionnaire is used, participants are asked to rate the 35 items related to the interface design of educational apps for children with a 5 point Likert scale that range from 1 for Not at all true,2 for Not very true,3 for Somewhat True , 4 for True to 5 For Very True.

This complete model offers a comprehensive structure for evaluating usability. It describes usability characteristics and how these are linked to UI design criteria. The metrics for accessing each criteria and the evaluation instrument for obtaining data for each metric. Hence this can be useful for obtaining quantitative and qualitative data for usability evaluation.

4 USABILITY STUDY

A usability study was carried out to ensure that the model is reliable and effective for evaluating the usability of mobile educational applications for children. Usability testing was conducted to test whether the metrics and evaluation instrument (User satisfaction questionnaire and task list) developed in the model can be used to collect the data for usability evaluation. Therefore both objective and subjective metrics were employed for this usability study. To validate the model this study used two educational apps Math Open and Barnyard Math designed for math skills of elementary school age children. These applications were installed in Sony Ericsson xperia arc s smart phone which was used for usability testing. For usability evaluation the

subjective data was collected using the 5 point Likert scale user satisfaction questionnaire developed in the model and objective data was collected through usability testing using task list to prepare tasks for each app.

A total of 10 children participated in usability testing. As suggested by Nigel Bevan [24] a minimum number of eight to ten participants are generally required in order to make reliable estimates to uncover the usability problems of an interface .The participants included a mix of boys and girls with different level of expertise from novice to experience. All children were of elementary school age i.e. 6 to 10 years and were recruited from elementary school.

Figure 2. Usability testing with children

The study was conducted in a quiet and comfortable room in school see figure 2. Before usability testing a consent form was signed by the parents for letting their children participate in the study. The parents were ensured that their personal information will be kept confidential. During usability test each participant was asked to experience both educational apps. All participants were required to complete 11 tasks with each of the two apps. The participants were given time to freely explore the applications before completing the tasks. For usability testing two evaluators were present in the room with one user at time. A teacher was also present during the test so that children may not feel uncomfortable with strangers. The participants were given a brief introduction in the start regarding the purpose of the study.

5 RESULTS AND DISCUSSIONS

The results for both subjective and objective metrics are presented separately and comparison of results is also presented for both apps to check significant differences in usability of the two educational apps.

5.1 Objective Usability Results

The data for objective measures were collected during usability testing and we summarized the data for each of the 21 objective metrics from the frame work. The mean score for each measure is presented in Table 5 for both apps.

Table 5. Results for Objective Metrics

Objective Metrics	Barnyard Math Mean	Math Open Mean
O1-Number of mistakes during interaction	2.2	3.6
O2-Number of collaboration/sharing option	0	0
O3-Number of mistakes in using gestures	0.6	0.1
O4-Number of mistakes during navigation	1.2	0.8
O5- Failure rate to use main menu	0	0.2
O6-Number of pedagogic feedback	2	4
O7- Failure rate for understanding pedagogic feedback	0.4	0.2
O8-Number of times voice instructions provided in a task	0	0
O9-Number of mistakes to enter /give input	0.1	1.7
O10-Time taken to load application	0.13	0.26
O11-Time taken to complete task	0.830	1.18
O12- Failure rate for finding help	0	0
O13-Number of task related clues/hints	0	0
O14-Number of icons/buttons not recognized in first attempt	0.1	0.8
O15-Number of mistakes in learning to use	0.3	1.2
O16-Time to learn a given task	0.38	0.4
O17-Number of difficulty levels for practice	5	14
O18- Failure rate for performance assessment/result	0.4	0.2
O19-Number of options for personalization/ customization	1	3
O20-Failure for using personalization/ customization options	0.1	0.3
O21- Failure rate for using controls	0.4	0.4

In Table 5 labels O1-O21 were used to represent the objective metrics. For comparative analysis of the two educational apps, score from table 5 is considered for analyzing the results of objective metrics of usability. The comparative analysis is carried out to determine whether one educational app has better usability than the other. Moreover this was useful to determine whether the model used was effective for usability evaluation and comparative analysis of different educational apps for children.

The results indicate that math open app has higher failure and number of mistakes for all objective measures in figure 3, except for navigation, pedagogic feedback and performance assessment/results. Therefore it is
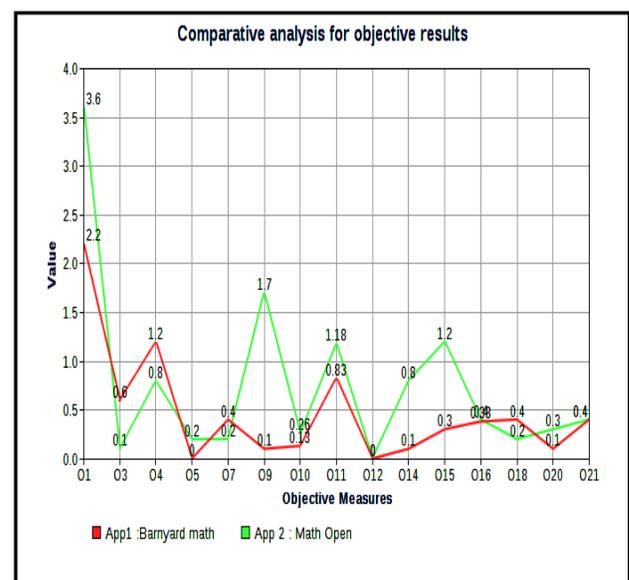

Figure 3.Objective Usability of Educational Apps

evident that barnyard math has better objective usability than Math open.

5.2 Subjective Usability Results

The data for subjective measures was collected through 5 point Likert scale user satisfaction questionnaire presented in the model .The questionnaire was filled by the participants after performing tasks, at the end of test session for each app. The analysis of data from 5 point Likert scale satisfaction questionnaire was done

according to the statistical procedure described by Boone et al [25]. For each subjective metric the questions from satisfaction questionnaire were matched with the metric (for example question number 5 and 8 relate to subjective measure "Appropriate Feedback") and mean score for these questions was recorded for each of these subjective metrics. A higher score indicates greater satisfaction level for users. Table 6 presents the results of subjective measures.The labels "S1 to S30" are used to represent each subjective metric.

Table 6. Results for Subjective Metrics

Subjective Metrics	Question No	Barnyard math		Math open	
		Mean	%	Mean	%
S1-Rating scale for multimedia usage	27,28	4.5	90%	4	80%
S2-Rating scale for appropriate feedback	8,5	2.3	46%	4.2	84%
S3-Rating scale for screen orientation	9	4.9	98%	3.8	76%
S4-Rating scale for navigation	11	1.1	22%	1.7	34%
S5-Rating scale for main menu	10	4.8	96%	3.5	70%
S6-Rating scale for pedagogic feedback	6	2.9	58%	4.3	86%
S7-Rating scale for voice instructions	7	1.4	28%	1.1	22%
S8-Ease to use virtual keyboard	2	4.5	90%	3.1	62%
S9-Rating scale for ease to understand output	1	4.3	86%	2.9	58%
S10-Rating scale for loading application	4	4.9	98%	4	80%
S11-Rating scale for time to respond	3	4.7	94%	4.5	90%
S12-Rating scale for task effort	17,18	4.35	87%	4.2	84%
S13-Rating scale for finding help	13	1.5	30%	1	20%
S14-Rating scale for usefulness of help	12	1.1	22%	1	20%
S15-Rating scale for tutorials	14	1.1	22%	1.1	22%
S16-Rating scale for appropriate language	15	4.2	84%	4.1	82%
S17-Rating scale for appropriate content	16	3.9	78%	4	80%
S18-Rating scale for ease of learning	19	4.5	90%	3.3	66%
S19-Rating scale for educational value	20	3.9	78%	4.8	96%
S20-Rating scale for suitability for all users	21	4.8	96%	4.3	86%
S21-Rating scale for learning activities	22	3.9	78%	3.5	70%
S22-Rating scale for performance assessment	23	3.1	62%	4.1	82%
S23-Rating scale for error messages	24	1	20%	1.6	32%
S24-Ease of readability	25	4.9	98%	1.6	32%
S25-Satisfaction with text	26	5	100%	2	40%
S26-Rating scale for engagement	34,35	3.25	65%	3.05	61%
S27-Rating scale for screen layout	29,30	3.95	79%	3.65	73%
S28-Rating scale for attractive screen design	31	4.1	82%	2.7	54%
S29-Rating scale for interface color	32	4.9	98%	2.6	52%
S30-Rating scale for icons and buttons	33	4.2	84%	3.3	66%

For comparative analysis, results for subjective measures are presented diagrammatically in Figure 4 and 5 for mean and percentage respectively. The results show that for most of the metrics barnyard math show better

subjective usability than Math open expect for feedback, educational value, performance assessment and pedagogic feedback where math open showed better results. It means that participants were more satisfied with barnyard

math app and had good experience using it. However both apps showed poor usability regarding help, tutorials, navigation, voice instructions and error messages. Furthermore the participants were unsatisfied with the interface color, text size, readability and virtual keyboard of Math open. These UI design attributes need to be improved. The results indicate that the user satisfaction questionnaire developed in the model is reliable and effective for collecting subjective data for evaluating the usability of mobile educational apps for children.
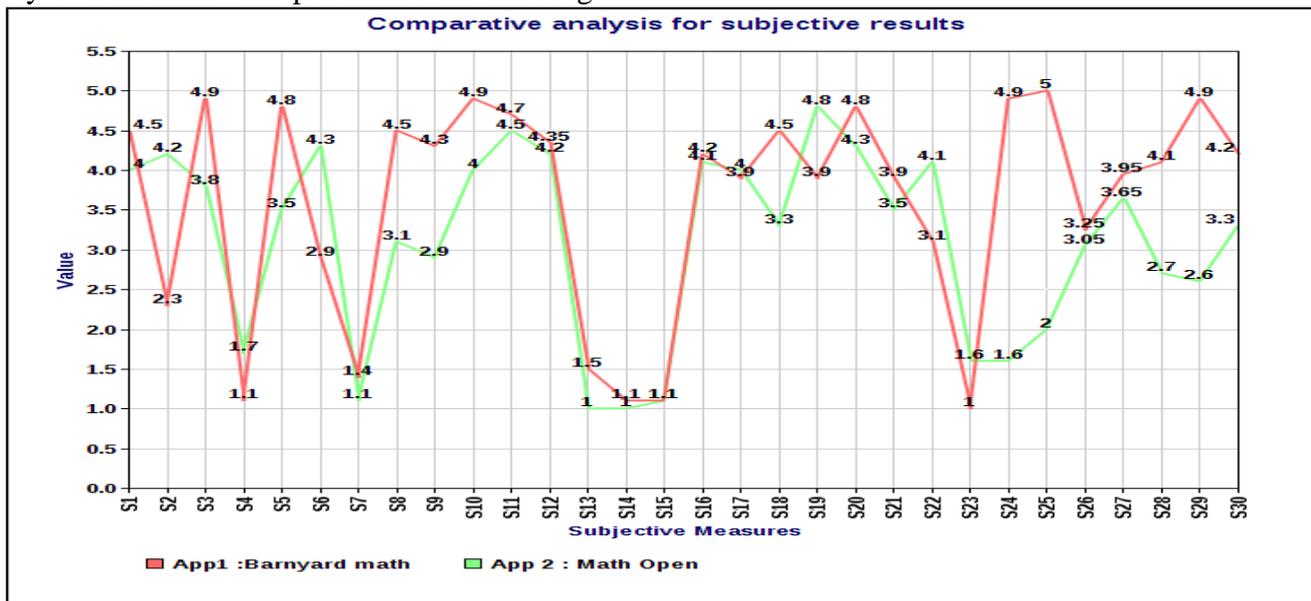

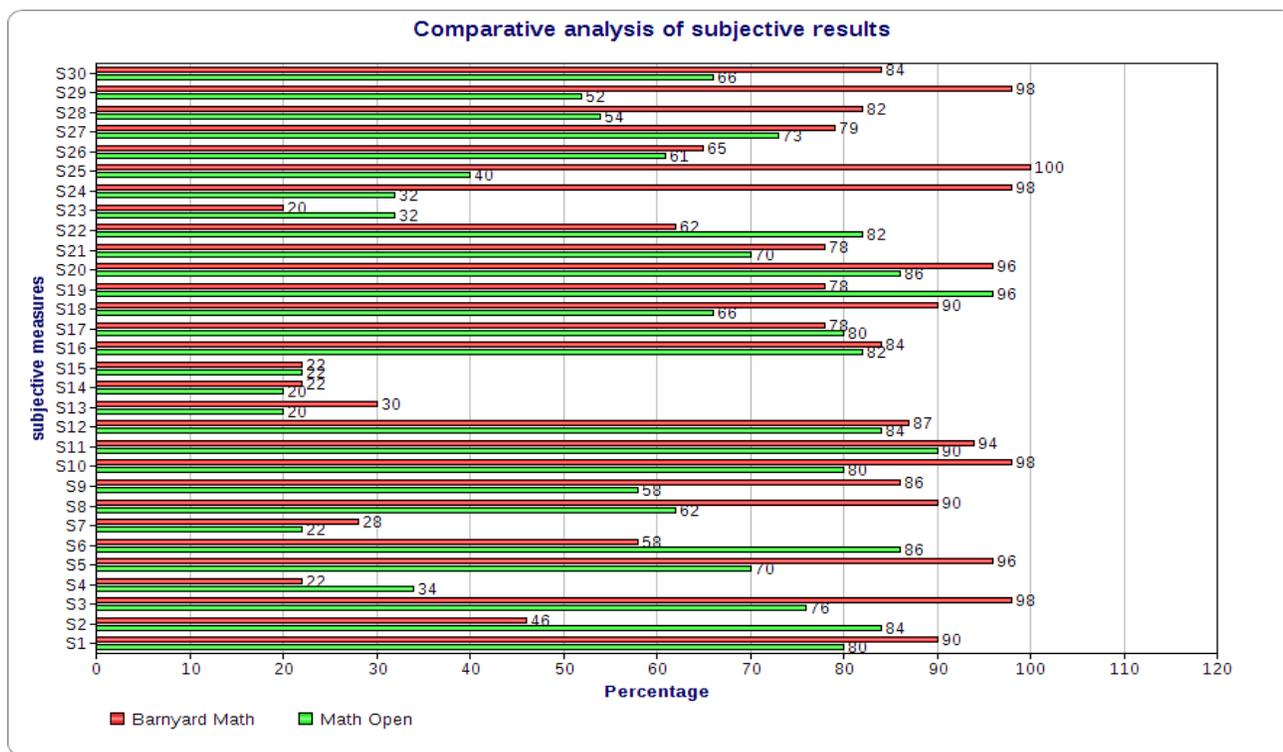Figure 4. Subjective Usability of Educational Apps

Figure 5. Subjective results showing percentage of satisfaction for each measure

28

The overall analysis shows that both subjective and objective results correlate. This relatively close correspondence between the results for subjective and objective measures indicate that the proposed subjective and objective metrics and evaluation instruments themselves are appropriate for use in evaluating the usability of mobile educational apps for children. The results also showed that the model is not only useful for evaluating usability and comparison of different application but also helpful to uncover usability issues and highlight the UI design areas for suggested improvements. Thus it is evident from results that the model proposed in this research is effective and reliable. However more experiments and usability studies should be conducted in order to validate the model with more educational apps and large sample of participants.

6 CONCLUSION

In this paper we have proposed a measurement model that is specific for use with the interface design of mobile educational applications for children. The paper reviews the current practices in usability and measurement models. A review of existing guidelines is carried out to develop usability guidelines for interface design of children's mobile educational apps.

The model provides a comprehensive structure for evaluating the usability. At the base level it presents the usability characteristics and the UI design criteria for educational apps for children and how these are related. This serves as the foundation of model presenting goals of evaluation. Then a list of objective and subjective metrics are developed to assess each goal (UI design criteria).Finally two evaluation instrument task list and user satisfaction questionnaire are developed to collect objective and subjective data for complete usability evaluation.

The paper provides a starting point for performing usability evaluation and will be helpful for evaluators and developers by serving as a guideline for evaluating the usability of educational apps. The validation of the model is done by implementing it in a usability study. Usability testing was carried out with two educational apps designed for elementary school age children for the purpose of validating the model. The main purpose of usability study was to determine whether the model is effective to collect subjective and objective data for usability, analyze and compare the apps, provide results to uncover the usability issue and limitation with regard to the UI design and highlighting the areas of improvement. The results of this study explain that the model is useful for evaluating the usability of mobile educational applications for children.

The paper highlights some of the directions for future work. The model can be generalized to be employed in different methods for usability evaluation (such as expert evaluation, inquiry and usability testing) to identify usability issues in educational apps for children in order to improve them. In addition further studies should be carried to check the effectiveness of this model with different devices and operating systems.

The rapid changes in mobile technology and a large number of educational apps being developed may cause the interface design criteria (goals) and metrics presented in this paper to be updated in future in order to match the needs of changing technology. The model can be modified based on new design guidelines. Therefore goals, questions and also metrics can be added or deleted. A new measure can be included in the model by developing a new goal or a new question. Thus the developed tasks and questionnaire can also be updated accordingly. The goals presented in this paper only focus on the interface design. This work can be extended to include other features related to hardware and software usability.

The guidelines and metrics presented in the model can also be used to develop more evaluation instruments such as checklist for evaluating usability.

Furthermore future work should also focus on expanding the validation of proposed model so that usability evaluators can employ it with confidence for evaluating the UI design of mobile educational apps for children.

7 REFERENCES

[1] Tafreshi, Fakhteh Soltani, and Taghi Miri,"User Interface Design and Evaluation for Mobile Educational Games for children." EUROSIS, MESM'2010, December 1-3, 2010, 134-138.

[2] Chou, Chientzu Candace, Lanise Block, and Renee Jesness. "A case study of mobile learning pilot project in K-12 schools." Journal of Educational Technology Development and Exchange 5.2, 2012, 11-26.

[3] Eyal Eshed, On Designing Mobile Education Apps, citizentekk, January 2014.

[4] Balagtas-Fernandez, Florence, Jenny Forrai, and Heinrich Hussmann. "Evaluation of user interface design and input methods for applications on mobile touch screen devices." Human-Computer Interaction, INTERACT 2009. Springer Berlin Heidelberg, 243-246, 2009.

[5] Huang, Kuo-Ying. "Challenges in human-computer interaction design for mobile devices." Proceedings of the World Congress on Engineering and Computer Science, Vol. 1, pp.236-241. 2009.

[6] Aziz, Nor Azah Abdul. "Children's Interaction with Tablet Applications: Gestures and Interface Design." International Journal of Computer and Information Technology (IJCIT) (ISSN: 2279 – 0764) Volume 02– Issue 03, May 2013, 447-450.

[7] Hussain, Azham, and Maria Kutar. "Usability metric Model for mobile phone application." PGNet, ISBN, 2009, 978-1.

[8] Seffah, Ahmed, Mohammad Donyaee, Rex B. Kline, and Harkirat K. Padda. "Usability measurement and metrics: A consolidated model." Software Quality Journal 14.2, 159-178, 2006.

[9] Caldiera, V. R. B. G., and H. Dieter Rombach. "The goal question metric approach." Encyclopedia of software engineering 2.1994, 528-532.

[10] Kunjachan, Mary Ann Chiramattel. "Evaluation of Usability on Mobile User Interface." University of Washington, Bothell, 2011.

[11] Bevan, Nigel. "International standards for HCI and usability." International journal of human-computer studies 55.4, 533-552, 2001.

[12] Hornbæk, Kasper."Current Practice in Measuring Usability: Challenges to Usability Studies and Research", International Journal of Human-Computer Studies, Volume 64, Issue 2, Pages 79-102, February 2006.

[13] Tafreshi, Nilsson, Erik G. "Design patterns for user interface for mobile applications."Advances in Engineering Software 40.12, 2009, 1318-1328.

[14] Tsai, Chih-Yung, and Chun Shih. "An analysis of preferences in smartphone interface design." Scientific Research and Essays 6.20, 2011, 4195-4206.

[15] Sharma, S., Singh, P., Sharma, R., Aditya, M, "Age Based User Interface in Mobile Operating System", International Journal of Computer Science, and Applications (IJCSEA) 2,2 :177-184, April 2012.

[16] Heather Nam, "Designing user experiences for children", UX design, May 17, 2010.

[17] Nielsen, Jakob, and Shuli Gilutz. "Usability of websites for children: 70 design guidelines based on usability studies with kids." Nielsen Norman Group, 2002.

[18] Nielsen, Jakob. "Kids' corner: website usability for children." Jakob Nielsen's Alertbox, 2002.

[19] Alsumait, Asmaa, and Asma Al-Osaimi. "Usability heuristics evaluation for child e-learning applications." Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services. pp. 425-430. ACM, 2009.

[20] Nokelainen, Petri. An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. Educational Technology & Society, 9 (2), 178-197, 2006.

[21] Sim, Gavin, Stuart MacFarlane, and Matthew Horton. "Evaluating usability, fun and learning in educational software for children." World Conference on Educational Multimedia, Hypermedia and Telecommunications. Vol. 2005. No. 1. pp. 1180-1187. 2005.

[22] Walayat Hussain, Osama Sohaib, Atiq Ahmed and M. Qasim Khan "Web Readability Factors Affecting Users of All Ages" Australian Journal of Basic and Applied Sciences, 5(11): 972-977, 2011.

[23] Meloncon, Lisa, Erin Haynes, Megan Varelmann, and Lisa Groh."Building a playground: General guidelines for creating educational Web sites for children." Technical Communication 57.4, 2010, 398-415.

[24] Nigel Bevan, "Practical Issues in Usability Measurement", Interactions, 13(6): 42-43, 2006

[25] Boone, Harry N., and Deborah A. Boone. "Analyzing Likert data." Journal of Extension 50.2, 2012, 1-5