Andrew Sullivan

11/5/13

CMSI 370

Assignment 1105

Evaluate the Evaluator:

The *Mobile OS User Experience Shootout* from Pfeiffer Consulting

*Introduction*

In this paper I will be looking at the usability study conducted by Pfeiffer Consulting titled: *How iOS 7 Stacks Up: Smartphone OS User Experience Shootout*. This study analyzed in depth the new iOS 7 released for the iPhones. It looks at several usability metrics and compares them to the four leading competitors’ operating systems on the market. These include Android (Samsung), Windows 8, Blackberry 10, and the previous iOS, IOS 6. Each test was designed and measured with non-technical users in mind for casual, non-professional use. The usability metrics that are tested for each operating system are cognitive load, efficiency, customization, and user experience friction. Pfeiffer says that the benchmarks being measured are “based on the Pfeiffer Consulting Methodology for User Experience Quantification”, which is a system that they developed to measure these usability metrics.

*Background*

When designing a user interface, such as one for a phone, the key is for the designer to convey his or her mental model of the system to the user. When done properly, this can lead to a very intuitive and well-designed user interface. When done poorly this often leads to confusion, inefficiency, and overall user dissatisfaction in the system. The report written by Pfeiffer Consulting analyzes several user interfaces, and attempts to rate their mental model and therefore their usability.

When studying the usability of a device, there are many usability metrics that can be used to test this. A usability metric is anything that tests how “usable” a system is. The goal is to quantitatively record how well a system performs its intended task. According to the International Standards Organization (ISO) standard 9241, the best usability metrics are effectiveness, efficiency, and satisfaction. According to Nielsen, there are five usability metrics, which are the most important. These are learnability, efficiency, memorability, errors, and satisfaction. Each of the metrics used in the Pfeiffer report can be connected to one of these five metrics, however many are changed in one way or another.

*Methods*

The first metric, cognitive load, is “the sum of elements you need to get familiar with in order to use a device spontaneously and intuitively” (Pfeiffer). For this metric they counted the number of apps, widgets, icons, or any user interface element that is pre-installed into the system. These elements are things that the user must learn how to use. The assumption made by Pfeiffer is that the more of these elements that there are, the worse it is for the user because it means that they must learn more before they can use the device properly. It may be unfair, however, to give an OS a bad rating based solely on cognitive load. One OS may have more elements to learn than another, but if they are designed in such a way that they are more user-friendly or efficient then they might have a higher level of user satisfaction. It is not a good idea to measure cognitive load simply by counting the number of elements. This way it completely ignores how these elements may work together and whether they are presented clearly or not. It would be much more effective to measure cognitive load in time, seeing how long it takes a user to use all of the new elements, as apposed to how many there are. This is the reason why they must test multiple usability metrics and this is pointed out in their report.

For this test iOS 7 received a score of 40, meaning that there are 40 different elements that the user needs to learn. This tied with Windows Phone 8 at 40. Blackberry 1o received 53 and Android received the worst at 162. iOS 6 got a score of 32 however. This means that the new OS has implemented new elements into its system. Pfeiffer addresses this saying that “iOS 7 is slightly less streamlined than the previous version of iOS. This difference is due to the addition of the Control Center, a new user interface element with a new set of icons that were not present in iOS before.” This is an interesting feature to point out, however during the examination of iOS 6, they point out that it “lags behind in terms of ease of use functionally.” Although iOS 7 may have a worse score in terms of cognitive load, Pfeiffer seems to side with the point that even with that, it is still more functional.

The second metric, efficiency, was measured by analyzing “access to key settings, integration with notifications, multitasking, and camera access, among others” (Pfeiffer). The unit of measurement for this was a scale from 1 to 10 of how efficient they thought each OS was. They rated each system by the amount of options each system has and also “the ease of discovery for a non-technical user, as well as quality implementation” (Pfeiffer). Windows Phone 8 received a score of 4, Blackberry 10 received 5, and Android received 7. iOS 6 scored a 6 out of 10 and iOS 7 scored a 7 out of 10. The reason for iOS 7 scoring higher than iOS 6 is very similar to the analysis of the previous metric. iOS 7 is more efficient due to its new elements that were not in iOS 6. So, although iOS 7 has a higher cognitive load, it does make the system more efficient. In order to balance these two metrics I think that it would have been good to include learnability as a usability metric, which would give a better overall understanding of how the cognitive load and efficiency play together. I also feel that using a count as the unit of measurement does not properly evaluate a system’s efficiency. The traditional unit for efficiency, time, gives a much better idea of how efficient an interface is by timing exactly how long it takes to perform certain tasks. I feel that time is a much more concrete way of measuring efficiency as opposed to a 1 – 10 scale. This is pointed out by Brian Fung in a criticism of iOS 7. He points out that although there are more shortcuts implemented in the operating system to more quickly change settings, it is even more difficult to get this done due to the new simplified design. “The icons look vaguely like their predecessors but rather than asserting themselves, their more abstracted appearance means they actually need the text to help explain what’ll happen next if you tap them” (Fung). So, although the new iOS 7 has more elements to access key settings, it may be take even longer to do, and thus be less efficient.

The next metric, customization, measures how customizable each device is. Pfeiffer says that “consumer-level customization is one for the key user experience aspects of connected digital devices,” giving it a large amount of importance in their final reports. Customization is directly related to user satisfaction. The mistake that Pfeiffer makes is that they assume that the more that the user can customize his or her device to match their needs/preferences, the more satisfied they would be with that product. Customization may lead to user satisfaction to some level, but it can also reach a point where too much customization may lead to user frustration while attempting to fine-tune their settings to be “just right”.

For this test Windows Phone 8 received the lowest score of 2. Blackberry received 4 and Android received 7. iOS 7 got a score of 6 out of 10 and iOS 6 scored 5 out of 10. Again iOS 7 scored one point higher than its predecessor iOS 6. Pfeiffer says that “iOS 7 offers similar customization options as the previous release, but adds dynamic type support, as well as comprehensive accessibility options.”

The last metric, User Experience Friction (UXF), is described by Pfeiffer as “the bad stuff, the aspects of a device that can annoy you in a niggling way, or, in extreme cases, drive you crazy. Basically, UXF occurs whenever a device does not do what you expect it to do – or lacks a key feature that should be available.” User Experience Friction is directly related to the usability metric of errors and even efficiency.

For the final test Windows Phone 8 scored the highest at 51. Blackberry 10 received a score of 38 and Android received 30. iOS 7 scored 17 while iOS 6 scored 14. This means that iOS 7 had more cases where something happened that was not expected or intended than iOS 6. This again may have been a result of the new added features and cognitive load. “While iOS 7 has added some useful efficiency features, some of the new additions also contribute to UXF. A good example is the Control Center: it is clearly useful, yet has the annoying habit of accidentally popping up” (Pfeiffer). In each of the tests, it seems that the iOS 7 and 6’s results are directly connected to the cognitive load. While iOS 7 has a higher cognitive load than iOS 6, which according to Pfeiffer is bad, these added features have created greater efficiency and customization. Pfeiffer addresses this saying that “these additions results in a slight increase of cognitive load over the previous release, albeit not at a point where they risk overwhelming even a casual users.”

*Conclusion*

At the end of the report they combined each metric to create each OS’s final overall score. The iOS 7 received a score of 73.25, the iOS 6 70, Android 57.25, Windows 8 47.25, and the Blackberry 10 received 56.37. iOS 7 scored much higher than its competitors with the runner up being iOS 6. I think that the four usability metrics that they used to review these OS’s were very good, and gave a very good idea of how each performs relative to one another. I do feel that they should have tested other metrics such as learnability however. Cognitive Load may have a lot of impact on the learnability of a device, however they mainly scored this metric based on the amount of items a user would need to learn, and not on how easy or difficult these items were to learn. This could have drastically changed the ratings in this category. Overall the Pfeiffer Report on the iOS 7 was very thorough and covered many important usability metrics, however they also left out many of the most important metrics as listed by the International Standards Organization (ISO) standard 9241 and by Nielsen. In addition, the way in which they measured many of the metrics was not clear, or was measured in a way that seemed to be much less effective than the traditional units of measurement.

*Abstract*

A consulting company by the name of Pfeiffer Consulting recently released a report titled: *How iOS 7 Stacks Up: Smartphone OS User Experience Shootout*. This study compares the newly released iOS 7 to other top operating systems on the market today. The article focuses on four usability metrics: Cognitive load, Efficiency, Customization, and User Experience Friction (UFX). Each operating system is reviewed and given a score or measurement under each metric. At the end the scores are totaled and we have a general overall score of each system. Pfeiffer used many strong usability metrics to test each system, however for a more in depth understanding of the overall quality it may have served them better to stick to the current consensus of usability metrics and units presented by Nielsen.

References

Fung, Brian. "Apple's Muddled Design Philosophy." *The Washington Post*. The

Washington Post, 19 Sept. 2013. Web. 2 Dec. 2013.

"The Pfeiffer Report." *The Pfeiffer Report Asking the Essential Questions about*

*Technology and Media BHow IOS 7 Stacks Upb Smartphone OS User Experience Shootout Comments*. N.p., n.d. Web. 05 Nov. 2013.