CHAPTER 6

TESTING AND RESULTS

In this chapter, the discussion will center on the usability testing of the CRIVE web site designed to improve the current National Archives of Malaysia web site. The participants were asked to compare the CRIVE web site with the existing National Archives web site in terms of its usability and subjective satisfaction. The results and lessons learnt from the testing will be discussed as well. In Section 6.1, the steps taken before conducting the usability testing are covered. Section 6.2 presents the testing being conducted. Hypothesis, participants, materials, procedures and problems encountered during the testing phase. Section 6.3 summarizes the findings while Section 6.4 discusses and analyzes the results.

6.1 PURPOSE, HYPOTHESIS AND VARIABLES

6.1.1 Purpose

The purpose of the testing was to compare the newly designed CRIVE interface with the existing National Archives of Malaysia web site. The testing wished to determine which interface was best for different tasks and also what interface features were preferred by the users.

6.1.2 Hypothesis

The hypothesis made was that the users of the CRIVE interface would execute tasks within shorter times compared to the tasks they were asked to complete

for the existing National Archives web site. It is assumed that the participants would be able to answer quantitative questions about the contents of a site more quickly when they are using the CRIVE web site. Finally, it is believed that the table of contents style interface would receive higher satisfaction ratings than the existing web page interface, and that user's would find the tasks easier to complete using the CRIVE interface.

6.1.3 Variables

The independent variable was the type of interfaces (the CRIVE interface and the existing site interface). The dependent variables were task execution times and subjective satisfaction.

To measure the task execution variable, the participants received tasks sequentially and wrote their answers on a sheet of paper before advancing to the next task. If the user answered wrong, they were asked to continue until they found the correct answer. Thus, errors were factored into the study as higher completion time. It may happen that the user may be unable to answer the question due to the impossibility of getting the right answer for that particular task. To measure the second variable, users filled out a user satisfaction questionnaire upon completion of the experiment.

6.2 TESTING

6.2.1 Participants

The CRIVE interface was tested by 14 participants, all of them were student volunteers and still pursuing their Masters degree in Computer Science at the Faculty of Computer Science and Information Technology, Universiti Malaya. From these 14 participants, only one of them is a male. There are four sessions of

the testing. Originally, the experiment was scheduled to be conducted in two sessions with 7 participants in each session. However, as not all participants were available at the scheduled time, thus, four sessions was held and the distribution of the testing is summarized in the Table 6.1. A graphical equivalence of this table is presented in Figure 6.1.

Table 6.1: Testing Distribution By Number of Participants Per Session

Session	Number of participants	Date
1	3	11.6.1999
2	6	12.6.1999
3	2	13.6.1999
4	3	14.6.1999

However, the testing results were not interrupted at all due to the changes in experiment scheduling. This is because once the participants finished doing their session's experiment, they were asked not to discuss anything related to the testing especially the testing questionnaire to the participants of the next sessions.
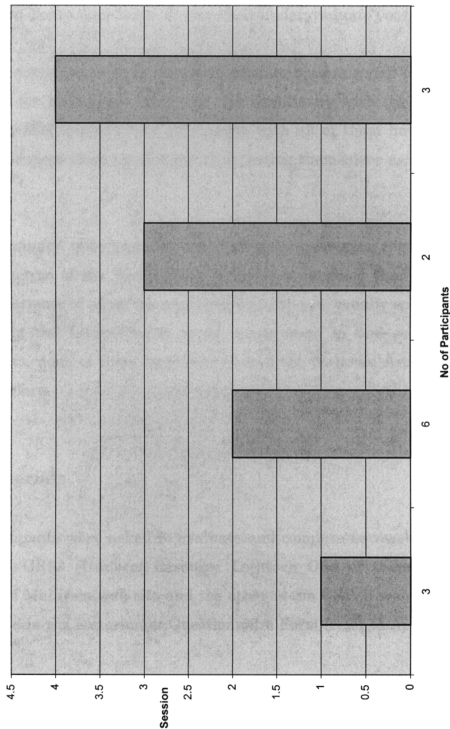
Distribution of Participants Per Session

Figure 6.1 The Distribution of Participants Per Session

The participants were asked to fill out an Experiment Consent Form (refer Appendix 1) and Preliminary Questionnaire (refer Appendix 2) before the experiment. From the questionnaire, it is determined that the participants were generally experienced with computers and the World Wide Web. This is not a surprise because all of the participants were either Computer Science or Information Technology major during their undergraduate years.

All of the participants were advanced computer users as all of them have used computers for more than five years. The familiarity with the World Wide Web was evenly distributed across treatments with all of them have been using the Internet for more than 2 years and thus, rating themselves as advanced users of the web.

The preliminary questionnaire asked several questions regarding experience with and usage of the World Wide Web. It is assumed that participants with longer experience of using the computer and the web usually spent more time per week using the Internet and spent longer time in one session. Of all the participants, none of them have ever visited the National Archives of Malaysia web site before.

6.2.2 Materials

The participants were asked to evaluate and compare two web sites. They were given two URLs (Uniform Resource Locator). One of them is the National Archives of Malaysia web site and the other is the CRIVE web site. These URLs were stated in the Experiment Questionnaire Form (refer to Appendix 3) given to them.

6.2.3 Procedures and Problems

The first step in conducting the experiment was to arrange a time and place for the sessions. The original plan was to have the testing in two sessions over the course of one afternoon with 7 participants per session. However, as not all of the 14 participants were available at the pre-scheduled time, a new arrangement was made. The new schedule have four sessions with the distribution of the participants is shown in the Table 6.1 above. Each session would test a same treatment of the independent variable. The experiment was held at the Workstation 4, Univesiti Tun Abdul Razak building. It is believed that this workstation would be a suitable testing area because there is sufficient number of computers for each session for training purposes and ergonomically comfortable environment. Furthermore, this workstation is reserved for the participants to do their works.

Prior to each experimental session, a number of computers for participants to use was set up. This setting up involved the following:

1. Start up MS-Internet Explorer 4.0 (IE4) and set the window size to the dimensions required for the experiment
2. Bookmark the experimental sites (the National Archives site and the CRIVE web site)

As soon as the participants ready, they were asked to sit at the computer that had been set up, and fill out the preliminary questionnaire. Once all participants were present for their session, the session started with an introduction and a training routine. The details of the training routine depended on the interface of the sites being tested, but each session consisted of a demonstration of all features of the interface. The participants could ask questions at any time during the training session. During the training, the tester performed a few examples of how the interfaces should be used which are similar in nature to the

experimental tasks. Guidance was provided to the participants at this stage and answered any remaining questions.

When all the participants were ready to begin, they started the experimental tasks. All of the questions require the subject to time the tasks they are doing. They will estimate the time by looking at their wrist-watch and record it in a piece of blank paper just before beginning executing each task. They then look back at their wrist-watch and record the time as they finished the task. The time measured here is not an exact time but rather are just average time. The next task is done accordingly as it is sequenced in the Experiment Questionnaire Form. As there are not too many participants in each session, thus no aides were needed.

The final stage of the experiment involved a short debriefing in which the participants were given a gratitude for their participation. The participants were asked to fill out a User Satisfaction Survey (refer Appendix 4) before they leave. After the experiment, the raw data was collected from each answer sheets for analysis purposes.

The approximate times for the various stages of the experiment are given below:

- Training and questionnaire – 10 minutes
- Experiment – 20 minutes
- Debriefing and survey – 5 minutes
- Total – 35 minutes

Apart from the difficulty of getting participants and free slots for them, other problems experienced are like one of the applet are not being displayed by the web browser and a very slow downloading of the web site from its server.

6.3 RESULTS

6.3.1 Task Completion Time

Table 6.2 displays the average time to complete each task for both web sites while Table 6.3 shows the average time taken to complete tasks for the CRIVE web site. Times were measured in seconds.

Both tables are presented in two separate line graphs as shown in Figure 6.2 and Figure 6.3.

Table 6.2 Time Taken To Complete Tasks for The National Archives Web Site

Task Number	Average Time Taken (second)
1	3
2	8
3	Not completed
4	Not completed
5	20

Table 6.3 Time Taken To Complete Tasks for The CRIVE Web Site

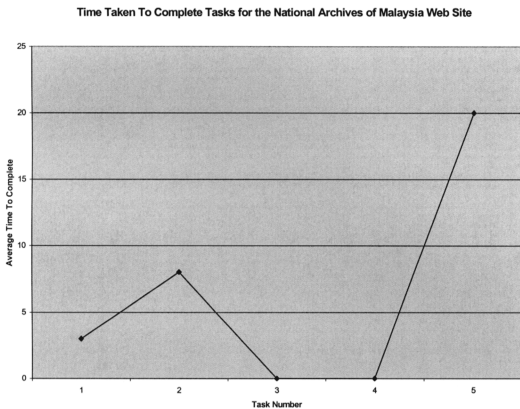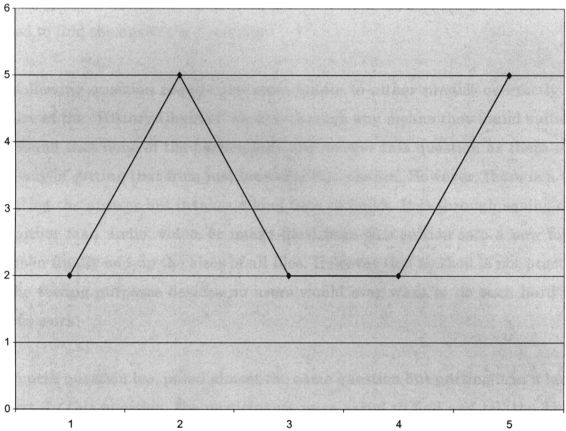| Task Number	Average Time Taken (second)
1	2
2	5
3	2
4	2
5	5

Figure 6.2 Time Taken To Complete Tasks for the National Archives of Malaysia

Time Taken To Complete Tasks for CRIVE Web Site

Figure 6.3 Time Taken To Complete Tasks for CRIVE Web Site

The National Archives Web Site

The first question posed in Table 1 in the Experiment Questionnaire Form is how many levels of hierarchy there is in the "Welcome" section. It is found that all participants can get the answer correctly within a very short time with the average of 3 seconds. They can do this just by looking at the web site's main page and clicking the hyperlink on the Welcome banner to get the answer.

The second question is asking the participants to name the section where the Memorial Tunku Abdul Rahman Putra is placed. This question is also quite easy to be answered. As the question implies the section name, thus the participants can very quickly answer by clicking on the Memorial hyperlink. The average time needed to find the answer is 8 seconds.

The following question require the participants to either predict or exactly say the size of the "History Channel" section through any means they found suitable. It is found that none of the participants can answer this question as there is no easy way of getting that from just browsing this section. However, there is a way of getting the answer but it takes a long time to finish. It is through saving each file (either text, audio, video, or image files) from this section into a new folder and then finally add up the sizes of all files. However this method is not practical for the testing purposes besides no users would ever want to do such hard and tedious work.

The fourth question too, posed almost the same question but putting it in a larger context. In this question, the participants were asked to find and tell the size of the whole site. Once again, none of the participants were able to produce any result. As in the previous question, the answer is possible to get but the time taken to calculate the size is just too much and impractical. The site also did not provide any way telling the size of the files or the collections that they have.

Finally, the participants were asked of how many file types they can find in this site. To be exact, there are only five file types available. They are: text file, Internet document file (HTML file), audio file, and image file. It seems that the site did not explicitly mention the file type but the users need to recognize the file type themselves. The average time taken in answering this question is 20 seconds.

The CRIVE Web Site

The questions for the CRIVE web site is posed in the Table 2 which can be found in the Experiment Questionnaire Form.

The first question asked the participants to find how many levels of hierarchy there is in the project web site. All participants answered correctly, that is 5 levels (of directories and subdirectories) hierarchy in the project web site. Average time taken to answer this question is just 2 seconds. The participants can quickly get the answer by referring at the bottom-most hyperlink at the left-hand frame. Once the user click this link, a figure showing a general overview of the whole site in a horizontal tree-like structure will be displayed on the right frame.

The second question was asking the participants to name the section where the "Malacca Sultanete" is placed. The participants do not have any difficulty finding their answer. On average, the participants take less than 5 seconds finding answer for this question. They use the approach of expanding each directories on the table of contents (on the left frame) until the nodes of each subdirectories being shown. From there, they can extract the answer directly.

The third question requested the participants to predict or tell in exact the size of the "History Channel" directory size. All the participants very amazingly answer this question very quickly with an average of 2 seconds. They can find the answer just by placing their mouse pointer over the "History Channel" directory and the answer is displayed at the status bar.

The fourth question also uses the same approach, and the participants can tell the size of the whole directory (or collections of documents) by placing their mouse pointer over the table of contents' National Archives Site Map label. The participants need only 2 seconds to answer the question.

The final question asks the participants to list the file or document types used in this project site. With the help of the Legend panel on the top-most applet at the left frame, the participants can answer the question within less than 5 seconds.

6.3.2 Subjective Satisfaction

The survey on user's Subjective Satisfaction posed six questions with the first question asking the participants to compare the difficulty level of using the two web sites being reviewed in this experiment. First question is divided into two with the first 5 tasks are based on the testing with the National Archives web site. The another five tasks are based on the CRIVE web site. The summary of findings from Subjective Satisfaction Survey is shown in Table 6.4 below.

Table 6.4 Subjective Satisfaction Results (1 for easy, 5 for difficult for Q1 and Q2; 1 for poor, 5 for excellent for Q3; 1 for not satisfied, 5 for satisfied for T6)

Question Number	National Archives Web Site	CRIVE Web Site
1 – Task 1	1	1
1 – Task 2	2	2
1 – Task 3	5	1
1 – Task 4	5	1
1 – Task 5	3	2
2	3	1
3	5	2
4	-	-
5	-	-
6	-	4

The result for the subjective satisfaction survey shows that for Table 1 (National Archives web site) tasks, Task 1 was considered easy by all participants. Task 2 was slightly easy when most users rated 2 in the continuum of 1 (the easiest) to 5 (the most difficult) to complete. All participants rated 5 or the most difficult for Task 3 and 4. In fact, the participants never completed these two tasks. For Task 5, the participants on average rated this assignment is moderately easy by rating it at scale 3.

For Table 2 (CRIVE web site) tasks , the users are evaluating the project's web site. For Task 1, the participants considered it to be easy by rating 1 on the rating scale. Task 2 is slightly easy when most participants rate the task in scale 2. A surprise is where 100% of the participants think that Task 4 and 5 to be the easiest to do. All of them rated 1 for these two tasks. Task 5 was rated 2 by most participants.

Question 2 asked about how easy it was for the participants to learn the interfaces they are using in this experiment. For the first web site, almost half of the participants rated it to be moderately difficult by rating it 4 on the easiness scale of 1 as the easiest. Another half of the participants rated it as moderately easy by giving a rate of 3. This may due to the reason of the difficulty of navigating back and through the web site, the lost of context once the users was already deep inside the site's document hierarchy. Sometimes, the site will launch a new web browser in order to display a new information which makes more active browsers available for users to choose from. Thus, this makes the participants become more irritated and dissatisfied. The same question is also asking the users to rate the difficulty level of using the project web site. Almost 80% of the users rated it as the easiest and another 20% rated 2 on the scale.

Question 3 asked of how well organized it is the web sites they are using. Not too surprisingly, many participants rated the CRIVE web site as more organized

compared to the National Archives web site. The project web site was highly rated as 5 in the scale with matrix of 1 as the poorest and 5 as the most excellent in site organization. National Archives web site was rated 2 on the site organization scale.

Question 4 is more general in nature. It asked of what the participants think of the idea related to the on-screen table of contents. All participants think that it was a great idea.

Question 5 was asking the opinion of the participants whether the table of contents help them to complete the tasks quicker than they would have if they are using a regular web browser and traditional navigation scheme. All participants once again agreed that the table of contents does help their tasks completion.

The final question asked participants general satisfaction with the new interface designed and developed for the project site. On the scale of 1 as not satisfied and 5 as satisfied, most participants rated the new interface in the scale rate of 4.

6.4 DISCUSSION

6.4.1 Discussion of Task Completion Times

Task 1 was a simple question asking the hierarchy levels of the both sites. For the National Archives web site, it is perhaps a little bit difficult to find the level of hierarchy without going to each section and click on them and later expand the scope from there on. However, for the CRIVE web site, it is very easy to get the hierarchy levels as there are two options available. The users can either use the View General Sitemap option found on the bottom-most of the left frame or they can expand the directories until they no longer can expand them. The nodes

represent the deepest hierarchy of the tree leaves and then the users can deduce the count the levels easily from it.

Task 2 for the National Archives web site is quite straightforward where the question asked where the Memorial Tunku Abdul Rahman Putra is located. The question itself implies the section name, thus it is not too difficult to answer the question. However, for an almost similar task for the CRIVE web site, it is not as straightforward. The participants were asked to find where the Malacca Sultanete collection is placed. As there is no such directory with such names found on the base directory, thus the option remains to the users is to expand the directories one by one and check through them to see the occurrence of the Malacca Sultanete, if any. Finally, the user will find it just right under the directory named The History Channel.

For Task 3 and Task 4 on the National Archives web site require the participants to find the size of "The History Channel" section and the whole site, respectively. It was expected that most of the users would find this task to be very difficult and impossible to do. There is no solution given or even hint given as there is no possible way of getting around it in a restricted amount of time. If the participants were given a leisure time to do this task, they might be able to find the sizes for both tasks. Extra works are needed where the subject must download each file from the National Archives web site and store them inside a new folder. After all the contents of the section was downloaded, then only the users can tell the size of the section. This applies to Task 4 as well. Task 3 and Task 4 for the CRIVE web site is easily conducted and the time taken to answer this question is very short. These two tasks can be accomplished by pointing the mouse pointer over the directory named The History Channel for Task 3 and the same procedure for Task 4. Once the mouse pointer was placed accordingly, the users will be presented not only with the directory size, but also other information such as the date of creation, the creator, and the parent directories to that particular directory. If the mouse is placed over a node or individual file,

then the filename will be given as well as the size of the file, date created, creator and parent directories. All these information is displayed in the browser status bar.

For Task 5 of both sites, it is quite confusing at first when the participants were all asking whether they have to navigate through the site to its final page in order to find if there are more files formats are to be found. Actually, both sites use not more than five file types as there is only HTML document, image files (in .gif and .jpg formats), downloadable file (in .zip format), and sound files (in .ra and .mid formats). There is no video file available from the National Archives web site. In the CRIVE web site, the only addition is the text file (in .txt format). The users have no difficulty finding these file types. For both sites, indeed the participants need to navigate through the site to find the answer. However, the CRIVE web site have a more advanced feature found in its control panel's Legend option. Even though there are many icons representing the file types, it does not mean that all of them are file formats. Some of them are just a symbol to represent the commonly used icon to represent file types in the MS-Explorer where many users have had experienced with.

6.4.2 Discussion of Subjective Satisfaction

Question 1 asked the users to rate the difficulty of each task. The question was intended to see whether there is a correlation between the task completion time and subjective difficulty of the tasks. There was a correlation between lower task completion time and lower difficulty subjective rating for Tasks 1, 2 and 5 which refers to Table 1 tasks. These tasks were meant to be easy, to get participants used to doing the tasks. Most of the users gave it a rating of 1, 2 or 3. The same thing also happening for Table 2 tasks which shows the same correlation pattern.

For Question 2, all participants think that the existing web site of National Archives is moderately easy and rated it on the scale of 3. Meanwhile, 100% of the participants think that the new user interface (CRIVE interface) is very good and easy to use. This may happen to their everyday use of MS-Explorer which have almost the same look and feel.

For Question 3, on average, the users feel that the CRIVE web site's interface is good in its organization. They rate this feature on the scale 4 which is very near to the excellent level. It is indeed that the new user interface which uses the TreeApp Navigator applet does give a better structure to the site compare to the textual sitemap provided by the existing web site.

Question 4 asked the users about the suitability of the on-screen table of contents for a web site. As all the users agree that it is suitable, the very reason is apparent. The users are exposed to hierarchical structure in their everyday life. Thus, they would also like to see such structure be used in the virtual environment as well.

Question 5 asks the user about their opinion of whether the table of contents would help them complete their tasks faster. All participants replied with yes which means that the table of contents does help them in that sense. This is due to the ability of the table of contents to display the document in another frame rather than launching a new browser. The navigation from the parent to nodes is also very quick as the expand and contract mode is supported. This feature eliminates the need for clicking on many hyperlinks and pages before reaching the final node.

In overall, all participants were satisfied with the new user interface and responded affirmatively. This indicates that interface such as the one designed for this thesis is in demand by users, and is motivation for further research in the area.