

Usability Metrics Study

Michael Fraser

September 26, 2013

Contents

1	Usability Metrics	3
1.1	Introduction	3
1.2	Self-Evaluations Pre-Experiment	4
1.3	Process	4
1.4	Results	5
1.4.1	Task 1: Find a File	5
1.4.2	Task 2: Connecting to Wifi	7
1.4.3	Task 3: Opening the Native Calculator Application . .	7
1.5	Satisfaction Survey Results	11
1.6	Judgement	13
2	Heuristic Evaluation	14
2.1	Task 1: Finding a file	14
2.2	Task 2: Connecting to a WiFi Network	14
2.3	Task 3: Opening the Native Calculator Application	15
3	Redesigning the Test	16
3.1	Choice of Hardware	16
3.2	Test Subject Experience Levels	16
3.3	The Testing Structure	17
3.3.1	Task 1: Find a File	17
3.3.2	Task 3: Opening the Native Calculator Application . .	17
3.4	Documenting the Results	17

List of Tables

1.1	Average Self-Evaluated Proficiency Pre-Experiment.	4
1.2	User Performance for Finding a File on Snow Leopard.	7
1.3	User Performance for Finding a File on Windows 7.	8
1.4	User Performance for Finding a File on Fedora.	8
1.5	User Performance for Connecting to Wifi on Snow Leopard. . .	9
1.6	User Performance for Connecting to Wifi on Windows 7. . . .	9
1.7	User Performance for Connecting to Wifi on Fedora.	10
1.8	User Performance for Opening Native Calculator on Snow Leopard.	11
1.9	User Performance for Opening Native Calculator on Windows 7.	12
1.10	User Performance for Opening Native Calculator on Fedora. .	12
1.11	Satisfaction Survey Averages for Each Task and System Com- bination.	12

Chapter 1

Usability Metrics

1.1 Introduction

The object of this study is to collect usability measurements and assess how well the tested systems comply with their associated guidelines document. There are three systems being tested and three tasks are to be performed on each system. The three systems being tested are:

1. the Windows 7 operating system,
2. the Mac OS X Snow Leopard operating system, and
3. the Fedora (Linux) operating system.

The three tasks being performed are accomplishable on each system, and are not necessarily more accessible on any one system. The three tasks are:

1. find a file,
2. connect to a Wireless Network, and
3. open the native calculator application.

Ten different individuals were asked to participate in this study. Most of the individuals are of the same background: senior Electrical Engineering majors. Three individuals have different backgrounds: one is a senior Mechanical Engineering major, one is a senior Computer Science major, and one is studying to become an elementary school teacher. The heavy focus on Engineering may skew the data, as engineering majors require a certain level of technological proficiency in their studies.

Snow Leopard	Windows 7	Fedora
6.4	6.8	3.7

Table 1.1: Average Self-Evaluated Proficiency Pre-Experiment.

The usability metrics that are being reported are Learnability, Efficiency, and Satisfaction. None of the tested individuals had interacted with Fedora before, and only a few had ever interacted with Linux operating systems at all. All the users have had some interaction with Mac and Windows operating systems before. This creates an imbalance in the data collected: measuring learnability for the Fedora system tests and measuring efficiency for the Mac and Windows tests.

1.2 Self-Evaluations Pre-Experiment

Before running the experiments, each individual was asked to judge their own proficiency with each of the three systems. These values were averaged according to operating system and the resulting data can be found in Table 1.1. As expected, the majority of the individuals were more comfortable with the more common operating systems: Mac OSX and Windows. Even though the majority of the people had never seen or interacted with Fedora before, the users believed they could at least figure it out. The individual’s experiences with Snow Leopard and Windows 7 would be interactions measured with efficiency, whereas experiences with Fedora interactions would be measured as learnability.

1.3 Process

Each individual was asked to sit down in front of three computers: a MacBook running OS X, a desktop PC running Windows 7, and a laptop PC running Fedora. The variance in physical design between the three computers also challenges the validity of this study, as the design of the computer may influence the metrics being observed.

The experiment was run with tasks being performed on Mac OSX Snow Leopard, then Windows 7, and then Fedora. This was later considered to be biasing the later tested systems, as the user begins to expect the next task for the later systems. The user was always unaware of their tasks when first facing the Snow Leopard machine.

Task 1: Each individual was told that a file titled "Find ME.txt" was hidden somewhere in the file system. The first task was to find that file and open it. A timer would run as soon as the individual was told to start, and the timer would be stopped when the individual opened the file.

Task 2: Each individual was then asked to "Connect to the student wireless network." One of the machines used, the desktop PC, was unable to connect to wireless at all. The task was measuring the time it took the user to open the list of available networks instead of actually measuring the time it took to connect to the network. This eliminated the need of the desktop PC to actually connect to a wireless network.

Task 3: Each individual was then asked to open the system's native calculator application. They were told to start this task and were timed until the calculator was open or able to be used.

After the three tasks were completed on one system, the individuals were asked to move to the next system. For each task on each system, the individual's actions were observed. Notes were taken on the steps the individual took to accomplish a task, any accidental errors along the way, and how the user reacted to obstacles or errors. At the end of each task, if the individual struggled to find a method of accomplishing a task, they were shown some alternative ways to accomplish the task.

1.4 Results

1.4.1 Task 1: Find a File

The results for the first task, find a file, can be found in Table 1.2, Table 1.3, and Table 1.4.

Windows 7

From the averages of these three tables, it would appear that most individuals were either more comfortable with Windows 7 over the other systems, or the system was more efficient. A common occurrence was opening the start menu in the bottom right corner and searching the file name there. Another common occurrence was the opening of Windows Explorer and running a search on the whole C: drive, which led the individuals to try a different approach as this search takes longer than necessary.

When setting up the directory path to hide the text file, the text file was accidentally placed on the Desktop, where it was unnoticed by the testers. When a few of the ten individuals were asked to find the file, they merely looked up and opened it off the Desktop. At this point, several tests had already been run without the file location being noticed. To stay consistent, the file was left where it was. This does skew the test data though, as no searching was necessary in a few occurrences for the Windows 7 system as it was for the other systems.

Snow Leopard

The next best average occurred with the Snow Leopard operating system. The most common search method of the system was using the top right search bar to find the file. It was also noted that even when the file was found, the double-clicking functionality of the Mac laptop were faulty. Several individuals double-clicked the file, expecting it to open, and nothing would occur. Since this was a hardware issue, and not an issue with the operating system itself, these errors should not be considered when comparing the different systems. To counter this fault, the timer was stopped when the double-clicking on the file was noted to occur.

Finding a file on this system was met with the least number of errors, excluding typing errors.

Fedora

Each individual underwent a learning experience trying to work with Fedora. Only a few individuals had used Linux before, and only one individual used it as their primary operating system. None of those users had ever worked with Fedora, as was evident when they were told to find the text file. The desktop screen of Fedora has few icons, and is primarily blank except for the background image. Most individuals, when sat in front of Fedora, began to mouse around in search of anything to click on. Most attempted mousing to the top right corner, where the battery, wifi, sound, and profile settings icons can be found. The main access to the applications of Fedora are accessed by mousing to the top left corner, which each individual eventually managed to find. It was not very intuitive for the users to move their mouse to the top left corner, as there was little indication that doing so would have any effect.

Accessing the applications window and opening the File System was the most common step the individuals took in finding the text file. From there,

they searched the entire drive to find the folder.

User Number	Time to Perform Task(s)
1	20
2	37
3	37
4	28.5
5	13
6	6.5
7	28
8	23
9	17
10	13.5
Avg.	22.4

Table 1.2: User Performance for Finding a File on Snow Leopard.

1.4.2 Task 2: Connecting to Wifi

When users faced the task of attempting to connect to a wireless network, they each followed the same general steps. They immediately began looking for the WiFi icon in the toolbars of each operating system. The WiFi icon is extremely similar between the different systems, so the user's knew to select it to access the wireless network settings. The results for this task, connecting to wifi, can be found in Table 1.5, Table 1.6, and Table 1.7. The average time it took to bring up the list of available wireless networks was less than four seconds on every system. There were no significant differences in the way the users completed the task between the different operating systems.

1.4.3 Task 3: Opening the Native Calculator Application

For the task of opening the calculator application, there were significant differences depending on the operating system.

Windows 7

When the users tried to find the calculator on the Windows system, most followed an expected path. They began by opening up the menu system in the bottom left and searched for the calculator application in the accessories

User Number	Time to Perform Task
1	11
2	39
3	37
4	5
5	1.5
6	7
7	37.5
8	6
9	4.5
10	2
Avg.	15

Table 1.3: User Performance for Finding a File on Windows 7.

User Number	Time to Perform Task
1	128
2	44
3	43
4	13.5
5	111
6	22.8
7	11.5
8	65
9	45.5
10	30
Avg.	51.4

Table 1.4: User Performance for Finding a File on Fedora.

User Number	Time to Perform Task
1	1.5
2	3.5
3	3.5
4	3.5
5	1.5
6	2.5
7	2.5
8	15.5
9	1
10	1.5
Avg.	3.3

Table 1.5: User Performance for Connecting to Wifi on Snow Leopard.

User Number	Time to Perform Task
1	1.5
2	3.5
3	3.5
4	3.5
5	1.5
6	2.5
7	2.5
8	15.5
9	1
10	1.5
Avg.	3.7

Table 1.6: User Performance for Connecting to Wifi on Windows 7.

User Number	Time to Perform Task
1	2
2	2
3	2
4	3
5	1
6	3.5
7	2
8	1.5
9	2
10	1.5
Avg.	2.1

Table 1.7: User Performance for Connecting to Wifi on Fedora.

subfolder. While most users knew of this path, likely from experience, it was not the most efficient method. A couple of users actually used the search bar in the start menu, having the system search for the calculator application and selecting it when the system found it. This method proved to be quicker than using the menu system. The data collected specific to this task on the Windows system can be found in Table 1.9.

Snow Leopard

For finding the calculator application on the Snow Leopard system, it was clear how experience with the system directly affected user efficiency. Users that had specified they were familiar with the Snow Leopard system knew about a calculator widget that is kept in a dashboard window. Those users simply used a keyboard shortcut to pull up the dashboard and the calculator widget was loaded. For users that were not as familiar with Mac systems, the typical method of finding a calculator was similar to how the users found the calculator in the Windows 7 system. The users would search through the menu system to find the applications list, and select the calculator from there. Several users who claimed to be unfamiliar with the system also used the search bar in the top right corner, as they found it easier to locate than the applications list. The results of the calculator application task with the Snow Leopard system can be found in Table 1.8.

User Number	Time to Perform Task
1	4
2	12.5
3	44
4	11
5	0.5
6	2.5
7	5.5
8	26.5
9	7
10	7
Avg.	12.1

Table 1.8: User Performance for Opening Native Calculator on Snow Leopard.

Fedora

When users attempted using Fedora to find a calculator application, they were unable to easily find the application. Users were at a loss as to where the application list could be found, and were searching for it for the majority of the timing period of the task. Only by moving the mouse around to random corners of the screen, and clicking on most icons they could see, were the users able to find the application list and then the calculator application. None of the users were really experienced with Fedora, let alone Linux, and so this experience was entirely new to them. If the users had been experienced, they would have likely been able to find the calculator application in times comparable to the other systems. The results of the calculator application task with Fedora can be found in Table 1.10.

1.5 Satisfaction Survey Results

After the different tasks were performed by each individual, the individual was asked to rate their satisfaction levels with each task performed on each system. The average satisfaction level for each combination of task and system can be found in Table 1.11. From these averages, using the Snow Leopard system seemed to be the most satisfying, while using Fedora was somewhat unsatisfying.

User Number	Time to Perform Task
1	5
2	7
3	4
4	4
5	5.5
6	4
7	5
8	6
9	6
10	3
Avg.	5

Table 1.9: User Performance for Opening Native Calculator on Windows 7.

User Number	Time to Perform Task
1	77
2	74
3	39
4	24.5
5	57
6	34.5
7	23.5
8	40.5
9	8.5
10	9.5
Avg.	39

Table 1.10: User Performance for Opening Native Calculator on Fedora.

Mac Task 1	Mac Task 2	Mac Task 3
8.5	8.5	7.6
Windows Task 1	Windows Task 2	Windows Task 3
7.6	8.2	7.2
Linux Task 1	Linux Task 2	Linux Task 3
6.5	8	5.8

Table 1.11: Satisfaction Survey Averages for Each Task and System Combination.

1.6 Judgement

Mac had the highest overall satisfaction ratings. Windows 7 quickest to open calculator on average. All systems accessed the wireless network list with similar levels of efficiency. Windows 7 had the best average time to find a file as well.

For the finding a file task, high efficiency is a must. Taking long periods of time to find a file lowers the satisfaction an individual would have towards that system.

For connecting to wifi, consistency with other systems is very apparent. It is also necessary for this step to be very efficient and error free, so users may switch between networks easily.

For opening the native calculator application, low errors and decent efficiency are necessary. Errors that could cause the wrong program to be opened would also lower the satisfaction with the system, and takes additional time to close the program and find the calculator application again.

From the results of the study, Windows 7 appears to have the most efficient design. The majority of individuals have used Windows before, and this experience may have skewed the data to favor the system. More people appear to enjoy using the Snow Leopard system, however. This is due to a significant decrease in errors noted in each task performed.

Chapter 2

Heuristic Evaluation

When the systems being compared are different operating systems, it is impossible to leave out user preference when considering efficiency. The tests were conducted with a limiting assumption that the individuals would have knowledge of how to use the different systems, when in actuality most people tend to stick to one operating system. Windows 7 users tend to know how to use Windows 7 more than they know how to use Snow Leopard or Fedora, just as Snow Leopard and Fedora users have greater knowledge of those systems than the others. The tests themselves allowed for additional error in data collection, as the hardware was different or the same systems were tested in the same order each time. Despite these errors, there are some noticeable design choices by the various systems that allow users of all operating systems to feel comfortable.

2.1 Task 1: Finding a file

For finding a file, most users attempted to search for the file as was expected. Using a search method for finding a file is convenient and efficient for users, as they do not have to manually search for a file. It is also a very common method for searches, such as searches on the internet.

2.2 Task 2: Connecting to a WiFi Network

The most significant design choice noted was the use of a similar icon for the wifi list selection. In each system, an icon representing signal strength of a wireless network were clear indications that the network settings could be modified by clicking that icon. In each system, that icon was accessible

from the desktop. The icon allowed users to see the relation to WiFi and networking without the operating system having to explicitly state what the icon's purpose was.

Each of the tested systems have been modified and modified over several versions. Windows 7 grew off of Windows Vista, XP, 2000, etc. just like Fedora and Snow Leopard each had versions before them. Each system has also had other operating systems that were developed and released before they were. With today's society, the idea of wireless strength is very commonly known. From mobile cell phones to wireless signal strength, the common user has an idea built in from interactions with previous versions that a wireless signal strength icon looks like a dot with ever-expanding lines emitting from it. When these systems were developed, the signal strength icon was kept consistent with previous systems, as that is what the common user recognizes. Easy access for the wireless selection list is also a must, as moving a computer from one location to another will allow it to view many connections easily.

2.3 Task 3: Opening the Native Calculator Application

When users tried to find the calculator application, the majority used the menu system to find the application on Windows 7 and Snow Leopard. The menu system allows the users to easily find the application, and the menu system was usually easily accessible. Some users even used a search option for this task as well as the first task, which shows that the operating systems allow multiple methods of accomplishing the same task. The multiple options to do any one thing illustrate the operating system's design being customer oriented, and add to the satisfaction of the user.

Chapter 3

Redesigning the Test

The previously discussed tests were somewhat flawed, but with improvements the tests could provide better data in relation to usability metrics. The major aspects that need to be accounted for in the updated tests are the choice of hardware, the test subjects experiences levels, the testing structure, and the documentation of results.

3.1 Choice of Hardware

In the previous test, a Dell Desktop was used for the Windows 7 system, a Dell laptop was used for the Fedora system, and a Macbook was used for the Snow Leopard system. To eliminate hardware factors, the test would ideally be done with systems of similar hardware. If each operating system were run on laptops only, and no external mice devices or other hardware were added, the hardware used for the tests would be similar. For the same reason, each laptop should have similar hardware specifications, such as clock speed, disk space, etc. While the differences due to specifications will be minimal, ensuring the laptops have very similar hardware specifications would eliminate any differences due to hardware.

3.2 Test Subject Experience Levels

In the previous tests, the majority of the users came from a science background and were familiar with Windows 7. While these test subjects worked for a preliminary data collection about the comparison of systems, the subjects themselves bias the data. To test the metrics of learnability and efficiency, both experienced and inexperienced users for all three operating

systems will be necessary. The experienced users with each system will be able to test the efficiency of a system, as they know the different methods to accomplish a task. The inexperienced users with each system will be able to test the learnability of a system, as they will not know how to accomplish the tasks given to them. Another way to improve the accuracy of the collected data would be to have more test subjects included overall.

3.3 The Testing Structure

The way the tests were set up previously biased different systems and therefore biased the test data. For the new set of tests, each task must be set up in the same manner. The second task, connecting to a WiFi network, was limited in the number of errors in the setup. The second task in the new tests will be setup in a similar manner as before.

3.3.1 Task 1: Find a File

There were some significant errors with the Find a File task in the previous test, the biggest error being that the file that needed to be found was accidentally placed on the Desktop for the Windows 7 system. The placement of the file made it possible for users to simply point at the file on the Desktop without needing to search for the file in any other manner. Attempts were made to eliminate similarities in directory architecture when the test was originally being set up, so users would not know the exact location of the file when moving to a second operating system. While this idea may be viable, the general structure of the architecture should stay the same, with the only difference being the name of the folders.

3.3.2 Task 3: Opening the Native Calculator Application

For the task of opening the calculator application, the main error stemmed from user expertise levels with the different systems. Any data collected about the efficiency of the systems were inaccurate. Otherwise, the calculator task was set up and tested decently.

3.4 Documenting the Results

When the tests were being run, any noticable actions by the users were marked down, such as errors and their method of accomplishing a task. The

results were placed into tables of the averages and notes were discussed previously. During the testing process, no discussions with the user were done though, that would have provided insight into what the user was thinking as they performed the tasks. No images were taken during the study either, which would illustrate the test process.