Mobile Assessment and Annotation Tool V2

Sibulele Khanya Kupelo

Submitted in partial fulfilment of the requirements for the degree of
Bachelor of Commerce
in the Faculty of Science
at the Nelson Mandela Metropolitan University

Supervisor: **Charmain Cilliers**

09 October 2017

Abstract

Write your abstract here.

Keywords

List of keywords here.

1. Introduction

The third report of the Mobile assessment annotation tool was focused on two activities of the Design science research methodology, namely, Outline and define requirements activity, the Design and develop artefact activity and the Demonstrate artefact activity. The Outline and define requirements activity focused on identifying and outlining an artefact that can be used to solve a problem and to define system requirements. The Design and develop artefact activity focused on creating a prototype that can be used to address the requirements defined in the outline and define requirements activity (Johannesson & Perjons, 2012). The aim of the Demonstrate artefact was to use the components created for the prototype developed for report three in order to prove their feasibility. In general, the activity was used to answer the question: How can the developed components be used to address the explicated problem (Johannesson & Perjons, 2012)?

Below is a list of requirements defined for the Mobile assessment and annotation tool:

1. **Assess Images:** The system should allow users to assess image submissions.

2. **Generate submissions:** the generation of submission files from source files on importing. For example, the programming and web modules have text files (. cs, Java, XML, fxml, .html, etc.) that need to be collated into a single submission file. Instead of having the student copy and paste files into a document, then exporting it, the system should allow a project to be imported and a submission file is generated from the source files.

3. **Moodle Integration:** The system must allow an examiner to upload files and download submission from the Moodle system. The system will not allow a moderator to perform this task since a moderator can only work with files that have been assessed by an examiner.

4. **Automatic emailing:** The system must allow an assessor to automatically email feedback files to every student that made a submission. A student must not receive other feedback files that belong to other students.

5. **Mark adjustment:** Adjust marks when a check mark has been removed.

6. **Moderating using different Colour:** Moderator uses a different colour to moderate submissions and they should not be allowed to modify original markers' annotations.

7. **Symbol recognition:** The system should be able to recognise symbols drawn by the users (Kupelo, 2017).

The third report focused on the assess images, generate submissions, Moodle integration, automatic emailing and symbol recognition requirement. The use case diagram in Figure 1 was used to contextualise these requirements. In the third report, Kupelo (2017) stated that the requirement assess images is associated with the use case called "Load information" and the use case called "Marking Submissions". Kupelo (2017) further argued that in this requirement if the image and pdf file submitted for assessment are different files then the image file must be named as "Assignment Name_ Subject Name_ Student Number_ Question Number". The purpose of this naming is for ensuring that the image file is inserted in a correct pdf submission file. Once the image has been attached to the pdf document a user must be able to draw annotations on it.

In report three the Moodle integration requirement was associated with the use case named "Download and upload files on Moodle". It was further discussed that under this requirement an assessor can only upload submission files into Moodle if they have been marked and each submission must be uploaded to a student account that corresponds with the way the file is named. Moreover, students must not receive feedback files that belong to other students. It was also further discussed that the download part of this requirement is associated with the "Load information" use case. Once the examiner has downloaded the files from will be saved on a tablet device so that they can be uploaded into the Mobile assessment and annotation tool app.

Furthermore, in the third report, it was discussed that the generate submissions requirement is associated with "Load information" use case and the "Marking Submissions" use case. This requirement is mainly performed by the system, an assessor only saves the source code into the tablet and the app imports the source files into a pdf document and loads the files into the app for marking. Additionally, in the previous report, it was also discussed that the automatic emailing requirement is associated with the "Automatic emailing of feedback files" use case. Only an assessor can perform this functionality and only students who submitted files for an assessment can be emailed feedback files. Lastly, Kupelo (2017) stated that symbol recognition is associated with the "Marking submissions" use case. In the report, it was mentioned that two symbols will be
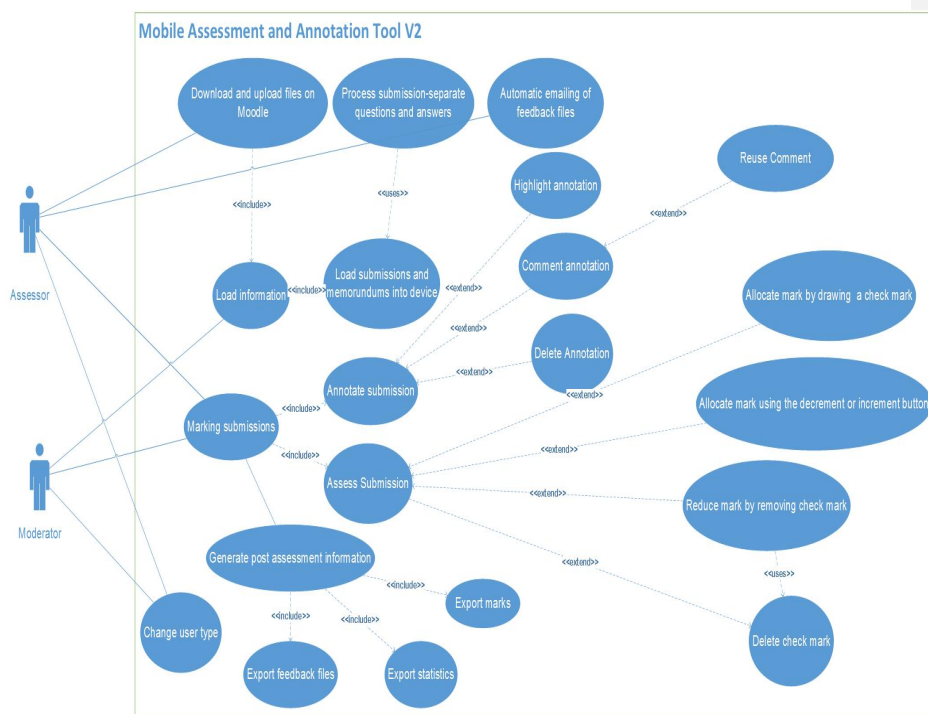
recognised, namely, cross symbols and check marks. Where the cross symbol is used for indicating that an answer provided by a student is incorrect.

Figure 1: Use case diagram

After demonstrating the prototype developed for the third report, the project stakeholder made a made few valuable suggestions that could help in ensuring that the prototype is performing effectively and efficiently. After taking these suggestions into considerations few changes were made on the prototype developed for report three and these changes affected the following requirements:

- Assess images
- generate submissions

- Symbol recognition
- Mark adjustment

This report will provide detailed information about the changes made on the prototype developed for report three and how these changes affected the abovementioned requirements. Moreover, this report will discuss the evaluate artefact activity of the design science research methodology, the evaluate artefact activity attempts to describe how well a developed artefact can solve a given problem. The Evaluate artefact has two sub-activities, namely, the Choose evaluation strategy and Carry out the evaluation.

2. Improvements on the Previous Prototype

A few improvements were made on the prototype developed for the third report, and these improvements affected some of the system requirements, namely:

- Assess images
- generate submissions
- Symbol recognition
- Mark adjustment

This section of the report provides detailed information about these prototype improvements and how the above-mentioned requirements were affected by these improvements.

2.1. Symbol Recognition

In report three, it was discussed that an intensive research will be conducted on Android gesture recognition and event handlers (Kupelo, 2017). The main reason for conducting this research was to try to find ways that can be used for substituting the symbol recognition functionality since the developer did not have enough time and lacked the expertise for developing the functionality. For this project, symbol recognition refers to the case whereby the app can differentiate in real time between a check mark or cross symbol drawn by a user when they assess a student submission using the Mobile assessment and tool app. According to Android (n.d.), there are seven gestures supported by devices running on the Android operating system, namely:

1. Scrolling: This gesture is executed when the user presses down on the device screen and then the user moves his or her finger as if they are dragging an item
2. Swiping or Flinging
3. Long press: This gesture occurs when a user touches the screen for a few seconds without moving their finger.
4. Show press: Pressing down without moving the finger up left, right or up
5. Tap: This gesture refers to a light touch on the device's screen with your finger.
6. Double tap: This gesture occurs when a user touches the screen rapidly twice.
7. Pinch: This gesture involves using two fingers, typically used for zooming in or out.

The initial objective was to use two of the above-mentioned Android gestures and standard symbols used for assessing submissions to replace the symbol recognition functionality, namely, scrolling and swiping. Where the scrolling gesture was going to be used for drawing the cross symbol for indicating that no mark has been allocated. In the first version of Mobile assessment and annotation tool app, the action up motion event was used to allocate mark for a submission after a user drew a check mark. The action up motion event occurs when the user releases the finger used to touch a view. On this version of the mobile app, a positive mark for a submission can only be allocated if a check is drawn by using the swiping gesture. Because the third-party software development kit (PDFTron PDFNet) used to implement the Mobile assessment and annotation tool already has its own gestures the developer could not use the Android gestures. Implementing the Android gestures clashed with the methods that come with the third-party SDK, thus there were many compile-time errors.

2.1.1. Implementation

To overcome the above-mentioned issue the developer decided to use the third-party SDK's gesture detectors since most core functionalities of the mobile app are developed using the third-party SDK (Wildervanck, 2016). The issue with gestures provided by this SDK was that it does not have the scrolling gesture (Developer, 2014). Thus, instead of using the swiping and scrolling gestures, the developer solved this issue by deciding to use a combination of the swiping and the long press gesture, for allocating a positive mark and indicating that no mark is allocated, respectively.

For a user to successfully allocate a positive mark for a submission, the user must activate the pen marking button and draw a check mark by using the swiping gesture. To draw the symbol for
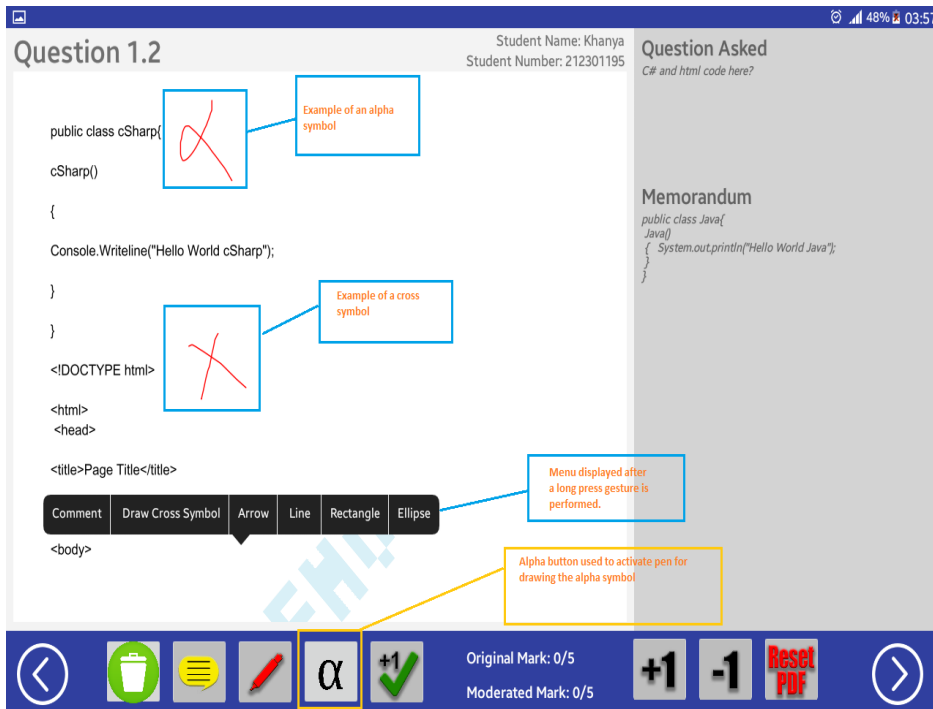
Figure 2: Screen containing the cross symbol, alpha symbol, alpha button and menu that displays after a long press

indicating that a mark has not been allocated the button for activating the pen tool must be deactivated, this prevents the app from awarding a mark when a user draws the symbol. After deactivating the button the user must perform a long press on the screen, this gesture will display a menu shown in Figure 2, and from the displayed menu options the user must select the "Draw Cross Symbol" menu item. This menu item activates the pen for drawing the cross symbol.

Drawing the cross symbol will require the user to perform the long press gesture and the other succeeding steps twice because when a user lifts the finger used for drawing the pen tool deactivates. To minimise the number of steps the user can draw a symbol that looks like an alpha

since it can also be used to indicate that a mark has not been allocated, refer to Figure 2. The button with an alpha symbol in Figure 2 can be used as an alternate way for activating the pen tool used for drawing the alpha symbol.

The alpha button added on the app acts as an interface metaphor for the alternate symbol used to indicate that a mark has not been awarded. Interface metaphors are user interface elements that are created to be comparable to real-world objects and utilised to take advantage of existing information from users (Averbukh, et al., 2007). According to Marcus (2014), systems that make use of interface metaphors makes it easy for users to learn and understand how a system works. Moreover, a system that has interface metaphors makes it easy for the user to memorise certain parts of the system.

These changes affected the mark adjustment requirement associated with use case named "Reduce mark by removing the check mark" in Figure 1. Prior the newly added functionalities to adjust a mark by removing a check mark a user had to perform the following two steps:

1. Select check mark
2. Click delete button

To prevent the app from reducing marks when a user removes the cross or alpha symbol a new extra step was added when a user performs the reduce mark by removing check mark use case. The new steps for performing this use case are as follows:

1. Select check mark
2. Activate pen tool
3. Click delete button

Lastly, the steps for removing the alpha or cross symbol are different from the steps executed when removing a check mark. To remove the cross or alpha symbol the user must select the symbol and then click the delete button.

2.2.   Assessing submissions with Images

After consulting with project stakeholders, it was found that the assessing submissions with images requirement was misinterpreted. Instead of the system processing a separate image and

submission and combining them together to form a single submission, the system should process a submission that already has images inserted into it. This section will provide detailed information about the improvements made in the implementation that addresses this requirement.

In the first version of the mobile the first version Mobile assessment and annotation tool for a user to perform the use case "Marking submission" the developer implemented a functionality that splits a submission file into smaller PDF files containing students' answers associated with a certain question (Wildervanck, 2016). To create these PDF files the app first extracts the text from the original PDF file and then the app inserts the text into a new PDF file. These new PDF files containing student answers are then displayed on the app using a PDF view that allows a user to assess and annotate the file containing a student's answer (Wildervanck, 2016). When the user compiles the feedback files after marking the student answers, the smaller PDF files containing the assessed and annotated answers are merged to form a single submission file for each student.

1)1~

UML Use Case Diagrams. Use case diagrams are usually referred to as behavior diagrams used to describe a set of actions (use cases) that some system or systems (subject) should or can perform in collaboration with one or more external users of the system (actors).
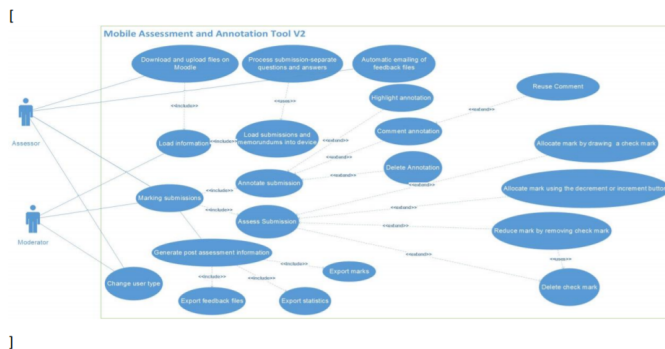
[

]

Figure 3: Student answer for question 1 with an image enclosed within square brackets

To address the assess submissions with images requirement an additional functionality was added to the functionality developed by Wildervanck (2016). This new functionality extracts images from a submission file and inserts them into the new PDF file containing a student's answer. When the PDF file containing a student's answer is loaded on the app's PDF view the image associated with the answer is also displayed and the system user is able to assess and annotate the image.

In order for this new functionality to work successfully, in the original submission file, the image must be enclosed in square brackets, as seen in Figure 3. The square brackets are omitted on the new PDF file, the main purpose of these brackets is to ensure that an image is inserted into the correct position in the new PDF file. Figure 4 shows how the student answer with images is displayed when the user is marking the answer.

Figure 4: Student answer with assessed and annotated image

2.3.   Generate submissions files from source files
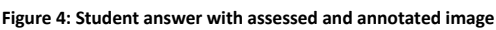
The prototype developed for the third report made use of a library called Aspose to address the generate submission requirement. To generate the submission files the library converted source files into PDF formats. After the conversion was successful the converted files were then merged to form a single submission file. The usage of this library was discontinued because it only converted three source files, namely, HTML, XML and XHTML (Aspose,2015). After the usage of this library was terminated a new functionality was implemented to address requirement. This section will provide information on how the new functionality addresses the requirement.

To address the requirement, the newly implemented functionality creates a PDF file that contains source code extracted from the source files saved in the student's submission folder. For this new functionality to successfully address the requirement, five rules must be followed, these are:

1. The student's submission folder containing source files must be named as <NMU_ Module Name _ Assessment Name_ Student Name _Student Number>, refer to Figure 5. The name of this folder is then used as the name of the new PDF file containing source code.

2. The source files must be organised in subfolders in the student's submission folder. For example, if the source files submitted are answers for question 1 then the source files must be saved in a subfolder called "Question 1" or "Q 1", this is depicted in Figure 5. Storing the source files this way allows the functionality to take the subfolder name and make the name a question number for the PDF document, before the subfolder name is made a question number the non-numeric characters are removed. After the question number has been set the functionality extracts source code from source files and inserts the code into the PDF document.

3. The user must ensure that there is a memorandum that has the same number of questions as the student's submission. The memorandum must be named as <NMU_ Memo _Module Name _ Assessment Name> (Wildervanck, 2016).

4. In the tablet device, the student's submission folder must be saved in the following path: MobileAssessmentToolPdfs -> Module Name_Assignment Name Folder -> student's submission folder, refer to Figure 5.

These rules help to ensure that the newly implemented functionality adheres to the question recognition algorithm. Wildervanck (2016) defines question recognition as a method whereby

individual questions and answers can be identified from the text and separated from another. Question recognition also helps to ensure that the system meets all other system requirements, for example, assessing and annotating submissions, generating and emailing feedback files (Wildervanck, 2016).
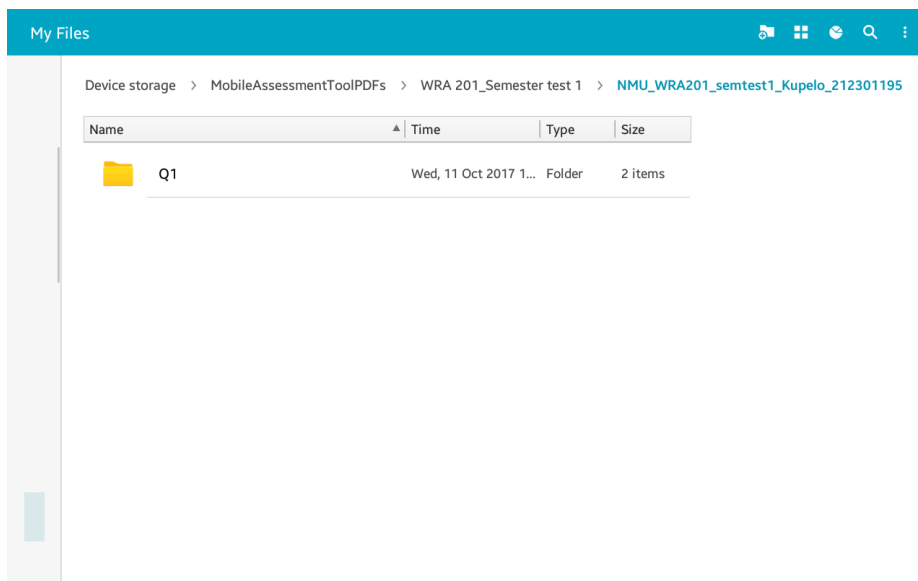

Figure 5: Student's submission folder containing source files

3. Evaluate Artefact

This section focuses on the fifth phase of the design science research methodology, which is, evaluate artefact. According to Johannesson & Perjons (2012), in this phase researchers are most interested in finding out how well can the developed activity solve the explicated problem and the researchers also want to know the degree in which the developed artefact meets the defined

requirements. The main output of this activity is an evaluated activity, together with information that describes how well the developed artefact works and why.

The Evaluate artefact phase is made up of two subactivities, namely:

i.   Choose Evaluating strategy: Determines how the evaluation is to be conducted

ii. Carry out the evaluation: Performing the actual evaluation (Johannesson & Perjons, 2012).

3.1.   Choose Evaluation Strategy

Johannesson & Perjons (2012) further argue that there are two strategies that can be used to evaluate an artefact, namely, ex-ante evaluation and ex-post evaluation. In an ex-ante evaluation, the artefact is evaluated without being used. The ex-post evaluation, on the other hand, requires the artefact to be utilised. For evaluating the Mobile assessment and annotation tool app, the researchers will evaluate in the ex-post evaluation. Evaluating in ex-post evaluation strategy will help the researcher to evaluate the usability, comprehensibility and learnability of the artefact.

For the evaluation of this artefact, the researcher will make use of a survey research strategy, and the data generation method to be used is a questionnaire. Research Methods Knowledge Base (2006), defines a questionnaire as a simple paper and pencil based instrument completed by respondents. The main reason for choosing a questionnaire as a data generation method for the survey it is because of the advantages it offers, these are:

i.   Low costs

ii. Makes gathering data convenient

iii. Provides accurate findings

iv. There is little or no researcher biases

v.   It is easy to find results that are statistically important (Explorable, 2017).

Lastly, Johannesson & Perjons (2012) argue that, although questionnaires are considered as superficial and restrictive to researchers in terms of gaining rich insights from the views provided by respondents, they are highly efficient in gathering views and insights about an artefact.

3.2.   Carry out evaluation

This section is divided into four subsections, these are:

1. Evaluation process design
2. Evaluation methodology
3. Evaluation results
4. Findings and recommendations

The evaluation process design subsection discusses how the evaluation for the final prototype of the Mobile assessment and annotation tool was conducted and the equipment used during the evaluation. This subsection will also provide detailed information about the tasks performed by the research moderator during the evaluation phase. The evaluation methodology subsection provides information about the metrics that were collected during the evaluation of the system. Moreover, the section provides demographic data of the participants who participated in the evaluation of the mobile app.

 Lastly, the evaluation results subsection provides qualitative and quantitative results of the data collected during the evaluation. The findings and recommendations subsection provides a list of findings and recommendations from the collected quantitative and qualitative data.

3.2.1. Evaluation Process Design

The evaluation carried out by the research moderator involved the second version of the Mobile assessment and annotation tool Android app which is intended for users who will be assessing and annotating student submissions. During the evaluation study of the mobile app, the system requirements listed below formed part of the evaluation. Each requirement is associated with at least one task provided from the list of tasks found appendix c.

1. **Assess Images:** The system should allow users to assess image submissions.

2. **Generate submissions:** the generation of submission files from source files on importing. For example, the programming and web modules have text files (. cs, java, XML, fxml, .html, etc.) that need to be collated into a single submission file. Instead of having the student copy and paste files into a document, then exporting it, the system should allow a project to be imported and a submission file is generated from the source files.

3. **Automatic emailing:** The system must allow an assessor to automatically email feedback files to every student that made a submission. A student must not receive other feedback files that belong to other students.

4. **Mark adjustment:** Adjust marks when a check mark has been removed.

5. **Moderating using different Colour:** Moderator uses a different colour to moderate submissions and they should not be allowed to modify original markers' annotations.

6. **Symbol recognition:** The system should be able to recognise symbols drawn by the users (Kupelo, 2017).

Firstly the assess images, generate submissions, moderating using different colour and mark adjustment requirement is associated with task 1, 2, 3, 5, 6, 7, 8, in these tasks, participants were required to assess submissions by adding and removing standard icons used for assessing submissions both as a moderator and assessor. Two submission files were provided to each participant for assessing and annotating, the first submission contained both images and text and the second submission was a PDF file with source code.

The Automatic emailing requirement is associated with task 4 where the participant had to email feedback files to students. Lastly, the symbol recognition requirement is associated with task 1, 3, 6, 7 and 8, whereby the user had to draw standard symbols used for assessing by using gestures. Lastly, the questions provided in the PSSUQ in appendix B is based on the newly added functionalities used to address the list of requirements above not the entire system.

The evaluation for the mobile app was conducted in an office at the Nelson Mandela University Computing sciences department. Conducting the evaluation in this helped the participants to take on the role of actual end users and it also helped in ensuring that the test results will be a greater predictor of the mobile app's performance in the place where it is normally used (Rubin & Chisnelll, 2008).

To perform the list of tasks (Appendix C) provided by the research moderator participants used a Samsung tablet. The tablet device used by the participants during the evaluation was running on Android version 5.0.2 (Lollipop). Participants also completed a questionnaire (Appendix B) related to the tasks they performed.

During the evaluation, the moderator performed the following tasks:

- Introduced the session to each participant and provided a detailed explanation about the purpose of the mobile app to be evaluated.
- Provided a participant with a consent form (Appendix A)
- Explained the tasks to be performed by each participant during the evaluation
- Recorded the amount of time a participant took to complete a task using a stopwatch
- Assisted in cases where the participant needs clarification with certain user interface elements

A subset of the components of the Mobile assessment and annotation tool was evaluated, the evaluation was more focused on the newly added functionalities of the mobile app. The requirement Generates submissions discussed in section 1 was not evaluated because this is a functionality performed by the system. Instead, users were evaluated on the output generated by the functionality handling this requirement. The first objective for conducting the evaluation was to discover and fix the mobile app's usability deficiencies. The second objective was to eliminate any difficulties and frustrations triggered by the way the mobile app has been designed.

3.2.2. Evaluation methodology

This section provides a detailed discussion of how the evaluation was carried out and the metrics that were collected during the evaluation. Moreover, this section provides detailed demographic data of the research participants (Musfid, 2016). The within-group design evaluation technique was used to conduct the evaluation with research participants. Rubin and Chisnelll (2008) describe this evaluation technique as a method where each participant performs all the given set of tasks.

The evaluation took a maximum of 20 minutes, on the first 5 minutes of the session the research moderator explained to the participant the main purpose of the system and requested the participants to complete a consent form. Additionally, the remaining 15 minutes was used for performing the tasks provided in appendix C and to complete an evaluation questionnaire. During the evaluation, three usability metrics were collected, namely, effectiveness, efficiency and satisfaction (NIST Industry USability Reporting project, 2001).

The effectiveness usability metric was used by the researcher to relate system usage goals to the precision and comprehensiveness with which these goals can be achieved. The effectiveness of a system can be measured by the percentage of task completion, the number of errors, the rate of recurrence of assists to the participant form the testers and the rate of access to help or documentation (NIST Industry USability Reporting project, 2001). This evaluation only focused on three measures of effectiveness, task completion rate, time taken to complete a task and the number of errors.

The completion rate was used to measure how well a task was completed by the participants. The 3-score measurement was used to determine the level of completion as indicated below:

- 1- Successful
- 0.5- Completed with minor errors
- 0- Failed to complete the task

Errors refer to cases where the evaluation participants were unable to successfully complete a given task or cases whereby participants had to attempts to perform certain parts of the task more than once. The other measure that was used for the measuring is "assists". Assists refer to cases whereby the participant cannot continue with a task, the evaluation moderator then provides immediate practical assistance in order for the evaluation to proceed (NIST Industry USability Reporting project, 2001).

According to NIST Industry Usability Reporting project (2001) efficiency metric associates the effectiveness level achieved to the total amount of resources used. Mean-time taken to complete a task is used as a measure of efficiency. Satisfaction describes a subjective response from the user when they use the mobile and to questionnaires are used as measures for this metric.

Lastly, for evaluating the Mobile assessment and annotation tool app, a maximum of five lecturers from the Nelson Mandela University computing science department was selected as evaluation participants. Rubin and Chisnelll (2008) argue that research has shown that if four to five participants have been selected to represent the entire end-user population, then 80 percent of deficiencies associated with the usability of a system can be revealed. The selected participants all had solid knowledge about how to use a smartphone or tablet device. Additionally, the participants

had at least one year experience of marking paper-based submissions. The table below summarises the demographic information obtained from the participants, where a participant's name is represented using a number to maintain their anonymity.

Table 1: Participant data

Participant Number	Gender	Profession	Paper-based marking experience (Years)	Electronic-Based Marking experience (Years)
01	Male	Lecturer	1	0
02	Female	Lecturer	20	10
03	Female	Lecturer	-	-
04	Male	Lecturer	22	5
05	Female	Lecturer	Many	8

3.2.3. Evaluation results

This section provides information about the usability evaluation of the Mobile assessment and annotation tool. The section is further divided into two subsections, namely, quantitative analysis and user qualitative analysis. The rate at which tasks were completed by participants will be discussed quantitative analysis subsection. Additionally, the quantitative analysis also contains the time taken to complete each task and the number of errors incured when performing a task. Lastly, the qualitative analysis section contains the results collected using the post-study system usability questionnaire (PSSUQ).

3.2.3.1. Quantitative analysis

3.2.3.1.1. Task completion time

Task	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Average in seconds (Task)	Standard Deviation
Task 1	12	22	20	10	5	13.80	7.09
Task 2	28	15	8	30	15	19.20	9.42
Task 3	5	2	6	15	20	9.60	7.57
Task 4	11	15	12	60	10	21.60	21.55
Task 5	4	5	10	20	3	8.40	7.02
Task 6	10	5	2	10	5	6.40	3.51
Task 7	3	5	2	30	11	10.20	11.61
Task 8	1	3	5	10	12	6.20	4.66
Average in Seconds (Participant)	9	9	8	23	10	12	6

Table 2: Table showing average time spent by a participant performing the tasks. The table also shows the average time for spent by participants on each task.
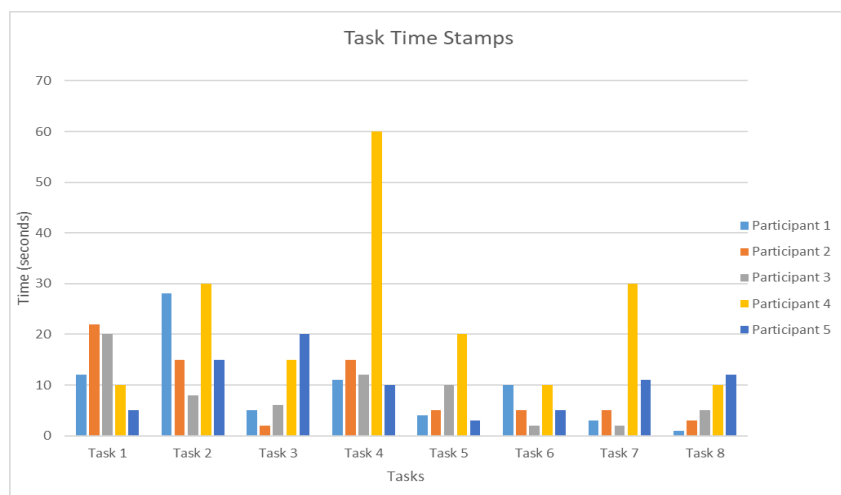

Figure 6: Task Time Stamps

The amount of time taken by a participant to complete a task was recorded by the research moderator, the recorded time included tasks in which the participants were assisted when they perform the given set of tasks. The mean time it took all participants to complete the given set of tasks is 12 seconds with a standard deviation of 6 seconds. Task 4, which required a participant to email feedback files students can be considered as the worst performed task since it has the largest

mean time of 21.60 seconds, refer to Table 2. As can be seen in Figure 6 the amount of time taken by most of the participants was excessively long and the outcomes this task differed a lot between the participants, with participant four having the largest outcome.

Task 2 is the second worst performed task by the participant, the goal of this task was to reduce a mark by removing a check mark and it took participants an average of 19.20 seconds to complete it. Lastly, as seen in Table 2 task 1 is the third worst performed task, the goal of this was to use the swipe gesture to allocate a positive mark for a submission. The mean time it took participants to complete this task is 13.80 seconds.

3.2.3.1.2. Task success rate

The second measurement that was also recorded during the evaluation is the task success rate, a task was considered successful if a participant did not get any form of assistance from the research moderator. Judging from the data in Table 3 and the graph in Figure 7 Task 2 is considered as the worst performed task with a success rate of 70%, while task 3 which required a participant to mark a student answer as incorrect had an average success rate of 90%. Task 4 and Task 5 had the same success rate as task 3, the goal of task 5 was to switch from assessing as an assessor to assessing as a moderator. Lastly, all other tasks had a success rate of 100% and the overall task success rate for all the participants is 91.3%.

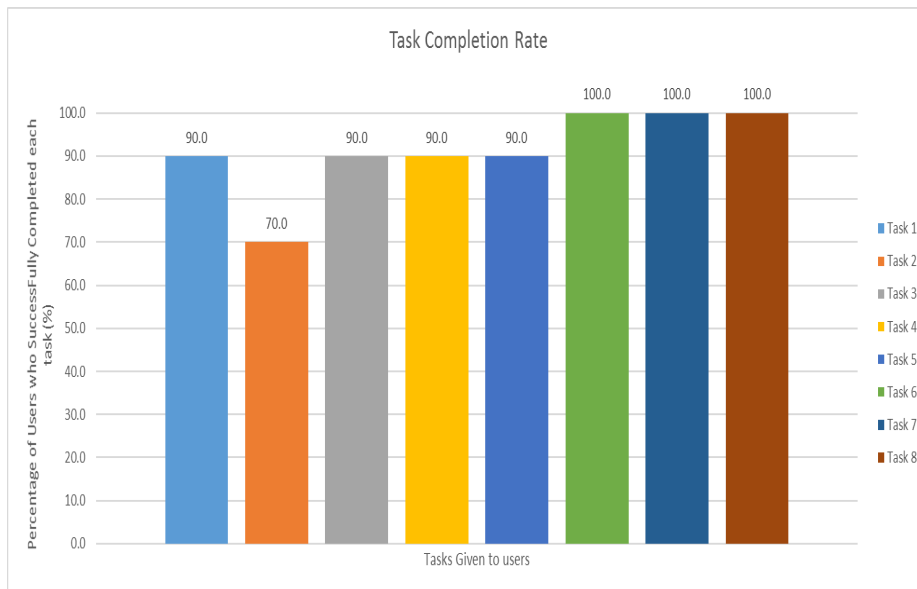| Task	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Average completion rate per task (%)	Standard Deviation (Task)
Task 1	1	0.5	1	1	1	90.0	0.22
Task 2	0.5	0.5	1	1	0.5	70.0	0.27
Task 3	1	1	1	1	0.5	90.0	0.22
Task 4	1	1	1	0.5	1	90.0	0.22
Task 5	1	1	0.5	1	1	90.0	0.22
Task 6	1	1	1	1	1	100.0	0.00
Task 7	1	1	1	1	1	100.0	0.00
Task 8	1	1	1	1	1	100.0	0.00
Average completion rate per participant (%)	93.75	87.50	93.75	93.75	87.50	91.3	14.60

Table 3: Task Completion Rate Data

Figure 7: Graphical Representation of Task Completion Rate

3.2.3.1.3. Number of Errors

Task	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Average number of errors per task	Standard Deviation per task
Task 1	0	1	0	0	0	0.20	0.45
Task 2	1	1	0	0	1	0.60	0.55
Task 3	0	0	0	0	1	0.20	0.45
Task 4	0	0	0	2	0	0.40	0.89
Task 5	0	0	1	0	0	0.20	0.45
Task 6	0	0	0	0	0	0.00	0.00
Task 7	0	0	0	0	0	0.00	0.00
Task 8	0	0	0	0	0	0.00	0.00
Average number of errors per participant	0.125	0.25	0.125	0.25	0.25	0.20	0.32

Table 4: Task Error Rate Data

The number of errors per task and per participant was the very last logged quantitative measure during the evaluation. From Figure 8 it can also be seen that Task 2 had the highest error rate with a mean of 0.60 errors made by each participant, while task 6, 7 and 8 had an error rate of rate of 0 meaning no errors were incurred when participants were performing these tasks. From Table 4 participant 2, 4 and 5 had the highest error rate per participant this was found to be 0.25 errors. The least was an error rate of 0 per task (no errors). The mean error rate among all tasks and participants was found to be 0.42.
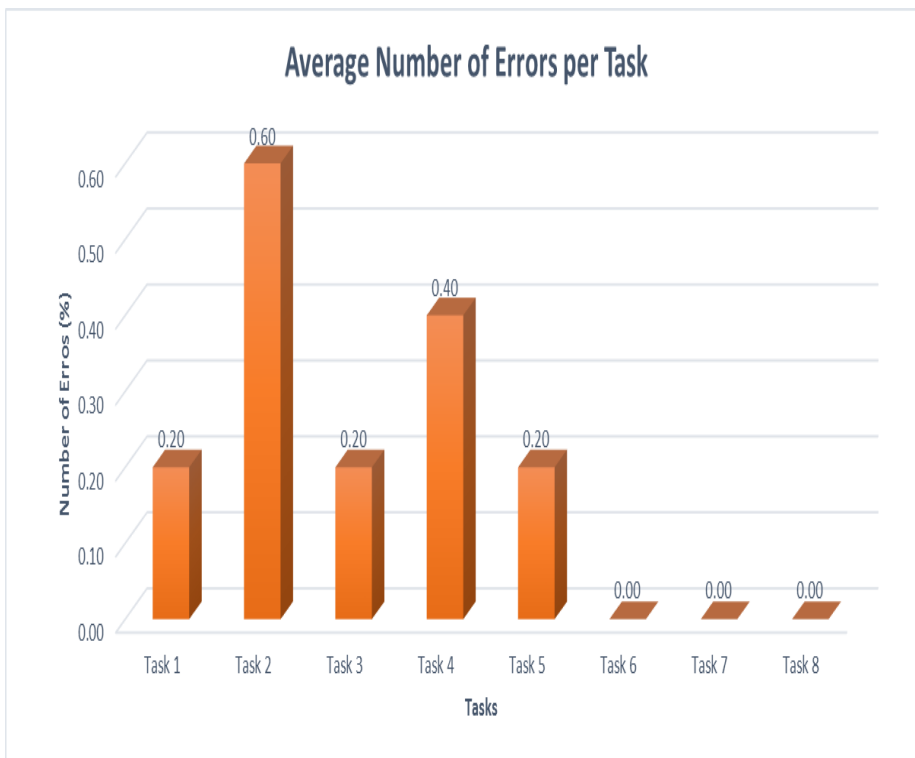
Figure 8: Graphical Representation of Task error rate

3.2.3.2. Qualitative Analysis

Questionnaire Responses	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Average in % (Question)	Standard Deviation
Q 1	80	80	60	40	40	60	20.00
Q 2	80	80	60	40	60	64	16.73
Q 3	80	80	80	60	40	68	17.89
Q 4	80	80	80	60	60	72	10.95
Q 5	80	80	80	40	40	64	21.91
Q 6	60	80	80	60	60	68	10.95
Q 7	60	80	60	60	60	64	8.94
Q 8	60	100	80	40	60	68	22.80
Q 9	40	60	0	80	40	44	29.66
Q 10	60	20	60	60	60	52	17.89
Average in % (Participant)	68	74	64	54	52	62.4	6.01

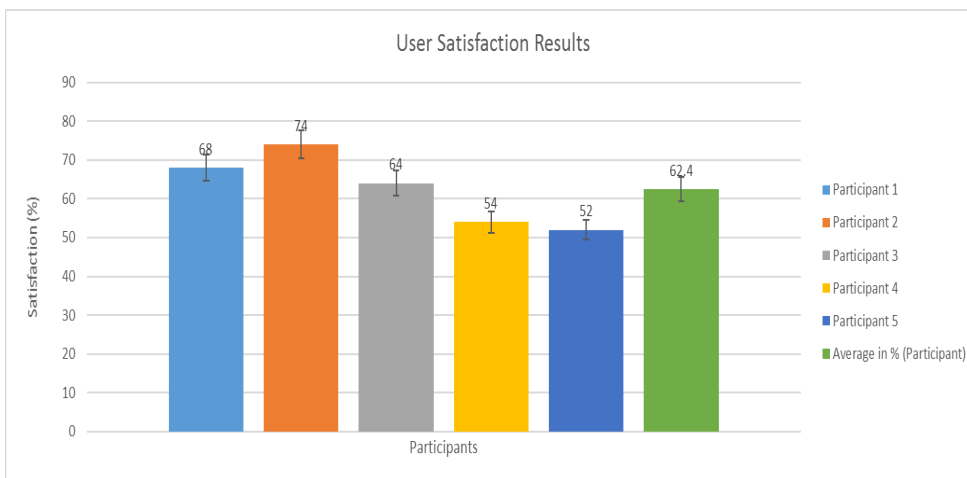**Table 5: User Satisfaction Data**

Figure 9: Graphical representation of user satisfaction per participant

This section provides a brief analysis of the data collected using PSSUQ, this questionnaire was distributed to participants after they have performed the given set of task. The PSSUQ consisted

of 10 different questions which are measured through a 5 point Likert scale. From the data represented in Table 5 the overall satisfaction of users after they have used the system was found to be 62.4% with an overall standard deviation of 6%. Based on the system usability scale (SUS) the mean overall user satisfaction value is below the standard SUS score which is 68% (Sauro, 2011). This, therefore, means the 62.4% percent is regarded as unacceptable and the general user satisfaction has a negative outcome.

3.2.4. Findings and recommendations

3.2.4.1. Findings

This section provides a list of findings and recommendations generated from the qualitative data and quantitative data gathered during the evaluation of the mobile app. As can be seen from the quantitative data in section 3.2.3.1.1 participants took the longest time to complete task 4. Two factors were identified as the main cause of the highest completion time, firstly users tried compiling feedback on submissions that have not been marked complete, secondly, users tried to email submissions without compiling feedback files.

Task 2 in section 3.2.3.1.1 was also listed as one of the tasks that were performed poorly by participants. In this task, participants were required to reduce a mark by removing a check mark. Participants kept on forgetting to activate the pen marking button to enable mark reduction. Moreover, most participants found the steps followed for reducing a mark tedious and confusing. Therefore, these usability issues were the major contributors to the high error rate for this task.

In task 1, two usability issues were identified which contributed to high error rate for this task, the first one was that participants kept on confusing the button for activating the highlighter tool with the button for activating the pen tool. Secondly, instead of swiping to allocate a mark, participants scrolled and this resulted in a mark not being allocated for a submission.

3.2.4.2. Recommendations

Based on the user feedback from the evaluation conducted with five participants, the researcher has decided to take these recommendations into consideration and the following changes have been adopted.

Formatted: Not All caps

Formatted: Not All caps

Formatted: Check spelling and grammar, Not All caps

Commented [KS((1]: Currently implementing

4. Summary

This report was divided into two sections, the first section (Improvements on the Previous Prototype) provided detailed information about the changes made on the prototype which was developed for the third report. Moreover, the section also provided details on how the changes made on the prototype developed for the previous report affected the system requirements.

The second section (Evaluate Artefact) of the report provided detailed information about how the evaluation for the Mobile assessment and annotation tool was evaluated and the equipment used to carry out the evaluation. The purpose of this evaluation was to determine if the final prototype meets the proposed system requirements. A maximum of eight tasks was given to 5 participants who are lecturers in the department of computing sciences at Nelson Mandela University and after performing the tasks the participants were asked to complete a post-study system usability questionnaire.

During the evaluation four metrics were measured for each participant, these are:

1. Time- taken to complete a task
2. Task completion rate
3. Number of errors per task
4. User satisfaction

The metrics above were grouped into two parts, namely, quantitative analysis and qualitative analysis. The quantitative analysis provided tabular and graphical results of the time taken to complete a task, task completion and the number of errors metric. The qualitative analysis metric provided tabular and graphical results of the user satisfaction metric. Lastly, this report provided a list of findings and recommendations acquired during evaluation.

4. References

Aspose, 2015. *File Formats and Conversions.* [Online]
Available at: https://docs.aspose.com/display/wordsjava/File+Formats+and+Conversions
[Accessed 15 August 2017].

Android, n.d. *Detecting Common Gestures.* [Online]
Available at: https://developer.android.com/training/gestures/detector.html
[Accessed 3 September 2017].

Canary Learning, 2016. *Software designed to make teachers' work easier.* [Online]
Available at: http://canarylearning.com/features/canarygrading/
[Accessed 21 September 2017].

Developer, P., 2014. *[Xamarin - Android] How can i use onScrollChanged and onScale.* [Online]
Available at: https://groups.google.com/forum/#!searchin/pdfnet-sdk/onscrollchanged%7Csort:relevance/pdfnet-sdk/jnjc1q4KNuI/tAFxix2VBTMJ
[Accessed 3 September 2017].

Johannesson, P. & Perjons, E., 2012. *A design Science Premier.* 1st ed. s.l.:s.n.

Kupelo, S. K., 2017. *Mobile Assesment and Annotation Tool V2,* Port Elizabeth: Nelson Mandela university. (Unpublished Honours Report 3).

Musfid, J., 2016. *Usability Testing Of Mobile Applications: A Step-By-Step Guide.* [Online]
Available at: https://usabilitygeek.com/usability-testing-mobile-applications/
[Accessed 01 October 2017].

NIST Industry USability Reporting project, 2001. *Common Industry Format for Usability Test,* s.l.: NIST Industry USability Reporting project.

NMMUa, 2016. *Corporate ID.* [Online]
Available at: http://corpid.nmmu.ac.za/

NMMUb, 2016. *General Prospectus for 2016.* [Online]
Available at:
http://www.nmmu.ac.za/www/media/Store/documents/apply/admission/quickdoc/prospectus/1-2016-General-Prospectus.pdf

PDFTron, n.d. *PDFTron Mobile PDF SDK.* [Online] Available at: http://www.pdftron.com/pdfnet/mobile/android_pdf_library.html [Accessed 2017 April 2017].

Rubin, J. & Chisnelll, D., 2008. *Handbook of usability testing: How to plan, design and concduct effective tests.* 2nd ed. Indianapolis: Wiley Publishing, Inc..

Usability.gov, 2017. *System Usability Scale (SUS).* [Online] Available at: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html [Accessed 02 October 2017].

Wildervanck, A. F., 2016. *Mobile Assessment & Annotation Tool.* Port Elizabeth: s.n.

Wilkinson, W., 2017. *Organization of a Traditional Academic Paper.* [Online] Available at: http://uncw.edu/ulc/documents/OrganizationofanAcademicPaper.pdf

Writing Guide, 2017. *Academic report/Thesis.* [Online] Available at: http://writingguide.se/writing-process/academicreport

Writing Skills, 2017. *The structure of a report.* [Online] Available at: http://libweb.surrey.ac.uk/library/skills/writing%20Skills%20Leicester/page_76.htm

Appendix A: Consent Form

Dear Lecturer

I have developed additional functionalities on the Mobile assessment and annotation tool app, which was developed in 2016. The app makes use of a digital tablet to overlay marking annotation onto online submissions and manages special notations made by the assessor while electronically assessing the online submissions. In order to evaluate the usefulness and usability of the new functionalities added onto the app, you are invited to participate in the assessment of the app at a time that suites you. Participation in this data collection activity is optional. The data collected as a result of this process will contribute to a report that will be submitted for assessment.

The study has received ethical clearance from the NMU Ethics Committee (reference number H17-SCI-CSS-002).

Should you have any queries regarding the study, please contact me, s212301195@nmmu.ac.za

What is expected of you should you choose to participate in the data collection process

Participation in this study will mean that you can be expected to complete the following activities:

1. Use the app to perform a given set of tasks. This will take a maximum of 15 minutes
2. Complete a questionnaire for a maximum of 5 minutes.

At no time will your identity be made public, nor will you be identifiable in the reporting of any of your responses. All reporting will be in a summarised fashion. You may at any time withdraw from responding to the surveys without any fear of reprisal.

There are no known risks to participating in this study.

What data will be collected

No personal data will be collected. The researchers wish to collect your opinion on the usability and usefulness of the app developed.

I agree to participate in the study (*please tick all the relevant boxes*).

☐ Yes   ☐ No (*many thanks for your time – no need to read further*)

☐ I understand that there are no risks to me as a participant in this study.

☐ I understand that no other personal data will form part of this study.

☐ No pressure was exerted on me to consent to participate.

☐ Participation in this study will not result in no additional cost to myself.

Full Name: _____

Signature: _____   Dated: _____

Appendix B: Questionnaire

NMMU Mobile Assessment and Annotation Tool Usability Evaluation

Approximate years of experience using paper-based marking: _____

Approximate years of experience using electronic-based marking: _____

	Strongly disagree				Strongly agree
1. Overall, I am satisfied with how easy it is to use this system.	1	2	3	4	5
2. It was simple to use this system.	1	2	3	4	5
3. I could effectively complete the tasks using this system.	1	2	3	4	5
4 I was able to complete the tasks and scenarios quickly using this system.	1	2	3	4	5
5. I was able to efficiently complete the tasks and scenarios using this system.	1	2	3	4	5
6. I felt comfortable using this system.	1	2	3	4	5
7. It was easy to learn to use this system.	1	2	3	4	5
8. I believe I could become productive quickly using this system.	1	2	3	4	5
9. The system gave error messages that clearly told me how to fix problems.	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system	1	2	3	4	5

Please list up to 3 things about the system that you liked / enjoyed and give reasons why.

_____

_____

_____

Please list up to 3 things about the system that you disliked / did not enjoy and give reasons why.

_____

_____

_____

Appendix C: Usability evaluation Task List

Mobile Assessment and Annotation Tool

Assessor/ Examiner

1. Allocate Marks

- Assess WRI 201 Semester test
- Activate the pen marking button (Button with check mark)
- Allocate full marks for the submission (use swiping gesture)
- Add extra mark by swiping on the screen

2. Reduce Marks by removing a Check Mark

- Select check mark
- Activate pen marking button
- Click delete button

3. Mark Student Answer as Incorrect

- Click on the alpha button
- Then draw a Symbol that looks like an alpha (α)
- Press back button on the tablet to finish marking

4. Email Feedback Files

- Open the navigation drawer
- Click on the compile submissions menu item
- Open the navigation drawer, click on Email feedback menu item. A dialogue box with modules to email will show, select the feedback files to email then click ok

Page **1** of **2**

Mobile Assessment and Annotation Tool

Moderator

1. Switch user type to moderator

- Click on the toggle button in the toolbar

2. Allocate Marks

- Change marking mode to submission mode
- Open a marked submission
- Activate the pen marking button
- Swipe on the screen to allocate a mark.

3. Reduce Marks by removing a Check Mark

- Select check mark
- Activate pen marking button
- Click delete button

4. Indicate that mark has not been awarded

- Click on the alpha button
- Then draw a Symbol that looks like an alpha (α)
- Press back button on the app to finish moderation
- When done exit the app

Page **2** of **2**