Presentation Design of E-Assessments for Blind Users Using Touch Screen Devices

Conference Paper · June 2013

DOI: 10.1109/ECAI.2013.6636214

Mohammed Fakrudeen
Anglia Ruskin University

Maaruf Ali
University of Suffolk

Presentation Design of E-Assessments for Blind Users Using Touch Screen Devices

Mohammed Fakrudeen[*], Sufian Yousef
[*]Dept. of Computing and Technology
Anglia Ruskin University
Chelmsford, United Kingdom
m.fakrudeen@uoh.edu.sasufian.yousef@anglia.ac.uk

Abdelrahman H. Hussein, Maaruf Ali
[*]College of Computer Sciences & Software Eng.
University of Ha'il
Ha'il, Kingdom of Saudi Arabia
ar.hussein@uoh.edu.sa
maaruf@ieee.org

Abstract **–The Web Content Accessibility Guidelines (WCAG) have been used by web developers to develop accessible web pages. However, touch screen accessibility varies with web site accessibility. This research is dedicated to the improvement of accessibility of e-assessment in mobile touch screen for blind users. We chose IMS' Question and Test Interoperability specification (QTI) as a basis to generate the components of the e-assessment. We then extended the components to derive the Accessibility and Usability features Guidelines (AUG) using the WCAG. Specifically, we designed a prototype according to the derived AUG to check the accessibility for blind users. We analysed the results and found that the performance of the prototype is promising.**

Keywords-E-assessment; LMS; QTI; WCAG; Accessibility.

I. Introduction

E-assessments are used to evaluate the knowledge of a person. As technology develops, e-assessment becomes more complex due to a variety of implementations in various devices (mobile,iPad, iPhone and Android devices). Also the complexities with a multitude of users (especially disabled) make the design too complex. These conditionsmakethe accessibility in e-assessments more intricate and challenging.

The blind users are the ones who face a real challenge in using E-assessment using touch screen based Smartphones. The main communication and interaction for the blinduser is through audio. Hence, the questions and options for the answer should be presented in such a way that which can be accessed and usableby the blind users. Thus, we have adopted the Question and Test Interoperability specification (QTI) issued by IMS Global Learning Consortium as a standard to framethe questions for e-assessments.

Chang *et al*. [1] and Zualkernan*et al*. [8] propounded QTI structure in flash and digital TV format. The research by Han *et al*. [2] developed a tool called 'dinsEditor'to produce questions in QTI format. The otherwork by Martínez-Ortiz *et al*. [4] and Zhang *et al*. [7] provide a mechanism to convert QTI structures to accommodate them in the Android phones for sighted users. However, no work has been dedicated tothe QTI presentation structure that can accommodate the blind users for touch screen based smart phones.

The main aim of our research is to provide an e-assessmentpresentation structure for blind users. The paper is structured as follows: in Section II, the structure of questions in QTI model was analysed. In Section III, components of e-assessment required for accessibility and usability of a blind user was framed. In Section IV, the accessibility and usability guidelines (AUG) based on the components were derived. In Section V, the presentation design for e-assessment was proposed. In Section VI, the user study was discussed. In Section VII, the results of the user study were measured. In the last section, we discuss and conclude our study.

II. QTI Model

QTI is the standard in the industry for designing and exchanging the assessments. QTI version 1.2 was mostly implemented by e-learning tools during 2002. In 2005, QTI 2.2 was released to integrate with other models such as IMS Simple Sequencing, IEEE Learning Object Metadata and IMS Learning Design [2].

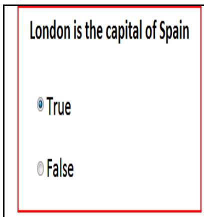
In QTI, a question (termed item) is builtfrom:

1. Question: the question to be asked, for example, "What is the capital of Spain?" (Stem).
2. Answer: the correct answer, in this case Madrid (Key).
3. Distracts: incorrect answers; in this case Paris, Berlin, London.
4. Rubric: Scoring or instructions for the assessment.

A single assessment can have many sections. A section can contain a numberof items. The material such as audio,video and images may be associated within a single item, a section or an entire assessment.The structure of the questioncan be explained through the metadata as shown in Figure 1.

Our research has focused onidentifying the component of e-assessments. Once the component is identified, the needed accessibility and usability features for each component is recognized to be implemented during the presentation and design of the e-assessment software.

978-1-4673-4937-6/13/$31.00 ©2013 IEEE




```
<item>
<presentation>
<material><mattext>London is the Capital of
 Spain</mattext>
</material>
<response_lid><render_choice>
<response_label>
<material><mattext>True</mattext></material>
</response_label>
 <response_label>
 <material><mattext>False</mattext></material>
 </response_label>
 </render_choice>
 </response_lid>
 </presentation>
 </item>
```

Fig. 1. QTI Structure for e-assessment.

III. COMPONENTS OF E-ASSESSMENT

To identify the components of the e-assessment, we took the software functional requirements such as asking questions,answering the questions, storing unanswered questions temporarily (flag), timing,scoring and so on at the requirement stage of the e-assessment software. With strong design implications, we converted this functional requirement to the components of e-assessments with relevant accessibility and usability inference.

As a consequence, we divided the components into three levels, namely: **presentation**, **response** and **inclusion**. The use of the case of the e-assessment shows that all three components are utilized by the blind users, however, the sighted user can access only the first two components. This is shown in Fig. 2, below.

Presentation: It is what we have presented to the blind user. It may comprise of a single item or multiple items. For blind users, items are presented through audio (screen reader or text-to-speech). For sighted users, the item may be in text, image or widget format. It consists of user identification, audio feedback, navigation, flagging and repeating the question, as shown in Figure 2.

Response: It is how the blind users respond to the items. It may be by using a screen reader which reads what the user types or by voice synthesiser to identify the response given through voice or through the widget, such as buttons. At present, the accuracy of converting speech-to-text in voice recognition is not achieved as expected.

Inclusion: Inclusion is an addendum given to disabled students to give a fair chance for attending the e-assessment. It is based on the disability of the examinee to modify the items or rubric (scoring) suitable for access or use without any difficulty. For instance, pausing during the e-assessment or extending the duration can be supplemented for blind users.

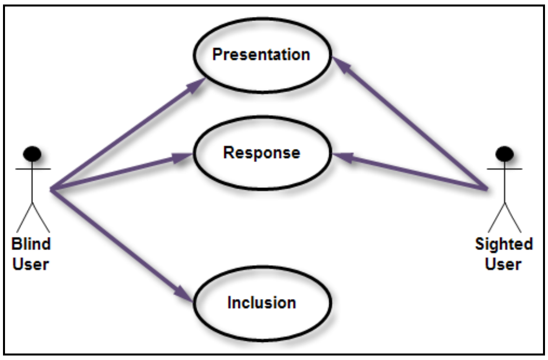


Fig. 2.User Case for e-assessment.

IV. GENERATING GUIDELINES FOR ACCESSIBILITY AND USABILITY ISSUES

Our approach consists of binding accessibility and usability guidelines that authorize developers to capture the functional requirementsof e-assessment in touch screen based Smartphonewithout depending on the expert. The information provided by the accessibility and usability guidelines (AUG) in the HCI literature is not directly related tothis rationale. We have analysedthis information from a developmental pointof view and elucidatedthe guidelines.

First, information regarding the AUG was extracted and categorized. We use WCAG guidelines for accessibility features as it was widely used. For the usability features, more detailed information was found in various literature [5][6]. This information has assisted to form a basis for recognizing usability issues.

The main obstacles in implementing AUG in touch screen based Smartphonefor blind user is the style of interaction. It is difficult for the blind users to identify the widget on the flat screen. Even if they identify the widget, the widget such as ‘list items’holding many items cannot be scrolled by the users. Therefore, we choose ‘button’ as pointof interactions that has to be accessible and easily usable by the user. Although QTI specifies 20 typesof questions, we adapted the question type in such a way that the user responds to the question through the button.

Under these circumstances, response by blind users to any closed format question type will be by just pressing the button. To put it more simply, only the number of clicks varies based on question type and not the type of interaction. For instance, multiple choice questions with four answer options, the blind user will press the buttons four times to understand the answer options. Consequently, blind users will press only two times for true/false or yes/no type questions.

Following this, AUG was framed according to e-assessment needs. Then we adopted suitable methods to solve it.Fig. 3 shows the design components for the e-assessment in the Touch Screen Device prototype.

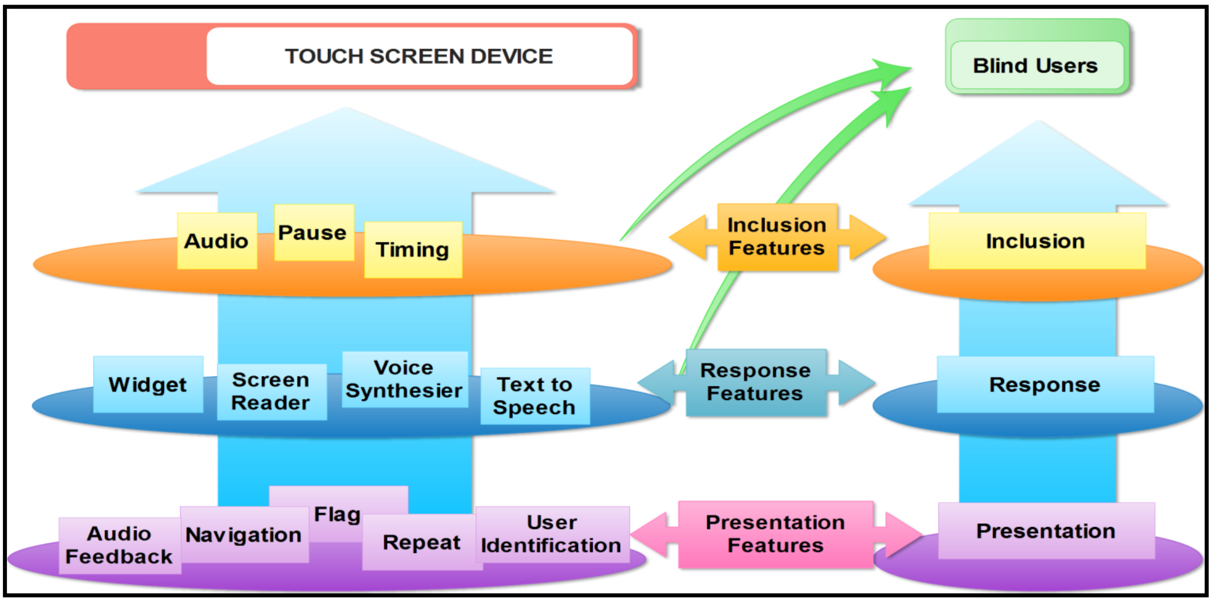


Fig. 3. Design Components for e-assessment in Touch Screen Device.

A. Accessibility issues for Presentation:

1) The blind user should know which question is currently under consideration.

2) The abbreviation should be properly pronounced. Currently, text-to-speech in Android phones pronounces abbreviations as words. For instance 'US' is pronounced as 'us'.

3) Unusual words must be properly pronounced. For instance, Roman-numeral format such as "I and II" in the options must be pronounced properly.

4) The blind user must have access to stimuli type of question with proper alternative text.

5) The description provided in alternative text of images needs to be understandable by a blind user. This needs the help of the expertise in the target user disability group.

B. Usability issues for presentation:

1) The blind user should know what type of question it is. For instance, in choosing the best answer, true/false or fill in the blanks.

2) Spelling and grammar in question and options must be proper.

3) Pronunciation of mathematical notation or expression should be pronounced properly.

4) Keywords must be properly understood by blind users.

5) The blind user should be able to understand the question. If needed, the repeat options should be made in order to play and hence listen to any number of times.

C. Accessibility Issues for Response:

1) The user must be aware of what type of gesture is being performed for each task such as to read the question and how to respond to the question (WCAG 2.1.1).

2) Consistent identification for response must be present on the screen (WCAG 3.2.4). For instance, options for the answer are kept in accessible locations and it should be maintained throughout the e-assessment.

3) Audio control is necessary if the response is thorough voice recognition (WCAG 1.4.2). Commonly, audio controls such as increasing or decreasing the volume are present in the mobile device itself. However, care should be taken to check these facilities before implementing it in the device.

4) Accessible widget should be placed for the user to perform the response to the question.

D. Accessibility issues for Inclusion:

1) Consistent navigation is provided to navigate through different questions (WCAG 3.2.3).

2) Should be designed to minimize the error (WCAG 3.3.6).

E. Usability issues for inclusion:

1) Provide feedback when interaction happens through haptic or through audio cues.

2) Provide feedback about instructions.

3) Provide flag operation for the questions.

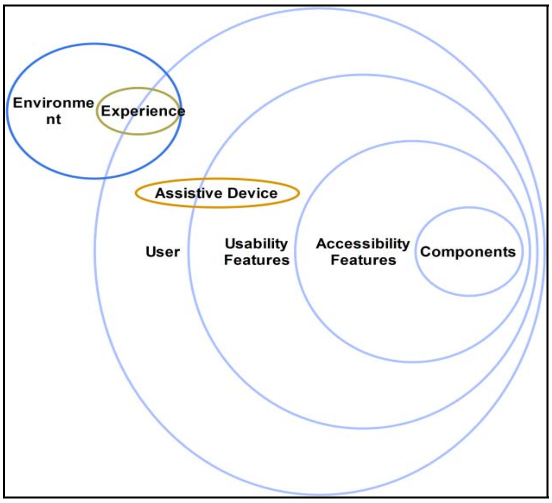


Fig. 4.AUG Framework for e-assessment.

Figure 4 shows the framework of e-assessment in which the components of e-assessment is the core. It should be accessible to blind users. The accessible components should be usable by blind users with assistive technologies such as text-to-speech or voice recognizer. For efficient operation of e-assessment, experience in using touch screen based Smartphoneandknowledge of how to use the e-assessment software is needed but is not mandatory.

V. Prototype Design

To evaluate the AUG,the QuizTouchprototype was developed for e-assessment to implement the accessibility and usability features. We aimed to design the prototype to achieve the following goals as sown in Fig. 5:

1. Only touch gesture is used.
2. Audio feedback must be provided on the touch screen, informing the cue.
3. Navigateto previous and next questions.
4. Each page contains only one question.
5. User can flag the question if they feel any difficulty and can be answered later.
6. Timing is allowed to inform the user about left over time to complete the e-assessment atregular intervals of time.
7. User must be able to repeat the audio of the question any number of times;
8. Answer options are cyclic;
9. Once the user identifiesthe answer for the question, s/he will press the answer buttonto reserve the answer for the present question.
10. The exam will be terminated by pressing the end button.
11. The scoring will be emailed to the concerned person automatically.

This featurecan be explained through the following user scenario.

User Scenario

John is a blind school student using QuizTouch app. His science teacher has uploaded QuizTouch app for a particular unit. She already explains to him how to operate QuizTouch. So John is well aware to hold the smart phone in his left hand and how to keep the left hand fingers over the marks. The buttons are placed near to the tips of the left hand fingers (marks). When he wants to press the button, he will touch at the tips of the left hand. He will confirm whether the button is pressed by hearing an audio cue of the button pressed. Hence John can take a quiz with confidence without any other assistance.

When John opens the QuizTouch app, it instructs John to tell his id. When John performs the same by using the voice recognizer, he receives information about the function of each button. He also gets the information about the duration of the quiz. Then quiz is started. The first question is heard through audio. John presses the options for the answer. On pressing four times of the option buttons, he receives an option cue to answer the question on each press. He feels the second option is the best answer. As answer options are cyclic, he presses twice to choose the second option. Then he presses the answer button present at the tip of the left hand middle finger. He presses the next button by pressing left hand ring finger to go to the next questions.

At this stage, the audio announces the time left for the quiz to complete. He felt that the second question is difficult and he wants to answer this at a later stage. So he presses the flag button by pressing below the left hand thumb finger. He moves to other questions to complete the answer. Finally, he presses the flag button to recall the flagged questions and answers them accordingly. Then he presses the end button by pressing left hand thumb finger to finish the quiz. He knows the quiz score will be emailed to his science teacher. Now he is anxious to discuss the quiz with his teacher.

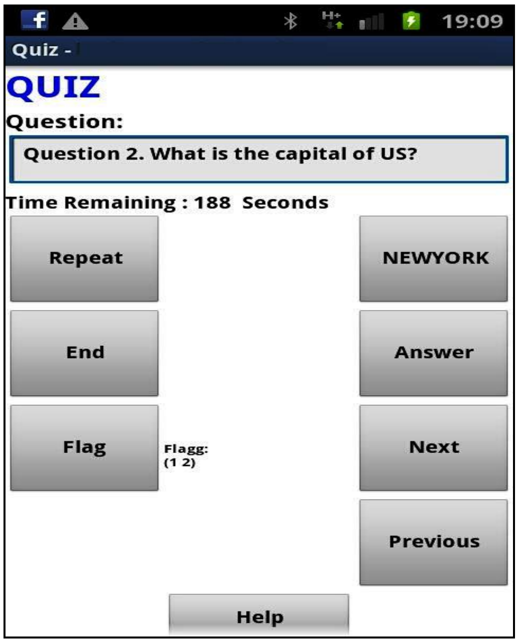


Fig. 5. Prototype Design for e-assessment.

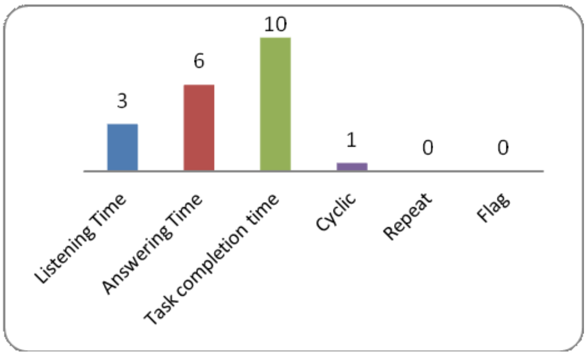


Fig. 6.Average Parametric Value for e-assessment.

VI. USER EVALUATION OF LISTENING STRESS

A. Participants

As it was difficult to find the blind person, we recruited three blind participants with an average age of 35 years. All participants, to some extent, have experience of using mobile phones with screen readers. None of the participants have experience of using touch screen mobiles, however

B. Equipment

To conduct the study we have developed a prototype to run on Smartphonessupporting the Android platform. The prototype was then tested on the Samsung Galaxy S2 mobile smartphone with blind users. The Smartphone running our prototype generates the speech according to the Android development code.

C. Procedure

The blind users were given the target name. On pressing each widget, the cue announcesthe name of it. The blind user has to repeat the task until they reach the desired target.

VII. ANALYSISOFTHE RESULT

User study analysis was performed to study the user behaviour with the QuizTouch prototype. The results were presented based on the participants' rating and our observation during the user study.

A. Data Collection

Since the cohort size of the participant is small, we conducted two trials for two participants each and three trials for one participant. Thus, in total seven trials were conducted. A single trial contains four questions. For each question, we recorded the time taken to listen to the question (listening time), answer the questions (answering time), the completion time of each question, the number of times the user use cyclic options to listen to the answer options (cyclic), number of times repeat is used for listening question (repeat) and the score for the question which was internally evaluated. Thus, 28 observations (7×4) were recorded.

B. Task Completion time

It is the time to complete a single question which is the sum of listening time and answering time. The average task completion time is shown in Figure 6. Because task completion time was normally distributed (Shapiro-Wilk W=0.957, p>0.05), we used the parametric test on completion time. We found significant effects of participant (F(2,25) = 9.251,p= 0.021) on the task completion time. But there was no significant effect of question (F (3, 24) =0.850, p=0.480) on task completion time.

C. Listening Time

The parametric test one way ANOVA was used to analysethe significance of listening time since it was normally distributed(Shapiro-Wilk W=0.926, p>0.05). The study found that there weresignificant effects of participant (F(2,25) = 5.017,p=0.015) on listening time. But the significant effect of question (F(3,24)=0.037, p=0.990) on listening time was not found. The average listening time is shown in Fig. 6.

D. Answering Time

The average answering time is shown in Figure 6. Because the answering time was normally distributed (Shapiro-Wilk W=0.916, p>0.05), we used the parametric test foranswering time.We found there were significant effects of participant (F(2,25) = 4.159,p= 0.028) on answering time. But significant effect of question (F(3,24)=1.695, p=0.195) on answering time did not exist.

E. Feedback Analysis

The prototype was rated by participants with a series of statements using a 5-point Likert-type scale (1=strongly disagree, 5=strongly agree). We perform reliability analysis to calculate and report Cronbach's alpha coefficient for internal consistency reliability of the Likert-type scale used. We found that Cronbach's alpha is 0.937 and 0.956 for standardized items which indicates good internal consistency of the items in the scale. A Friedman test revealed a significant effect of the feedback evaluation on rank ($\chi 2$ (2) = 30.720, P <0.05).

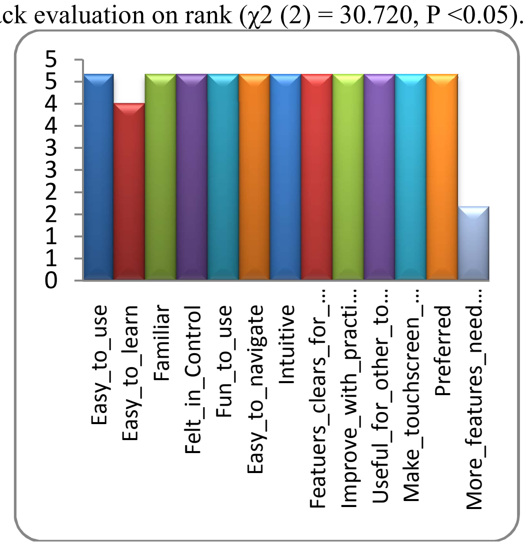


Fig. 7.Feedback in Likert Scale for e-assessment.

Figure 7 show that on average, all features are strongly agreed (5 points) with standard deviation of 0.5 exceptmore features are needed by the blind users.It is interesting to note that only a number of blind users disagree (2 points) to need more features.

VIII. Discussion and Conclusion

The AUG for blind users is difficult to frame for a touch screen based smartphone application, since the style of interaction varies between the sighted and the blind user. The literature reviewsuggests AUG for general apps and there are no specific guidelines for e-assessments. Thus a more detailed analysis of the nature of the problems was studied in this research.

At the beginning of the user study, the participants lacked in experience in handling the touch screen device. When we showed them how to use the touch screen based on the marking, they tried with the demo prototype, until they became comfortable with handling the device. Once they were confident, the actual app prototype was loaded and tested for the study session.

As our intention is to find accessibility and usability of apps and not to assess the blind users' knowledge, we asked simple and easy questions. Thus our study analyses shows that the question and answer session does not have much impact on listening time, answering time and completion time. On average, participants took three seconds to listen to the question. But it depends on the readability of the text. Interestingly, they took double the time of listening time to answering the question. This does not actually reflect the capability of the blind user to answer the question since it involves the operational time to press the button.

The analysis of repeat and cyclic options shows the easiness of the question. Out of 28 questions which were used by participants, on average, none of the questions were repeated by the user. In the study, 112 answer options were checked by the blind users, in total 17 times cyclic options were used.

The feedback analysis shows that participants were very much satisfied with the prototype. On average, they rank all the feedback to 5 (Strongly Agree). They are not of a view to include more features in addition to the existing features (2 – Disagree).

While our e-assessment touch screen prototype provides basic access and is usable by blind users, still the research has to be made on different question types. For instance, the question based on the images, usability and accessibility issues has to be identified. We use only 'button' as a pointof interaction for whole e-assessment except the voice synthesizer was used to identify identification. But more control has to be studied in order to facilitate the easinessin using the e-assessment app.It should be thought that this study is suggesting the presentation design for an e- assessment appfor blind users.This conclusion serves as a guide for developers to redesign the e-assessment keeping the blind users in mind.

References

[1] Chang, H., Liu, P., Shih, T. K., and Chen, Y. 2008. Developing QTI compliant assessment platform on digital TV. In Proceedings of the 2008 IEEE International Symposium on IT in Medicine and Education (ITME 2008) (Xiamen, China, December 12-14, 2008).

[2] IMS Global Learning Consortium: IMS Accessible Portable Item Protocol (APIP): Best Practice and Implementation Guide. Available at: http://www.imsglobal.org/apip/apipv1p0cf/APIPv1p0_Best_v1p0cf.html [10/3/13]

[3] Han, S., Kim, J., Lee, Y., Cha, J., and Choi, B. 2010. disnEditor: A browser extension for QTI-compliant assessment item authoring. In Proceedings of the 10th IEEE Int. Conference on Advanced Learning Technologies (ICALT 2010) (Sousse, Tunisia, July 5-7, 2010).

[4] Martínez-Ortiz, I., Moreno-Ger, P., and Sierra, J. L. 2006. <e-QTI>: a reusable assessment engine. In Proc. of the 5thInternational Conference on Web-Based Learning (Penang, Malaysia, July 19-21, 2006).

[5] Nielsen, J. (1993). Usability Engineering. John Wiley & Sons.

[6] Shneiderman, B. (1998). Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley.

[7] Zhang, P., Wills, G. B., and Gilbert, L. 2010. IMS QTIEngine on Android to support mobile learning and assessment. In Proceedings of the 2010 International Computer Assisted Assessment (CAA) Conference (Southampton, UK, July 20-21, 2010).

[8] Zualkernan, I. A., Ghanam, Y. A., Shoshaa, M. F., and Kalbasi, A. S. 2007. An architecture for dynamic generation of QTI 2.1 assessments for mobile devices using Flash Lite. In Proceedings of the Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007) (Niigata, Japan, July 18-20, 2007).