Evaluating efficiency of interactive notifications on mobile devices

Joakim Ljungren

Department of Computing Science
Umea University, Sweden
id12jln@cs.umu.se

1 Introduction

1.1 Definition of efficiency

It is important to perceive different associations of the word efficiency to properly grasp the content of this paper.

The term itself literally means the accomplishment of a task with minimum expenditure of time and effort[6]. This definition is accurate for the context of this research. However, in the perspective of usability testing, such as in this study, efficiency can also be seen as one aspect out of three that affects the usage of a interface. The other subcategories are then effectiveness and satisfaction [1].

Usability:

- Effectiveness
- Efficiency
- Satisfaction

This research aims to evaluate the efficiency part of the usability for a specific type of use case, which is interacting with notifications. A use case accordingly means the exchange of interaction between a person and a system for the user to achieve a goal. The actual use case of this explicit research inquiry follows in the next chapter. The reason behind choosing efficiency alone as the evaluation aspect is addressed in the Background chapter.

Effectiveness as a factor of usability means the accuracy and level of completion with which users reach a goal. Indicators of this are solution quality and error rates. The measure of effectiveness is therefore the outcome of the user interaction with a system [1].

Efficiency is very closely related to effectiveness but not quite about the same thing. In the usability perspective, efficiency can be referred to as the relation between the completeness of achieving a goal and the resources it took to get there. Task completion time and learning time are both indicators of efficiency. Thus, efficiency measurements are in terms of time [1].

Satisfaction is the part of Usability where the users subjective experience comes in. It is the qualitative factor covering what attitude users had towards using something and what their preference lies [1].

This study will use task completion time as an indicator of efficiency. Consequently, the time it takes for a user to complete a task.

1.2 Research ambition

There is today a common way of how user interaction with notifications on modern mobile devices work, mainly referring to smartphones. Which is that when user attention is demanded in a mobile application, a notification through the operating system is issued. This can in many cases be information that is easy for the user to directly consume from the notification, but if it requires some kind of response in user input, the application from which the notification came has to be opened. The application UI is opened for instance by tapping this section of the screen which is represented by the notification. The operating system interface, where the notification is managed, is therefore serving as a link to the application.

There is an alternative approach to this, where the application UI is not required to be opened. The user interaction with the notification event remains with the operating system (OS).

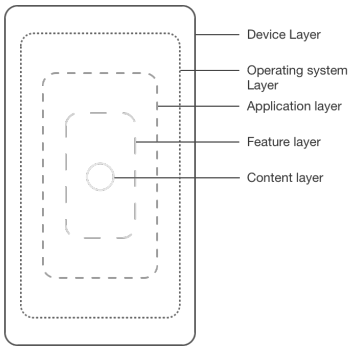

Fig. 1. The layers of a mobile user interface. Inspired by a figure of mobile infrastructure[1].

A user interface (UI) of a specific application can in a way be seen as a partial layer of the mobile interface as a whole. The parent layer in this perspective would be the operating system layer where the notifications appear. In the traditional described use case of response to a notification, it is necessary to open the application UI to interact with the actual application functionality. However, there are some examples of when the application functionality transcends the operating system layer.

[0] `http://www.dictionary.com/browse/efficiency?` A definition of the term efficiency, accessed 2016-10-18

[1] `https://blog.intercom.com/messaging-apps-just-getting-started/` The potential of conversational UI:s, accessed 2016-10-14

In the latest versions of modern popular mobile platforms like iOS and Android it is possible to directly respond to information inside the actual notifications[2]. Thus, interacting with the application from outside the application layer. This phenomenon is what is referred to as interactive notifications in this article. The current methods of interactive notifications on these different platforms vary a great deal and they have different flaws in usability.

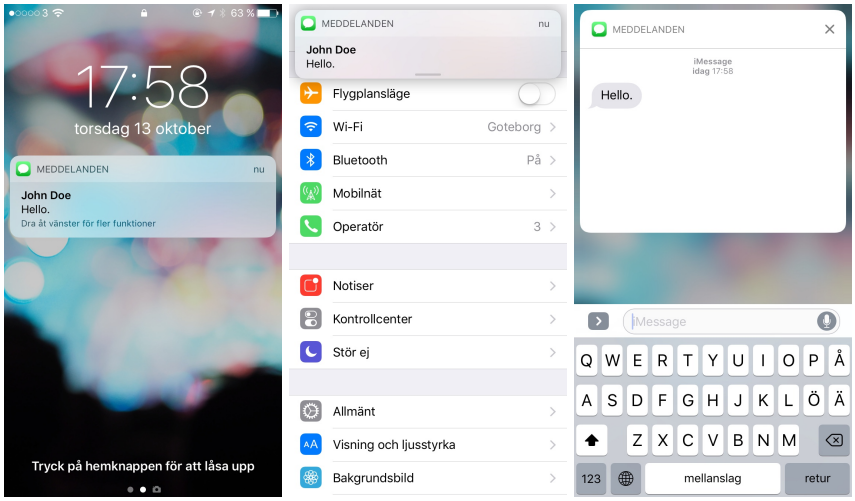

Fig. 2. iOS 10 example of interactive notifications with possibility to respond to a message by expanding the notification card rather than opening the application.

1.3 Hypothesis

The hypothesis is that the approach of direct response inside notifications is more efficient for the user to perform compared to the traditional approach of opening the application UI. This regarding messaging applications, by being able to respond to messages more effortlessly.

1.4 Purpose

This paper will try to decide if the stated hypothesis is actually true. It is expected so since it is something that is getting more common on mobile platforms and intuitively should be a more efficient option in interaction. But, it could be so that the user rather needs the greater context of the application UI to make quick decisions. There is a possibility that the full occupation of the

[2] `https://blog.intercom.com/the-end-of-apps-as-we-know-them/` The potential of notification interaction, accessed 2016-09-21

users attention-span on the screen is necessary for the user to process the specific content properly and being able to respond quickly. The occupation of several notifications that are available to expand in the same view could be distracting and thus the option of responding inside interactive notifications ends up less efficient. The method of opening the application UI could prove to give a better clarity of what system the user is interacting with and in that case maybe represents the faster interaction process compared to interactive notifications. The research question is therefore: *What is the most efficient option for responding to messages on a mobile device?*

2 Background

This section includes literature studies which are composed of related topics in user experience research and design of interfaces in mobile platforms. This earlier work has been taken into account to eliminate surrounding questions, to be more confident in chosen methods of testing and to pinpoint what is relevant.

It can be argued that multiple UI:s on your mobile device become less and less relevant and there are now a lot of third-party content that could be used in a more unified manner. We could be moving towards having one dynamic interface instead of one for each app[3]. Especially since the biggest technology conglomerates no longer are competing with one functionality, they want to engage in all of the users activity[4]. Clear examples of this could be Google Assistant[6], Google Now[5], Siri[7], Alexa[9], Cortana[8] and other artificially intelligent assistants or multi-tasking environments. This abstraction raises many questions of how the user experience will change and if it really will be all for the better. This is regardless of what type of interface that will rule in the future, like touch screens, voice control, gesture control, or something else. It is still a question about the general user experience. The research concerns this uncertainty as in questioning if users really prefer the unified and dynamic type of interface, but doing so relevant to what is happening today.

[3] `https://www.subtraction.com/2014/08/26/what-is-a-card/` Designing by "cards" or minimal UI:s, accessed 2016-09-21

[4] `http://avc.com/2014/05/app-constellations/` Article with discussion about how we in the future might have interfaces governed only by the big companies, accessed 2016-09-21

[5] `https://www.google.com/search/about/learn-more/now/` Google Now, a more unified interface for accessing all information, accessed 2016-10-24

[6] `https://assistant.google.com/` Google's AI, probably soon to be inseparable from Google Now, accessed 2016-10-24

[7] `http://www.apple.com/se/ios/siri/` Siri, an AI interface for Apple devices, accessed 2016-10-24

[8] `https://support.microsoft.com/sv-se/help/17214/windows-10-what-is` Cortana, Microsoft's counterpart of an AI assistant, accessed 2016-10-24

[9] `https://developer.amazon.com/alexa` Alexa, Amazon's intelligent assistant, accessed 2016-10-24

The significance of this study comes from comparing the usability of two methods in interacting with mobile interfaces and with notifications in particular. Usability can be seen to be dependant on three factors, efficiency, effectiveness and satisfaction. These conditions should be individually evaluated when measuring usability and assumptions about their correlation should be avoided. It is also important to being able to control the effectiveness of tasks to be able to measure the efficiency as a part of usability [1].

Since the real-world context of interacting with notifications on a mobile device is often done in stressful environments, on many different locations and in many contrasting situations, one could argue that the test of this research should be effected in a similar fashion. There is research that acknowledges this as a fact of usability studies on mobile devices [5] In spite of this, according to another evaluation of this statement, conducting a field test instead of a laboratory test when evaluating usability on mobile devices takes up more time and does not guarantee substantially different results [4]. However, it seems as the more similar the laboratory environment is to an actual environment, the smaller difference is there between the results of a field-type test and laboratory-type test [4]. It is also hard to know what environment that will be best suited for testing mobile interaction since mobile devices are as said used everywhere.

Possible framework for how in more detail to do the usability tests [7].

Possible source of design guidelines for the prototypes [2].

- Find more sources about general guidelines to mobile UI, the design should be be unbiased, but probably better if it is based on conclusions from some unprejudiced earlier work.
- Include further recognition about earlier related research and explain with this at hand in detail how I will perform my study.
- Think about including images and figures to better explain why it is relevant to the future of our mobile interfaces.

3 Method

This section describes the user testing in detail. It depicts the chosen methods and explains why the research was executed in the way it was. Tools and materials like software development programs and devices that were used was chosen out of project suitability and accessibility.

The study included two concept prototypes for testing two different approaches to mobile UI. The approaches represents contrasting ways of interacting with a messaging application. Two user scenarios were developed as a simple native application for a mobile device. The prototypes are interactive and represents a typical sequence of inputs and outputs for each case. These interfaces are close to one of a real operating system on a mobile device, when it comes to operating the specific interaction. These cases are about directly responding to a message inside a notification as opposed to opening the application and responding to a message there.

The prototypes used in the tests are meant to be of a more general state rather than being similar to either of the popular existing platforms, like Android or iOS. This was to avoid involving biased views of brands and to not mimic templates of commercial design guidelines. This could have affected the test results due to previous user habits and earlier phone usage. However, the hardware device used in the tests was an iPhone 6.

The prototype was implemented with React Native, a Javascript native application framework, using Xcode as a development environment on a Mac, developing for iOS.

For the reason of more credible test results, the research isolates a single aspect of the usability, efficiency. Efficiency is explicitly related to the hypothesis that interactive notifications saves time and it is therefore most relevant to the investigation. Efficiency is according to personal experience also more accessible to test than other factors of the user experience. The evaluation does not cover tests about the effectiveness or satisfaction of interactive notifications, seeing that this would be extensive work for the time-frame available. Also because it is more important to properly evaluate the most relevant aspect of the usability rather than assuming too much of the prospects.

The investigation has been narrowed down from the wide concept of notifications, as in focusing solely on conversational applications. One reason for this is how interactive notifications could be especially beneficial in the aspect of more quickly being able to reply to messages. Another part of it is how the trend of AI:s and many service applications are moving towards having more of a conversational UI[1]. It is because of this even more relevant to reflect upon how user interaction is handled in critical moments of demanding attention to a conversation.

With reference to earlier research about usability testing on mobile devices, the testing will be executed in a relatively closed environment with as similar characteristics to a real setting as possible [4].

3.1 Testing

User tests was performed in two iterations with different levels of detail in the prototypes, low fidelity (lo-fi) and high fidelity (hi-fi). The assignments in the user tests were the same, where the test with a lo-fi prototype was to prepare the hi-fi procedure.

In both iterations the users tried the two different versions of notification interaction and the time it took them to complete each assignment was measured. The first test with lo-fi prototypes were acted as a qualitative test and served the purpose of pilot testing. The lo-fi prototype consisted of UI mockups sketched on paper and pretended to be used as on a smartphone to see if the chosen methods for the test are reasonable.

The considered test in this paper is the one performed with a high fidelity prototype. Several of these quantitative tests was completed. The measurements where analyzed to see if they are correlating in some way. A comparison of all the

gathered data delivered a good foundation for a conclusion of what the answer to the research question is.

The procedure of both tests was the same in terms of intuitiveness and accomplishment.

The test person will receive a notification with a question in form of a text message. First, they get the mission of opening it inside the application UI and answer it there. The time it takes to complete this task is measured. They will then get a similar question with the other mission of instead expanding the notification and reply to it there. The time to complete this is also measured.

The questions will be of easy character that requires the same level of thought process to reply.

Low fidelity test

- Think about the test procedures (important). What intervals of interaction should be measured? What will the messages be and should they answer? One notification at the time or several? more to specify...
- Is it allowed/appropriate to time the test participants without their consent? since telling them about it might create stress.
- inform about things like number of test participants, gender and age distribution, et.c.
- explain what exactly the required tasks of the user were
- Explain in further detail the execution of the test
- Describe implications that arose

High fidelity test

- inform about things like number of test participants, gender and age distribution, et.c.
- Explain in further detail the execution of the test
- explain what exactly the required tasks of the user were
- Describe implications that arose

4 Results

- Show what the results were for each test
- Illustrate resulting data in tables or figures
- Consider if the results are trustworthy or if there was something that might have affected the outcome.
- Examine if the results reflect the hypothesis or not

5 Conclusion

- With the results in hand and circumstances acknowledged, how is the initial research question answered?
- What can be learned from this research?
- What insight will the world gain from my study?

References

[1] Erik Frøkjær, Morten Hertzum, and Kasper Hornbæk. Measuring usability: are effectiveness, efficiency, and satisfaction really correlated? In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 345–352. ACM, 2000.

Establishes that there was in their study no apparent correlation between efficiency, effectiveness and satisfaction in usability testing. It states that it is not valid to make assumptions about the overall usability from tests about one aspect of it in terms of complex tasks. Thus, that efficiency, effectiveness and satisfaction in those cases should be evaluated individually. It measures efficiency as task completion time.

[2] Jun Gong and Peter Tarasewich. Guidelines for handheld mobile device interface design. In *Proceedings of DSI 2004 Annual Meeting*, pages 3751–3756, 2004.

[3] Kasper Hornbæk. Current practice in measuring usability: Challenges to usability studies and research. *International journal of human-computer studies*, 64(2):79–102, 2006.

The paper evaluates methods of how to measure usability with the purpose of that the review will improve the background on how to perform such tests.

[4] Anne Kaikkonen, Aki Kekäläinen, Mihael Cankar, Titti Kallio, and Anu Kankainen. Usability testing of mobile applications: A comparison between laboratory and field testing. *Journal of Usability studies*, 1(1):4–16, 2005.

Conducts a study if user experience tests of mobile applications differ when executed in the different contexts, such as labratory testing versus field testing. It concludes that there is no sign of advantage for field testing when it comes to results of studying user interfaces and navigation issues, while the field method is supposed to demand more time. This study address more of the qualitative type of testing.

[5] Christian Monrad Nielsen, Michael Overgaard, Michael Bach Pedersen, Jan Stage, and Sigge Stenild. It's worth the hassle!: the added value of evaluating the usability of mobile systems in the field. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, pages 272–280. ACM, 2006.

(Recently found, have not read yet.)

[6] Ahmed Seffah, Mohammad Donyaee, Rex B Kline, and Harkirat K Padda. Usability measurement and metrics: A consolidated model. *Software Quality Journal*, 14(2):159–178, 2006.

The paper reviews existing standards of usability testing and forms from this a unified model of going about these tests, called the QUIM model.

[7] Dongsong Zhang and Boonlit Adipat. Challenges, methodologies, and issues in the usability testing of mobile applications. *International Journal of Human-Computer Interaction*, 18(3):293–308, 2005.

An article that takes several studies in consideration and suggests a specific framework for testing usability of software applications developed for mobile devices.