

Usability Testing of Mobile Applications Store: Purchase, Search and Reviews

Wilson Prata*, Claudia Renata Mont'Alvão, and Manuela Quaresma

LEUI, Pontifical Catholic University of Rio de Janeiro (PUC-Rio). Rio de Janeiro, Brazil
wilsonprata@gmail.com, {cmontalvao,mquaresma}@puc-rio.br

Abstract. The objective of this research was to evaluate the process of purchase, search and evaluation of apps in the two main mobile stores available in Brazil, Google Play Store and iPhone App Store. It takes place in Brazil, from September to October of 2012. Sixteen users were selected based on a socio-demographic profile. The results helped to evaluate the overall opinion and attitude of the Brazilian users regarding each system. The debriefing confirmed issues in both stores, some similar between them but others quite particular. It also helped to indentify some minor's preferences regarding each task in each store.

Keywords: Smartphone, App Store, Mobile Usability Test.

1 Introduction

Over the last few years, Smartphone had increased their presence in the every day life. Since they are devices with high technologies their development and adoption had first happen in the most advanced countries but over the popularization and mass adoption of this technology the price decline and other markets could benefit from this solution. It is possible to observe this grown when examining the expansion of this market; in 2012 the global Smartphone grew 42% comparing with the previous year [1].

Nowadays, Brazil is the 5th most important Smartphone market with 131% of mobile subscription rate [2]. Smartphone correspond to 26% of this market. Android devices have almost 40% of the market share and iPhone 7%. In 2015 Smartphone will represent 57% of local market [3].

One of the reasons to this mass adoption is the capability download and install third part applications, those apps are usually distributed by a specific content deploy solution called App Store. Although there are some researches related to the usage of mobile application stores, there are few concerning the singularities of the Brazilian users and market. Even when those researches are conducted, they usually follow a mass-market analysis approach in which the focus is in the economic findings, bypassing usability constrains.

* Corresponding author.

As results, many issues related to the Brazilian environment that could have impact on the usage and adoption of Smartphone and mobile app stores are completely ignored. Perform a local usability study help not only to confirm the problems that affect the majority of audience but it helps as well to identify the problems that concern the local implementation, which in theory is as good as the original one. Some of the issues that could emerge due to cultural constrains are: misunderstood of labels of actions and areas of the app stores, organization of the apps by category, search terms and result's expectation.

2 Problem

If the app stores do not follow the usability guidelines - or do not follow a User Center Design methodology - during their development and implementation, users may find it difficult to choose, acquire and evaluate applications in those solutions what could result in frustration. These difficulties occur when users have an expectation that was not met by the systems, either by a dissonance regarding the previous communication about the service or because users cannot develop the mental models expected by the developers of the system.

To evaluate if these difficult happen or not in Brazilian scenario, and if so, which are the most problematic aspect of it, a field research was conducted. The objective of this research was to evaluate the process of purchase, search and evaluation of apps in the two main mobile stores available in Brazil: Google Play Store and iPhone App Store. It takes place in Brazil, from September to October of 2012.

3 Methodology

The objective of this research was to evaluate the process of purchase, search and evaluation of apps in the two main mobile stores available in Brazil: Google Play Store and iPhone App Store. Those stores were selected because they were the most popular app stores available in Brazil during the period of research, as reported by some research institutes [4, 5] and confirmed by a field research [6].

Although search, acquisition and evaluation are not all the possible actions that can be performed in an app store, those were selected because of the relation between them and the relevance in the usage of the store. Limited to these tasks, the research would like to identify:

1. Critical points in the data inputs to accomplish those tasks.
2. Expectation and opinions regarding labeling, navigation, organization and search systems.
3. Relevance and efficiency of the evaluation process and the information gathered and published for user's experience.

Based on the research objectives a research technique was chosen. Usability Testing is an established technique among various evaluation methods regarding Human Computer-Interaction. A Usability Testing as a process that employs people who are

representative of a population, as participants in a test that evaluate whether a product meets or not a specific usability criteria [7].

However, some questions arose when the object of evaluation is the usability in a mobile device. Based on the considerations and comments made by the literature in the area, it is clear that it is not possible to assume that a task evaluation technique can be arranged in the same way on desktop and in a mobile device. There is also the inability to reproduce in the laboratory all the context implication necessary to task evaluation on mobile devices.

The equipment and procedure of a mobile Usability Testing is similar of the tests on desktop but there are some particularities. The test was a comparative test of two systems, based on the research objectives the session it was define the profile of the participants, the tasks and scenarios, the session structure and procedures. There were five tasks to be performed. To each one a scenario was developed. Those are the tasks that were developed:

- Task 1: download the application Angry Birds
- Task 2: evaluate the application
- Task 3: download an application from featured area
- Task 4: download a weather application
- Task 5: download a internet browser

However, those tasks only could not provide all the data needed to complete the research objective. To better identify the satisfaction on the stores, it was used the SUS questionnaire and to enable the user to expose what they like most and what they dislike in each store, an debriefing session was conducted in the end of the usage of each system.

The session started with the reception of the participant and explanation of the methodology. The participant was requested to fill the consent form and a questionnaire. Based on the questionnaire an interview was conducted to confirm if the user match the defined profile.

Those procedures are the same of a desktop Usability Test. However the recording equipment need some adaptation. It was used a laptop to record audio and video of the user interaction with the systems. The particularity of a mobile usability test take place at this point, since it is a touch screen technology and there is no screen capture software that can record the interaction in the app stores. To enable this record, instead of the screen capture, it was used an external camera. An acrylic base was designed and it was used for fixation of the device and the webcam during the test. The webcam was plugged on a Macbook Air and a free video capture software was used to record the session. The debriefing sessions were recorded using a Smartphone. All equipment, analog and digital, fit in a backpack, which enables the test to be done in different places.

A pilot test was performed before the final sessions. It was conducted with five participants, two males and three females; four of them aged between 22 and 28 years and the other one with more than 40 years. It was possible to identify that the session would take about 50 minutes each. Scenarios and tasks proved clear and quick understanding by all participants.

4 Results

4.1 Task completion

It was defined two levels of success for task completion. They could be completed without difficulty, with difficulty or not completed. It was defined that a task was completed without difficult if it was completed at once. Thus, the participants didn't need to follow a predetermined path; in fact they could choose any acquisition process available (e.g.: browse by categories, use the search tool, acquire the app from feature area, etc). This criterion was chosen because one of the research objectives was to identify the preferences of users in face of these possibilities. However, it was decided that once the participant had chosen a process and had defined a mental model of acquisition, it was expected that this model would be used until the end of the task. Any difficulty or hesitation within that process, indicate that the task was performed with some degree of difficulty.

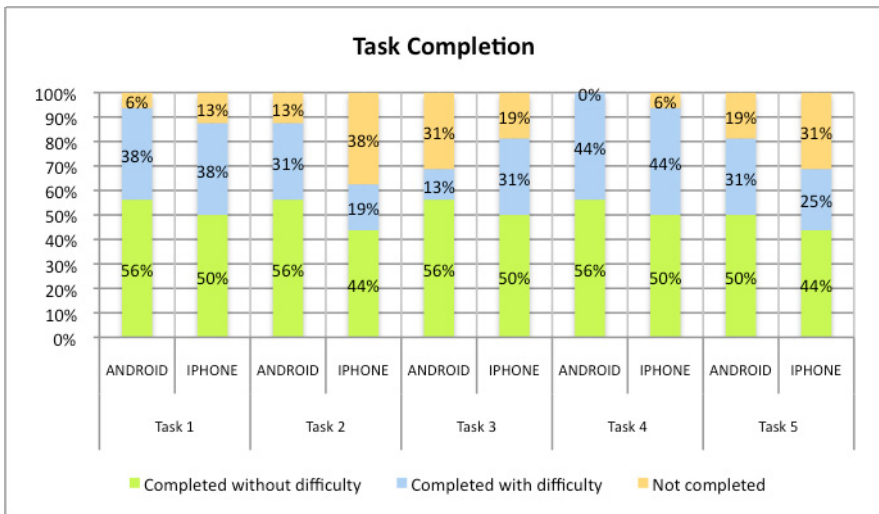


Fig. 1. Task Completion in Google Play Store and iPhone App Store

All tasks had an expressive number of tasks completed without difficulties but all of them had a percentage of non-completed tasks (Fig. 1). Google Play Store had a larger number of tasks completed without difficulties. The amount of tasks completed with difficulties was the same in both systems. The iPhone App Store got a superior percentage of tasks that was not completed. Regarding the user previous experience with the usage of apps stores, the participants with experience in the use of app stores had a better performance with Google Play Store. Participants without previous experience had a slight better result with the usage of the iPhone App Store.

4.2 Task 1: Download Angry Birds

A significant number of participants completed the task without difficulty in both systems. Those participants or had browsed in "games" category or had used the search tool. The majority preferred to use the search engine. There are two reasons why the participants completed the task with difficulty. The first one was the decision to change the acquisition method during the task performance. The second reason was because some participants ended up accessing another version of the application, but before installing the application they realize that it was not the correct version.

The reason of non-complete tasks on the Google Play Store was because one participant did not find the application in the category that he was browsing. This participant had no prior experience with app stores and also commented that he didn't used the search tool because the application name was in English and he did not feel comfortable to search for a term in other language. In the iPhone App Store, two participants did not complete the task they downloaded a different version of the application but they were sure that were downloading the one requested in the scenario. The use of the term "free" to indicate that it app version was free result in an additional problem to some participants since the term was not translated to Portuguese.

4.3 Task 2: Comment on Angry Birds

To perform this task 2 the participants should have had completed the task 1. Therefore, the task was presented only to the participants who completed the first one. Just one participant in iPhone and two participants in Android failed to complete it. The majority of participants in both stores used the same navigation model used in the previous task to complete the task. A few participants used a different model in Google Play Store; they accessed the application page using the options "My Applications" available on menu.

In Google Play Store the difficulties occurred due to separation of the reviews and the button that lunch the evaluation pop-up. Another problem was the interaction with the keyboard. When participants finished their comment, the keyboard overlaps the confirmation Button. Some users did not realized that they should have to press "back" to get out of edit mode.

In iPhone App Store the main difficulty was in realize in which part of the store the evaluation should be made. Participants tried to use a similar model as the one used in Google Play Store, they tried to find an equivalent to "My Applications", which was not possible. Another difficulty was to understand that the comment should have be made on the application page. There were few participants that didn't realize that it was mandatory to rank and to comment. After the system feedback about this obligation, the participants completed the task.

The participants understood that the labels in Google Play Store were clear than the ones in iPhone App store. Some participants were confused and uncertain about the completeness of the task in iPhone App Store since there is no feedback after evaluation. The participants who failed to perform the task were the ones that could not identify where the evaluation should have done.

4.4 Task 3: Download an Application from Featured Area

The featured area is usually the first area displayed at the store. This area has banners to promote some applications; these banners are larger than a standard snippet and not having a defined structure. The objective of this task is to confirm if this effort to advertising and promote some applications gather users attention.

The participants who completed the task quickly identified the featured area. In the Android system the participants downloaded the applications that were being promoted in first banners. Participants in iPhone hesitate to click on the banners and other users even noticed then, although the banners are bigger than the snippets. When comparing the banners of the two stores, it is possible to observe that the banners in Google Plat Store have more textual information.

Some participants who used the iPhone system completed the task with difficulty. The iPhone used in the test did not have the final version of it operational system; because of this some applications were not compatible and could not be installed. Many of these applications were highlighted in the featured area and when the participants tried to install one of these, a message appeared requesting the user to update the Smartphone. Other users had difficulty because they do not identified that the first screen was the featured area.

The users who fail to complete the task were those who downloaded applications in different area and didn't realize it. Many participants complain about the absence of indication of the category and price of the applications promoted on banners.

4.5 Task 4: Download a Weather Application

In this task the participant should have to download a weather forecast application. The objective of this task was to confirm which acquisition process is preferred by the user, if it was an exploratory navigation, the use of the search tool or browse in categories. The scenario was designed to minimize any suggestion to any of these search modes to make it clear which model among these the participant preferred.

That was the task with the highest success rate, only one participant failed to complete it. The participants made use of both the search and browsed in categories, but with prevalence of search. It was observed that the participants did not use the term informed in the scenario, they used synonyms in Portuguese. Most of the applications use those words in the application name. Related with this fact, the participants that complete the task with some difficult were the ones who had difficult to decide which app to download. In both systems the participants choose the app in the first page of the list of search results or in the first page of the category list, but they accessed multiple applications, comparing the images and description to determine which application download.

The participant who failed to complete the task on the iPhone had no prior experience with app stores. He tried to search of an application for weather forecast in featured area and there was no application of this type available there.

4.6 Task 5: Download an Internet Browser

The last task asks the participant to install a web browser. Browsers are specialized tools, but they are common tool for the participants based on the socio-demographic profile of these users. This task was the task with fewer percentage of completed without difficulty, and a significant number of participants failed to complete the task. Participants who were able to complete the task made use of the search tool and they had used the names of browsers for desktop as search terms, such as "Chrome" and "Opera".

Participants who did not complete the task tried to perform the task in the same way, but ended up downloading applications that were not browsers. One of the most downloaded apps downloaded by mistake was "Chrome to phone", it is an app that transferred to the phone's browser the bookmarks saved in "Chrome" for desktop. In the iPhone App Store, the same thing occurred, but the application were or "Browser secrets" or "Google search".

There participants who had difficulty completing the task tried to download a browser exploring categories, but none of them could between the available categories one that could have a browser. Two participants did not complete the task. These participants did not find the application in the categories and didn't try the search tool. All participants who completed the task, either with or without difficulty, had made use of the search tool to complete it.

4.7 Participants Comments, Opinions and Preferences

The usability problems regarding data entry in apps stores were most of then related to the keyboard configuration. Since there are a significant use of special characters in Portuguese's words (e.g.: é, ó, ê, ç) the participants found it difficult to type those characters, especially in iPhone. In both stores some participants complained about the obligation to type a title and a comment in the app comment session. Such division is seen as redundant and even retrograde. Another recurring complaint in the iPhone App Store was the obligation to insert the password to download free apps.

About the quantity and quality of information in iPhone App Store, the participants complaints about the labels in the installation button. When they decide to download an application they should click on the button labeled "Grátis" (Free) after this, the button changed the label to "Instalar App" (Install App), in participants opinion there was no need to click twice to install a free app.

Comments and reviews in store are relevant information to help users to decide whether or not to download an application. The lacks of classification criteria, filters and evaluation of the comments in the comments area were missing features in participant's opinion.

The search engine is mostly used when the users already know which application to download; the categories are used when the users want to download an application of specific kind, but they don't have one in mind. Both stores have auto complete search, however, each store had a different approach. In the Android store, the suggestions are search related terms. In the Apple store, the suggestions are specific

applications. The participants used both when performing the task but a preference for one or another could not be identified.

Regarding search results, the participants commented about the relevance of the evaluation on the snippets in the list of results but the participants pointed that there is no information about which category the application belongs.

The participant found that the search icon in iPhone App Store is clearer than in Google Play Store. In the Apple store, the search tool is located in a tool bar on the bottom of the screen and it has an icon and a label. In the Android Store, it is located on the top of the screen and there is no label.

The comments related about store organization complained about the structure of the Android store, with too much horizontal navigation. Regarding the application page, participants found it clearer and agreeable in Apple store, however, many indicated that the Android page had more visual appeal. A positive point noted by participants was the fact that the Android store display videos and images on the top of the application page and those are on the middle of the page in Apple store.

5 Conclusion

Trough Usability Testing technique it was possible to identify both qualitative and quantitative data regarding the use of the two major app stores for Smartphone. These information pointed out that there are several opportunities for improvement. In general, the Android system obtained a completion rate greater than iPhone regarding the tasks of search, acquisition and evaluation of applications. However, the iPhone App Store a got a greater satisfaction rate.

Regarding the Android store, the main complaints were relate to those points:

- Control while filling the comments;
- Excessive horizontal navigation;
- Lack of label in the icon search;

Regarding the iPhone store, the main complaints were about:

- Lack of feedback after the completion of the app evaluation.
- Difficulty to type specific characters in
- Double click to download free apps
- Requiring password to download free apps.

Through the comments it was possible to identify that the users of mobile applications stores create their mental models through association with similar experiences in performing similar tasks in other devices or digital media. This means that users expect to find in the store the same functionality they find in the everyday tools of internet such a ways to point the relevance of a comment or advanced search features.

With the expansion and adoption of the apps stores as the default solution to distribution of most of the digital content available. Users will have different expectations and needs as new content becomes available, many of them already present in the routine of these users through other distribution solution. It is necessary to study and

understand the characteristics of each of this new platform and the adaptation and improvement necessary to the distribution of content the available in the store, since not only applications are been distributed by digital stores.

Besides these possibilities and new applications of content stores is important to point that many improvements must still be made in app stores for Smartphone. This research took place from September to October of 2012. During that period the iPhone App Store was updated and the improvement in the interface could have solved many problems that were identified here. However, those changes could make emerge new problems, so it is important keep track of this changes and evaluate if they solved the problems and improved the usability of the stores. The Google Play Store started to offer books and it impacted in the layout, organization and usage of the search tool.

Some difficult identified by the research were related to the hardware of each device and are inherent to the touch screen technology. To overpass these constrains, it is necessary to work on the long term evolution of the technology, those information could be used as reference and parameters while new technologies emerge and are adopted by the manufactory of Smartphone and others mobile devices.

References

1. Gartner, Inc., <http://www.gartner.com/newsroom/id/2237315>
2. mobiThinking, <http://mobithinking.com/mobile-marketing-tools/latest-mobile-stats/a>
3. IDC, <http://www.idc.com/getdoc.jsp?containerId=prUS23624612>
4. Cavallini, R., Reis, T.: *1a Pesquisa #Mobilize Consumidor Móvel (2011)*, http://www.xydo.com/toolbar/26516473-consumidor_m%C3%B3vel_2011
5. Rocha, R.: *Relatório Mobilize de Inteligência de Mercado Mobile #4.*, <http://www.mobilizado.com.br/wp-content/uploads/2012/04/Relat%C3%B3rio-Mobilize-IM-4-Mercado-Mobile-Abr-2012.pdf>
6. Prata, W., de Moraes, A., Quaresma, M.: User's demography and expectation regarding search, purchase and evaluation in mobile application store. *Work: A Journal of Prevention, Assessment and Rehabilitation* 41, 1124–1131 (2012)
7. Rubin, J., Chisnell, D.: *Handbook of Usability Testing, Second Edition: How to Plan, Design, and Conduct Effective Tests*. Wiley Publishing, Inc. Indianapolis (2008)