Making Mobile Touch Screens More Usable Using SUM Metrics Model

Allan Telles Bessa, Simone Bacellar Leal Ferreira, Adriana Cesário de Faria Alvim, Ney Wagner Freitas Cavalcante, Aline da Silva Alves

Graduate Program in Informatics (PPGI) Federal University of the State of Rio de Janeiro, Brazil (UNIRIO)

{allan.bessa, simone, adriana, ney.cavalcante, aline.alves}@uniriotec.br

Resumo

A interação com os celulares vem se modificando ao longo do tempo por meio de toques na tela e telefones inteligentes. Diante desse cenário, os projetistas encontram barreiras para construírem interfaces que atendam tanto aos usuários com limitações quanto aos usuários comuns. Com o objetivo de minimizar esses obstáculos, a presente pesquisa adota uma abordagem de avaliação de usabilidade e acessibilidade com foco na terceira idade a fim de identificar os prós e contra nos sistemas Android, iOS e Windows Phone, bem como propor recomendações nesses sistemas para tornar a interação desses públicos mais fácil e agradável.

Essa pesquisa envolveu 170 respostas a questionários com perguntas a respeito do perfil cadastral e de uso. Vinte usuários participaram de experimentos nos três sistemas supracitados, originando assim em uma lista recomendações, como também evidenciou: cada sistema possui vantagens que podem ser compartilhadas entre si visando à evolução e facilidade ao uso.

palavras-chave: Usabilidade, Acessibilidade, Idosos, Ecrãs Táteis, Métricas.

Abstract

Cell phone interaction has changed with the advent of touch screen and smartphones. Given this scenario, designers find obstacles when trying to build interfaces able to cater to special users as well as the ordinary one. This present research aims for minimizing such obstacles by adopting usability and accessibility evaluations focused on the elderly in order to identify pros and cons of systems like Android, iOS and Windows Phone, as well as propose recommendations on these systems so as to make interaction easier and more pleasant for this group.

This research involved 170 answers to a questionnaire about registration and use profiles. Twenty users took part in the experiment by testing the three systems mentioned above, and then produced a list of recommendations that made several this clear: each system has advantages that could be shared among each other's for the purpose of development and facilitation.

keywords: Usability, Accessibility, Elderly People, Touch Screen, Metric.

1. Introduction

Interaction on cell phones has changed over time due to the appearance of touch screens, and was boosted in 2007 with the arrival of the iPhone [Apple 2012; Gartner 2012], which also strengthened platforms such as Android, Windows Phone among others [Android 2012; GSM-ARENA 2012; Nokia 2012]. This type of interface, also present in ATMs, product search terminals, etc., can make this experience a little easier and more pleasant for users [Häikö et al. 2007].

As this data entry mechanism is relatively new, considering that the first cell phone was commercialized in 1983 by Motorola [Abreu and Moraes 2005], few studies have been found that evaluate these devices on the grounds of their usability and accessibility by young and old audiences alike.

Factors such as the growth of the elderly population, the representation of this older public with regard to their ability as mobile users, as well as the low use of services such as sending text messages (SMS) [CETIC 2013] become relevant in understanding the possible difficulties that this older public might have when using touch screen cell phones.

However, accessibility difficulties, common to specific users such as the elderly [Ferreira and Nunes 2008], are capable of being replicated by groups of youngsters, thus generating a potential usability problem too. Therefore, a given problem may not necessarily be restricted to a specific group.

Thus, the aim of this work, based on a model that combines usability metrics - the Single Usability Metric (SUM) - is to evaluate the ease of touch screen cell phone use with Android, iOS and Windows Phone systems for young audiences and seniors. Emphasis is given in the latter case to those that do not have severe visual, sound or motor limitations so as to be able to identify the pros and cons of each one and suggest possible improvements to the interface in these systems and make recommendations for the design of such interfaces with a focus on the elderly.

2. Theoritical Framework

2.1. Usability and Accessibility

The term usability represents ease of use and determines whether a product or system is difficult to forget, attains satisfactory levels when being handled by users, has no operational errors, solves efficiently the tasks proposed by the user, among others [Ferreira and Nunes 2008].

This ease of use, the primary concept of usability, consists in determining the time and effort required for users to understand and learn to use the systems. During user interaction, the "easiness" factor also reflects the number of mistakes made by individuals. However, this does not mean that a system that is easy to use automatically falls within the framework of easy to learn, since they can present inefficiencies, either due to efficiency-when-using factor (what the user can do) or by the safety factor (how the user performs a task) [Nielsen 1993]. These factors are included in usability attributes efficiency, effectiveness and user satisfaction in part 11 – guidance on usability of ISO 9241 [ISO9241 1998], and in studies conducted over-90s by other researchers according to correlation adapted below (Table 1).

Table 1: Correlation of usability attributes by some researchers (Adapted from Seffah *et al.*, 2006)

Constantine e Lockwood (1999)	**ISO9241 (1998)**	**Preece *et al.* (1994)**	**Nielsen (1993)**	**Schneiderman (1992)**	**Shackel (1991)**
Efficiency in use	Efficiency	Throughput	Efficiency of use	Speed of performance	Effectiveness (speed)
Learnability	-	Learnability (ease of learning)	Learnability (ease of learning)	Time to learn	Learnability (time to learn)
Rememberability	-	-	Memorability	Retention over time	Learnability (retention)
Reliability in use	-	Throughput	Errors	Rate of errors by users	Effectiveness (errors)
User satisfaction	Satisfaction (comfort and acceptability of use)	Attitude	Satisfaction	Subjective Satisfaction	Attitude

Due to these usability attributes, Ferreira and Nunes [2008] developed a taxonomy of usability, in which were established non-functional requirements (NFR) to conduct a usability analysis based on two basic interaction activities: one related to the display of information and the other to data entry.

To have users focusing their attention mainly on their tasks, the process of software development must be “user centered”, that is, its interface must be designed with the objective of satisfying the expectations and needs of users. The design of an interface that considers users’ characteristics and the NFR usability is a difficult process for many reasons, but most of this difficulty can be traced to the lack of attention on NFRs during the system definition process [Ferreira and Nunes 2008].

The usability NFRs related to the information display are inherent to the forms of presenting information, such as texts, images, the proper use of colors etc., broken down into thirteen requisites: consistency, feedback, levels of skill and human behavior, human perception, metaphors, minimization of memory load, efficiency in dialogues, functional classification of commands, direct manipulation, display of only relevant information, use of clear labels, proper use of windows, project independent of resolution [Ferreira and Nunes 2008].

With respect to data entry, the NFR adopt three procedures as good practice for building an interface: aid mechanisms, prevention and treatment of errors [Ferreira and Nunes 2008].

Accessibility refers to the possibility of anyone being able to use or access the product, i.e., it consists of introducing users to life's pleasures without exclusion, regardless of whether they are limited or not, within the social context: social, computing, and communications, among others [Ferreira and Nunes 2008].

Among specific audiences, who in general have accessibility problems are those with limitations related to activities of daily life, such as visual, hearing, motor and cognitive impairments. Limitations are also present in older adults, in users who do not have broadband and in the functionally illiterate, individuals who lack basic reading skills [Ferreira and Nunes 2008; Thatcher et al. 2006].

The Brazilian population has increased exponentially in the last 60 years, having gone from 50 million to approximately 190 million people, i.e. a growth of over 350%.

Longevity has also increased; average life expectancy jumped from about 40 to 70 years between 1960 and 2010 [IBGE 2010; Moreira 1997].

In Brazil, the aged or elderly population falls within the age bracket of 60 years and older [Brasil 2003].

This age group expands every year, and in 2010 already represented 11% of 190 million people, totaling about 21 million people [IBGE 2010].

2.2. Usability Metrics

This research was based on usability attributes – efficiency, effectiveness and satisfaction - used as the basis for creating metrics to measure facility in using an interface [Seffah et al. 2008]. These metrics are divided into performance and perception. The first portrays the attributes of efficiency and effectiveness, whilst the metric of perception reflects qualitative data based on behavior or user satisfaction [Seffah et al. 2008; Thatcher et al. 2006].

Examples of quantitative metrics or of performance include the percentage of completion of a given task, number of errors, execution time, number of paths for reaching a feature, number of mouse clicks, and frequency of use of the help mechanism, among others. Collecting them is achieved by means of log analysis, automatic screen and video capture, as well as among other techniques [Seffah et al. 2008; Thatcher et al. 2006].

Regarding qualitative or metric perception, the calculation is subjective, i.e. dependent on user behavior at the moment. These metrics are exemplified by collecting preferences and the level of user satisfaction with a particular task or product and can be collected by means of questionnaires and interviews [Diniz et al. 2006; Seffah et al. 2008; Thatcher et al. 2006].

2.3. Related Work

The work presented by Larsson and Radi [2009] addressed the touch screen in the features test of cell phones that use text entries, conducted with ten elderly people. This study highlighted the reduction in the size of phones and the increase in resources linked to them, which consequently hampers their use by the elderly. Nevertheless, there are sophisticated handsets with larger screen size, as is the case of the 2012 Samsung Galaxy Note released with a 5.3-inch screen [Samsung 2012]. The authors emphasize that the test participants reported a lack of feedback in the act of pressing the button when typing, even if this act on the touch screen is displayed directly on the screen. However, they felt comfortable due to the fact that they do not exert "pressure" on the key.

In another work [Gonçalves et al. 2011], the authors conducted a case study with ten elderly people in order to observe this public interaction with the touch screen phone Samsung Galaxy 5 and identify usability problems and accessibility. Problems such as lack of standardization of buttons, screen size and letters placed too close together were identified, among others. After this practice of observation and non-functional requirements, a framework was adopted to support the interface design process. However, the study was based on a single touch screen model and on a specific operating system, Android 2.2, as well as on a number of ten participants, which may possibly turn out to be limiting factors in this work.

The article written by Stößel and Blessing [2010] evaluated gesture interaction on touch screen phones and whether elderly people had behavior similar to younger generations. This study involved two groups: 22 younger and twenty elderly people, who undertook 34 tasks on touch screen phones. In certain tasks, such as to execute a zoom on an image or a map, both groups presented different behavior. The younger group used two fingers and the other one rather to use only one finger with two taps. The authors recommend standardizing the “touches” in such a way as to make them more accessible to any user.

In [Mol 2011], usability recommendations for mobile apps were established with a focus on seniors by means of a case study that adopted a set of questions and answers developed by the author and run on touch screen cell phone based on metrics established in the literature. This study also gave examples of good practices in the use of interface elements and proposed new tests for existing features in mobile phones, such as calendars, messages and others. However, not only did he use a single model touch screen device, but he also did not conduct a comparative study in order to validate if the same problems happen with younger groups.

With respect to metrics-related studies in Hussain and Ferneley [2008], the authors developed a specific model for establishing usability metrics focused on mobile apps by means of a paradigm to measure software projects and processes, the Goal Question Metric (GQM) [Basili 1992]. However, the authors failed to mention usability metrics and only emphasized the GQM.

Sauro and Kindlund [2005] proposed a model to standardize usability metrics based on the aspects of effectiveness, efficiency and satisfaction proposed by ISO 9241 and ANSI. This method, called Single Usability Metric (SUM) by the authors, uses performance metrics, such as the successful execution of tasks, the number of attempts (maximum two) and the time it takes for doing tasks in seconds, as well as the metric of perception such as user satisfaction,…, with values from 1 to 5 to be normalized following the Six Sigma process improvement model scale.

Martin and Weiss [2006], using SUM, depict a comparative study regarding the download process of value-added services, such as wallpapers, ringtones and games. Thirteen different types of cell phones were used to identify on these devices and services the respective scores from the standpoint of usability. It is worth mentioning that in the 200 interviews conducted these authors did not report in detail the profile of users involved, such as level of education, age, and experience in use, among other itemizations.

In the specific context of qualitative metrics, the Software Usability Measurement Inventory (SUMI) [Kirakowski 1996] proposes to measure user satisfaction by means of a questionnaire with fifty declarations. The user can answer by choosing one of three options at his/her disposal: "agree", "do not know" and "disagree". The research shows that more than three responses hinder the level of relevance and of rank of the participants.

3. Research Method

The steps involved in this research study are: (A) determining user profiles in the case study; (B) study and selection of touch screen operating systems; (C) definition of usability metrics technique; (D) preparation of case study (E) Interpretation of results. Steps D and E are detailed in sections 4 and 5, respectively.

A. Determining user profiles in the case study

In this research work, we selected young and elderly users to check if the same usability issues occur in both groups. After this stage, groups were created and separated by educational level, by the most used service on their mobiles such as SMS [CETIC 2013], and by the use of conventional-type cell phones. Those users in both the young and elderly groups who were already using touch screen cell phones were excluded from the study. This "cut" was made because the study was slated to evaluate the first experience in using this type of cell phone, so prior experience was not desirable, since people familiarized with this kind of technology could influence the results, thus compromising a valid measure.

Thus, young and elderly participants with at least some experience of higher education were separated into four groups of five persons each (two from each age group). Groups of five became appropriate because a survey conducted by Jakob Nielsen [Nielsen 2012] established that 85% of usability problems are identified from this number of users.

B. Study and selection of operational touch screen systems

The selection of operating systems on touch screen type mobiles occurred due to their ranking as smart phones: the Android Google system and the Apple iOS system occupy the first two positions in the ranking, and the Microsoft Windows Phone the third position [Gartner 2012; TELECO 2012]. We chose three systems in view of the concern about the runtime in the case study versus the availability of participants for these evaluations. In addition, these systems occupy 92% of the smart phone market [TELECO 2012]. After this definition, mobile models of each system were established with similar physical characteristics, such as screen size and processor. Thus, the cell phones used were: iPhone 4 with iOS, Sony Xperia U with Android and Nokia Lumia 800 with Windows Phone.

C. Definition of the technique for measuring usability

Sauro and Kindlund [2005] developed a method called SUM to standardize four usability metrics: 1) successful task completion; 2) errors counts; 3) task time; 4) task rating [Thatcher et al. 2006].

The SUM method was used in this research as the measurement technique due to the fact that it considers both perception and performance metrics in a simplified model.

3.1 Research Limitations

Due to the difficulty in finding volunteers, three of the twenty participants had direct involvement with the technology, one belonging to the young adult group with incomplete higher education, one belonging to the young adult group with college degrees (YC) and another to the elderly group with college degrees (EC). Probably this influenced some results. Conducting the tests with users took place in open environments, which does not ensure the same results as in a laboratory environment, nor does the fact that testing occurred in just one Brazilian city, a limiting factor in view of the reality of Brazil's population with cultural differences. The choice of some of the functionality may have caused a limitation, as it does not assess the interaction capabilities of the mobile phone as a whole. The tasks used were performed on three types of apparatus, which made the tests lengthy and tiresome for the users, in this way jeopardizing their level of satisfaction.

4. Establishment of The Case Study

A case study of the techniques of direct observation and questionnaire was established to identify, by means of a comparative study and of a model that combines usability metrics, possible problems of usability and accessibility, as well as to identify pros and cons of touch screen mobile operating systems. Therefore, this study involved two steps: collection of users through a questionnaire and direct observation.

In the first stage, young and elderly users were identified in order to take part in direct observation, this step also enabling obtainment of certain data such as level of education and familiarity with sending text messages (SMS), the most widely used functionality on cell phones, second only to the voice [CETIC 2013].

To select users, a questionnaire was drawn up, with sixteen questions divided into two sections: personal data and mobile use profile. For the data to be based not only on the elderly on the internet, in principle adepts of technology, the survey was conducted, in such places as public squares, churches, educational institutions and companies, and administered directly to people known to the researcher. Heterogeneous environments were sought out in an attempt to collect different usage profiles and different levels of schooling.

This same questionnaire was duplicated in the Google Docs tool [GOOGLE 2006], enabling it to be distributed by e-mail and by a social network so as to become available on the internet and thus increase the capillarity of collection.

In the second phase, that of direct observation, five procedures were used: 1 – adoption of a mechanism to capture user interactions by means of the Mobile Observation Device 1000 [MOD1000 2012]; 2 – creation of new tasks to be performed on all three cell phones, wherein each one has been explained by the researcher to avoid uncertainty about the commands to be executed: T1 (Task 1) - turn on and unlock the cell phone, T2 (Task 2) – locate the message icon, T3 (Task 3) – insert a new message, T4 (Task 4) – type the phone number, T5 (Task 5) – type the text message, T6 (Task 6) – send the message, T7 (Task 7) – go back to initial screen, T8 (Task 8) – lock the cell phone and T9 (Task 9) – read a message; 3 – selection of twenty users participating in four groups: young people with complete or incomplete university studies and the remaining two groups for elderly users; 4 – performance of above mentioned tasks by all twenty users on the three cell phones so as to collect performance metrics; 5 – completion of satisfaction questionnaire and interview to obtain perception metric.

5. Interpretation of Results

5.1 Analysis of Questionnaire

The questionnaire was implemented face-to-face with elderly people directly at the *Academia da Terceira Idade* – ATI in a suburban neighborhood of a large capital city. ATI is a quality of life program for the elderly run by a municipal agency: Special Secretariat for Healthy Aging and Quality of Life [SESQV 2009]. With this collection means, 25 elderly participants were introduced to the survey, in addition to two others under the age of 60, bringing the total to 27 users contributing in this public square.

To complete the sample, an online questionnaire was used highlighting the search of elderly users and 6,233 users received e-mail messages asking whether they wanted to participate in a research study. However, only 45 responded, accounting for a low effectiveness rate of 0.72%. It is noteworthy that these 45 included 36 seniors and nine youngsters.

Against this backdrop of poor adhesion, another means of disseminating the research was activated, the social network Facebook. Messages were sent to 150 friends and thirteen elderly users identified through this online channel, *i.e.* 8.6% of adhesion.

Environments such as companies, educational institutions, churches and distribution to friends of the researcher were used to obtain other users, but this time geared to Young people. This heterogeneity is due to the fact that attempts were being made to collect different levels of schooling and user profiles.

Thus, participants' responses were collected using a pen, getting: Forty young users through two schools, one middle school and one private university; 35 users in three companies, 33 young and two elderly people; and ten close friends, two young and eight elderly people. Therefore, the first step of the case study by questionnaire contained 170 users, with 84 seniors and 86 young people.

The usage profile of respondents varied according to the type of cell phone: conventional or touch screen. There was a larger number of the latter in the younger group, *i.e.*, of the 86 young people involved, 61 (70.9%) had touch screen mobiles and 25 (29.1%) had conventional mobiles. As for the elderly, there is a reversal of this analysis: of the 84, 22 (26.2%) have touch screen mobiles and 62 (73.8%) conventional ones (Figure 1).

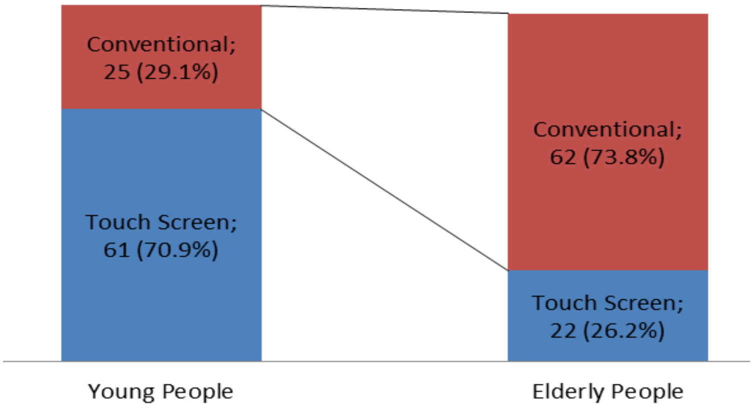

Figure 1: Users according to type of mobile; inversion of use.

Regarding the SMS analysis, a larger number of users make use of this service with front touch screens as compared to conventional ones, as shown by the percentages of 86.7% and 43.7%, respectively, in the "Total", regardless of the type of user or of schooling (Table 2).

Table 2: Cell phone usage profile of young and elderly people versus type of cell phone.

Educational Level	Touch Screen		Conventional	
Group	Send SMS	Do not send SMS	Send SMS	Do not send SMS
Younger Complete	20 (95.2%)	1 (4.8%)	6 (75%)	2 (25%)
Younger Incomplete	38 (95%)	2 (5%)	12 (70.6%)	5 (29.4%)
Elderly Complete	10 (55.6%)	8 (44.4%)	10 (31.2%)	22 (68.8%)
Elderly Incomplete	4 (100%)	0 (0%)	10 (33.3%)	20 (66.7%)
Total	**72 (86.7%)**	**11 (13.3%)**	**38 (43.7%)**	**49 (56.3%)**

Nevertheless, an analysis of the evaluation of usability and accessibility through the second stage of the case study becomes necessary in order to demonstrate the possible difficulties of users as well as the pros and cons of touch screen systems depicted in the following section.

5.2 Analysis of Direct Observation

The “Task” column (Tables 3, 4, 5 and 6) refers to tasks mentioned above in the "Establishment of The Case Study" section.

Of the 170 users obtained through the questionnaire, twenty were selected to perform the second phase of the case study, that of direct observation.

With the help of the equipment MOD 1000 and of Google Picasa software, more than 15 hours of recordings were generated, totaling 4.33 Gb of physical storage.

Some failures and similar behavior during task execution between user groups were identified for each operating system. Some of these tasks failures weren't shown since they were not relevant evidence on this case study. It was noted that in the Android system the main failures of similar behavior refers to the metaphors used for the *enter* and *backspace* buttons that confused users (Figure 2) when typing a message, regardless of the user group: 25% of them were confused by these buttons (Table 3).

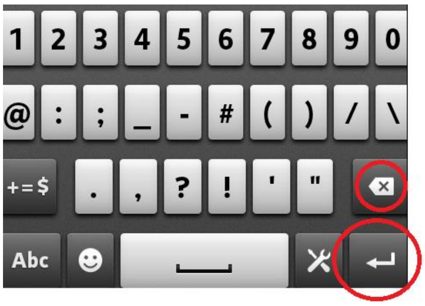

Figure 2: Confusing *enter* and *backspace* metaphors on the Android.

Table 3: Similar behaviors by user group on the *Android.*

Task	Behavior	Number of Users on Educational Levels				%
		Younger Incomplete	Younger Complete	Elderly Incomplete	Elderly Complete	
T1	Powered camera button	1	0	2	0	15%
T4	Used text message field	0	0	3	0	15%
T5	Erased with *enter* instead of *backspace*	1	1	1	2	25%
T7	Powered options on screen	0	0	2	0	10%
T8	Powered options on screen	0	0	1	3	20%
T9	Slid finger rapidly	0	1	2	0	15%

In this Android system, the *enter* button is also represented with a left arrow and *backspace* with an arrow, but with an "x" in it (Figure 2). One possible improvement would be to alter the "x" metaphor of the *backspace* by using an "eraser" and another button with the written word *enter*.

In the iOS system the task of turning the system on and off became complex for sixteen users (80%) that showed similar behavior (Table 4). Of these participants, six powered the bell and another ten powered the middle button, which has as one of its functions "return to home screen", thus underlining the difficulty to perform this action. Therefore, adjustments are needed to reduce the mistakes, despite the possible involvement of the design. However, exchanging the bell button which has the function of powering the system on and off can minimize this situation.

Table 4: Similar behaviors according to user groups on the iOS.

Task	Behavior	Number of Users on Educational Levels				%
		Younger Incomplete	Younger Complete	Elderly Incomplete	Elderly Complete	
T1	Powered the bell button	1	1	2	2	30%
T1	Powered the middle button / back to initial screen	0	3	3	4	50%
T5	Did not locate the character (*)	1	1	0	1	15%
T5	Typed message on the display screen	0	0	0	2	10%
T7	Powered the return button with the enter function	4	2	2	3	55%
T8	Powered the options on the screen	0	0	0	2	10%
T8	Powered the bell button	1	1	0	0	10%

The task of returning to the home screen showed the problem of the word "return", which has the function of scrolling down to the line below or *enter*, *i.e.*, 55% of users made this mistake (Table 4). As a recommendation for this and other systems, changing the word "return" or "*retorno*" in Portuguese to "enter" may reduce this percentage of errors (Figure 3).

Figure 3: The word "return" or "retorno" confuses users into thinking they will return to an earlier screen or to an earlier operation on the iOS.

Users on Windows Phone showed a strong tendency - 60% (Table 5) to connect the phone via the button below the device (Figure 4), which in this case activates the camera.

Table 5: Similar behaviors by user groups on Windows Phone.

Task	Behavior	Number of Users on Educational Levels				%
		Younger Incomplete	Younger Complete	Elderly Incomplete	Elderly Complete	
T1	Powered the button camera.	1	4	2	5	60%
T2	Associated the envelope metaphor to SMS	0	1	1	1	15%
T5	Did not locate the character (*)	2	4	3	1	50%
T5	Erased with *enter* instead of *backspace*	1	1	1	2	25%

Figure 4: Camera button on the Windows Phone is confused with the power button.

All behaviors mentioned in the three systems have been influenced in one of the performance metric, the "number of errors", according "2.2 Usability Metrics".

Both the performance metrics collected by the videos generated during the execution of tasks on the cell phone and the associated perception metric obtained by the questionnaire were entered into a spreadsheet support, available on the website of the authors of the SUM method [SUM 2005]. The results are summarized in Figures 5 and 6. Scores have values from 0 to 100. The higher the score, the easier it was to perform the task or use the system involved according to these authors [Sauro and Kindlund 2005].

With the consolidation of the SUM scores caused by these metrics such as successful completion of task, number of errors and run time, as well as the metric of satisfaction, according to "3. Research Method", it was found that, besides the aforementioned gaps between groups of users in the three systems, improvements are needed in order to increase these scores and are hence classified in terms of attributes of usability, efficiency, effectiveness, and satisfaction itself.

The Android system was easier for the group of young users with incomplete higher education who attained the score of 77.7, while the Windows Phone earned the highest score in the group of elderly users with incomplete higher education, obtaining 52.4 (Figure 5). iOS obtained the highest scores of 82.2 and 59 for the groups with college degrees, formed by young and old respectively. Thus, it is that the iOS appears twice among the four groups (Figure 5).

Regarding the groups analysis, elderly people obtained lower scores than the younger groups, execution time of the tasks being one of the factors that influenced on results. In education levels side, it was not evidenced differences to consider, in fact, the problems identified affected both young and elderly people (Figure 5).

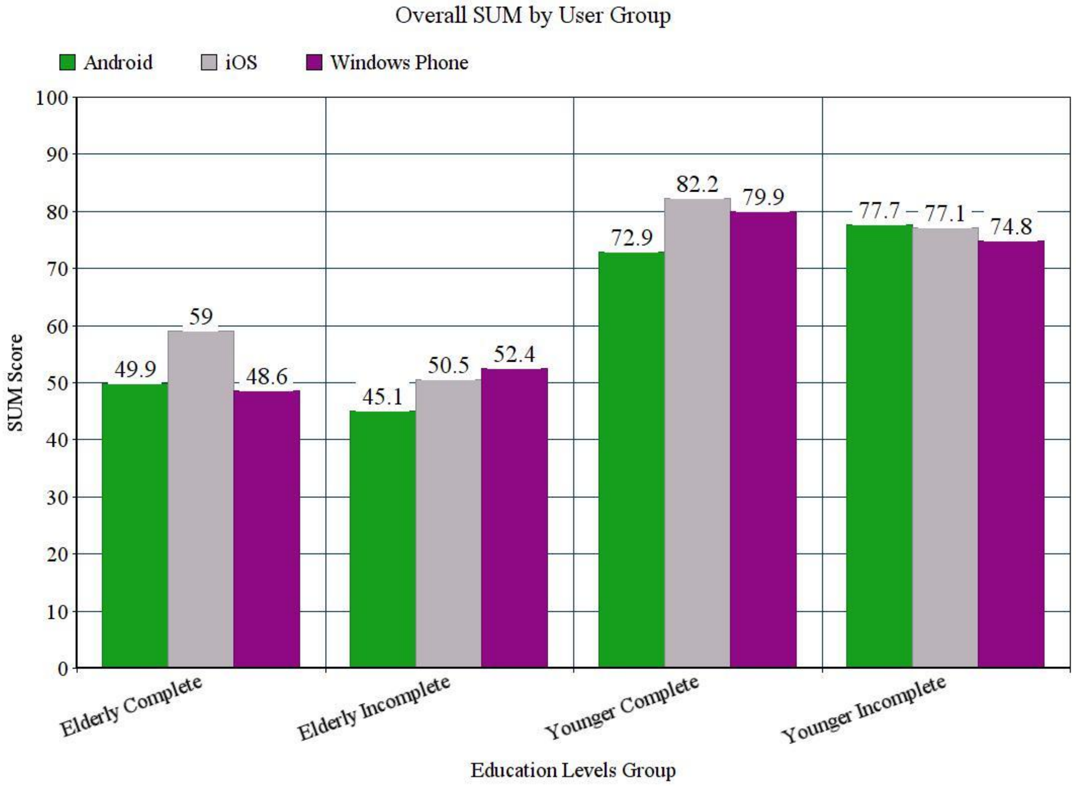

Figure 5: Overall SUM average by group of users and the highlights of each one.

Nonetheless, this does not go to prove that iOS is the best system: three tasks with higher scores are present in each system: Android with regard to tasks 1, 3 and 6; iOS for tasks 5, 8 and 9; Windows Phone for tasks 2, 4 and 7 (Table 6). Therefore, each system has better usability as compared to another depending on the type of task.

Table 6: Overall SUM average of nine tasks in the systems, by tasks.

Task	SUM Score in each System		
	Android	iOS	Windows Phone
T1 – Turn on / Lock it	**55.4**	49.4	47.6
T2 – Locate the message icon	91.7	94.7	**95.1**
T3 – Insert a new message	**85.2**	82.3	67.3
T4 – Type the phone number	31.7	63.3	**81.0**
T5 – Type the text message	38.2	**44.3**	22.3
T6 – Send the message	**96.5**	92.8	65.3
T7 – Go back to initial screen	40.1	36.9	**71.2**
T8 – Lock it	46.3	**59.7**	57.2
T9 – Read a message	67.6	**81.1**	68.2

6. Recomendations

SUM made it possible to assess which tasks were easier to perform in the systems, as well as those that need improvements and developments, in addition to obtaining observations made by the research participants and their comments. Given these results and with the support of the taxonomy of usability [Ferreira and Nunes 2008] according to the Usability and Accessibility section, Table 7 presents some recommendations for possible improvements to the interface in these systems based on the results of this research. Once these recommendations were developed, it is interesting to make more tests for measuring impact.

Table 7: Recommendations for the three systems tested among users.

Recommendations	Applies to:		
	Android	**iOS**	**Windows Phone**
More spacing between letters on the screen, as in the iOS.	X		X
Use the same numerical keyboard as in the Windows Phone.	X	X	
Make shortcuts available in the “factory” version, as in the Android model, to activate, for example, GPS and Wi-Fi.		X	X
Allow a rapid finger movement to “slide to unlock”, as in the iOS and Windows Phone.	X		
Insert a text field to type numbers in a new message, rather than navigating on five screens. Adopt the same system on the iOS and Windows Phone.	X		
Insert a sentence in gray on the text message Field, as in the Android and Windows Phone: “Type your message here...”.		X	
Change the position of the camera button from the lower right-hand corner to the upper left hand corner.	X		X
Change the position of the on/off power button with the bell button.		X	
Change the position of the on/off power button with the increase/decrease volume button.			X
Change the increase/decrease volume button from the upper left-hand to the upper right-hand corner. corner.		X	
Complement the function “return to initial screen” with "house" icon to "house with arrow to the left".	X		
Change icon for return to initial screen function from a “square” to a “house with arrow to the left”.		X	
Change icon for return to initial screen function from “initiate windows” to “house with arrow to the left”.			X
Insert the same Android icon for the on/off power button.		X	X
Change the icon for new message function to an “envelope”.	X	X	X
Change the icon for send message function to an “envelope”.			X
Change the icon for messages function from a “face” to an “envelope”.			X
Change the icon for the backspace button from one to an “eraser.”	X	X	X

Insert icons for raising and lowering volume on respective buttons.			X
Insert an icon to represent the camera button.			X
Show texts with the names of buttons without the user having to click on button "…" for it to appear.			X
Add the word *enter* on the respective button.	X	X	X
Alter the current label that represents special characters to "symbol".	X	X	X
Add the word "unlock" as in the iOS example.	X		
Add the sentence "slide upwards to unlock".			X
Add to the sentence "slide downwards to power off" another sentence: "or simply touch the same button to lock the screen".			X
Show a *pop up* with a tip: "press power on/off button once to lock your cell phone...." followed by the option "do not show this message again" when the user turns on the device.	X	X	X

7. Final Considerations

This study applied an approach to evaluating usability in touch screen mobiles in order to identify the pros and cons of Android, iOS and Windows Phone systems, and came up with a list of recommendations based on usability requirements made apparent by the difficulties of young and elderly users.

A case study was implemented in two stages. The first by means of a questionnaire, where forms were distributed electronically and filled out manually, and containing sixteen questions divided into two sections: personal data and use of mobile profile. Thus, 170 responses were obtained which showed certain results, such as: 1 - increased utilization of touch screen mobiles by young people, with 70.9%, as compared to 26.2% by elderly users; 2 - increased use of SMS in touch screen mobiles with 86.7% as compared to 43.7% in conventional cell phones; 3 - lack of influence of level of schooling for using SMS in either type of cell phone. 4 - low percentage of use of SMS by elderly users in conventional type cell phones, a little over 30%, regardless of educational level.

Despite these favorable results for the touch screen, the second stage of the case study, through direct observation, showed that this type of interaction becomes subject to modifications by the difficulties presented by both young and elderly users with regard to some of the functions.

Running tests in this second stage involved the application of nine tasks related to sending, receiving and reading messages and operations with the phone itself, such as, turning it on, locking it and unlocking it. During these tests, the MOD 1000 equipment was used for collecting user interactions on mobile phones, Google Picasa software for recording audio and video, and supporting forms to record errors counts, task completion, task time and task rating as informed by users.

From these artifacts and from the consolidated data, a model that combines usability metrics, SUM, was adopted to generate scores to identify which tasks presented difficulties.

It was found that the level of schooling did not influence the difficulties presented and, in fact, the problems identified affected both young and senior users. Even so, the elderly, regardless of education, had lower scores than the young users, thus revealing the existence of the age factor.

Given the results and the scenarios presented, the approach of evaluating usability and accessibility by means of SUM and of applied case studies enabled us to identify situations in which systems were easier and those which were more difficult to use, thus resulting in improvements and suggestions. Thus, some recommendations were pointed out that are applicable to non-functional usability requirements in the tested systems.

The research presented important results in respect to the lack of influence of the educational level of the user when using touch screen mobiles and also the greater incentive for this type of interaction with other value-added services like SMS. In addition, an evaluation was presented on how to approach not only the young but also the elderly user, *i.e.*, by portraying issues of usability and accessibility.

Nevertheless, this study was limited to testing certain functions in the three systems and in specific versions, making it necessary to run tests on the current versions and other features most used in mobile phones aside from SMS, such as sending e-mails, instant messaging and social network apps, among others. Other future research work will possibly portray touch screen devices such as tablets, in which screen size may have an influence in obtaining higher scores in the SUM or other method, depending on the type of task to be performed.

References

Abreu, Leonardo Marques de; Moraes, Anamaria de (2005), *Usabilidade de telefones celulares com base em critérios ergonômicos.* Master's thesis, Department of Arts and Design, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil.

Android (2012), Online document available at http://www.android.com/devices/, last visited on August 16, 2012.

Apple (2012), Online document available at http://www.apple.com/br/iphone/, last visited on August 12, 2012.

Basili, V. R. (1992), *Software Modeling and Measurement: The Goal/Question/Metric Paradigm*. Technical Report CS-TR-2956, Department of Computer Science, University of Maryland, MD 20742.

Brasil (2003), Law 10.741, October/1/ 2003. *Diário Oficial [da] República Federativa do Brasil*, Brasília, DF, 3 out. 2003.

CETIC (2013), *Centro de Estudos sobre as Tecnologias da Informação e da Comunicação*. Online document available at http://www.cetic.br/pesquisas-indicadores.htm, last visited on July 22, 2013.

CONSTANTINE, L. L.; LOCKWOOD, L. A. D (1999). *Software for Use: A Practical Guide to the Models and Methods of Usage-Centred Design*. New York: Addison-Wesley..

Diniz, E. H.; Petrini, M.; Barbosa, A. F.; Christopoulos, T. P.; Santos, H. M (2006), "Abordagens Epistemológicas em Pesquisas Qualitativas: Além do Positivismo nas Pesquisas na Área de Sistemas de Informação". In: *Proceedings of the XXX Encontro da*

ANPAD - Associação Nacional de Pós-Graduação e Pesquisa em Administração, Salvador, BA, Brazil.

Ferreira, S. B.L; Nunes, R. R. (2008), *e-Usabilidade.* Rio de Janeiro, Editora LTC.

Martin, R.; Weiss, S. (2006), "Usability benchmarking case study: Media downloads via mobile phones in the U.S.". In: *Proceedings of the 8th Conference on Human–Computer Interaction with Mobile Devices and Services*, pp. 195–198.

Gartner (2012), Gartner Says Android to Command Nearly Half of Worldwide Smartphone Operating System Market by Year-End 2012. Online document available at http://www.gartner.com/it/page.jsp?id=1622614, last visited on December 13, 2012.

Gonçalves, V. P.; Neris, V. P. A.; Ueyama, J. (2011), "Interação de Idosos com Celulares: Flexibilidade para Atender a Diversidade". In: *Proceedings of the X Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais e V Congresso Latino-americano de Interação Humano-Computador* (IHC+CLIHC 2011), v. 1, pp. 343-352, Porto de Galinhas, PE, Brazil.

GOOGLE (2006), *Google Announces Google Docs & Spreadsheets*. 2006. Online document available at http://googlepress.blogspot.com.br/2006/10/google-announces-google-docs_11.html, last visited on March 5, 2013.

GSM-ARENA (2012), *GSM Phone reviews*. Online document available at http://www.gsmarena.com/zte_s302-3222.php, last visited on December 28, 2012.

Häikö, J.; Wallin, A.; Isomursu, M.; Ailisto, H.; Matinmikko, T.; Huomo, T. (2007), "Touch-Based User Interface for Elderly Users". In: *Proceedings of 9th International Conf. on Human Computer Interaction with Mobile Devices and Services*, pp. 289-296, Singapore, 2007.

Hussain, A.; Ferneley, E. (2008), "Usability metric for mobile application: A goal question metric (GQM) approach". In: *Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services*, Linz, Austria.

IBGE (2010), *Censo Demográfico*. Online document available at http://www.censo2010.ibge.gov.br, last visited on October 9, 2011.

ISO 9241 (1998), *ISO 9241-11: Ergonomic Requirements for Office work with Visual Display Terminals. Part 11: Guidance on Usability*. ISO, 1998.

Kirakowski, J. (1996), "SUMI: The Software Usability Measurement Inventory: Background and Usage". In: P. W. Jordan, B. Thomas, B. Weerdmeester, & A. L. McClelland (eds.), *Usability Evaluation in Industry*, (pp.169-178), London, Taylor and Francis.

Larsson, F.; Radi, I. (2009), *The elderly and text input: A study about mobile phones – traditional, senior-friendly and touch screen*. Degree in Computer Science, University of Gothenburg, Gothenburg, Sweden.

MOD1000 (2012), *Mobile Observation Device*. Online document available at http://www.measuringusability.com/products/mod1000, last visited on December 12, 2012.

Mol, Artur Martins. (2011), *Recomendações de usabilidade para interface de aplicativos para smartphones com foco na terceira idade*. Master's thesis, Postgraduate

Program in Informatics, Pontifical Catholic University of Minas Gerais, Belo Horizonte, MG, Brazil.

Moreira, Morvan de M. (1997), *Envelhecimento da população brasileira*. Doctorate thesis, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil.

Nokia (2012), *Nokia Lumia 800*. Online document available at http://www.nokia.com/br-pt/produtos/lumia/, last visited on December 12, 2012.

Nielsen, J. (1993), *Usability Engineering*. London, Academic Press.

Nielsen, J. (2000), *Why You Only Need to Test With 5 Users*. Online document available at http://www.useit.com/alertbox/20000319.html, last visited on November 16, 2012.

Preece, J.; Rogers, Y.; Sharp, H.;Benyon, D.; Holland, S.; Carey, T. (1994). *Human-Computer Interaction*. Addison Wesley.

Sauro, J.; Kindlund, E. (2005), "A method to standardize usability metrics into a single score". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 401–409.

Samsung (2012), Online document available at http://www.samsung.com/br/consumer/cellular-phone/cellular-phone/smartphones/GT-N7000ZBLZTO-spec, last visited on March 29, 2012.

Schneiderman, B. (1992), *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Second edition, Reading, MA: Addison-Wesley.

Seffah, A.; Donyaee, M.; Kline, R. B.; Padda, H. K. (2006), "Usability measurement and metrics: A consolidated model". *Software Quality Journal*, v. 14, n.2, pp. 159-178.

Shackel, B. (1991), "Usability – context, framework, design and evaluation". In *Human Factors for Informatics Usability*. pp 21-38. Ed. by B. Shackel and S. Richardson. Cambridge University Press, 1991.

Stößel, C.; Blessing, L. (2010). "Mobile Device Interaction Gestures for Older Users". In: *Proceedings of the 6th Nordic Conference on Human-Computer Interaction*. Reykjavik, Iceland.

SUM (2005), *Single Usability Model*. Online document available at http://www.measuringusability.com/SUM/index.htm, lasst visited on November 16, 2012.

SESQV (2009), Secretaria Especial de Envelhecimento Saudável e Qualidade de Vida. Online document available at http://www.rio.rj.gov.br/web/sesqv/exibeconteudo?article-id=126402, last visited on March 28, 2012.

TELECO (2012), *Mercado mundial de smartphones*. Online document available at http://www.teleco.com.br/smartphone.asp, last visited on November 20, 2012.

Thatcher, J.; Burks, M. R.; Heilmann, C.; Henry, S. L; Kirkpatrick, A.; Lauke, P. H.; Lawson, B.; Regan, B.; Rutter, R.; Urban, M.; Waddell, C. (2006), *Web accessibility: Web standards and regulatory compliance*. New York, Friends of ED.

Allan Telles Bessa possui Mestrado em Informática pela Universidade Federal do Estado do Rio de Janeiro (UNIRIO), pós graduado "*lato sensu*" pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) em Análise, Projeto e Gerência de Sistemas e graduado em processamento de dados pelo Centro Universitário da Cidade/RJ. No campo profissional, atua há 20 anos na área de TI com ênfase em gestão de projetos, infraestrutura e desenvolvimento de sistemas.

Simone Bacellar Leal Ferreira é Professora Associada dos cursos de Sistemas de Informação (doutorado, mestrado e graduação) da Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Possui Doutorado e Mestrado em Informática, ambos pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Graduação em Oceanografia, pela Universidade do Estado do Rio de Janeiro - UERJ (1983). Coordena um grupo de pesquisas em acessibilidade e usabilidade da UNIRIO. Autora do livro e-Usabilidade (ISBN 978-85-216-1651-1) pela Editora LTC.

Adriana Cesário de Faria Alvim possui Graduação em Comunicação Social, Mestrado em Informática e Doutorado em Informática, todos pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Membro do Programa de Pós-graduação em Informática, da Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Tem experiência em metaheurísticas para problemas combinatórios difíceis, atuando principalmente nos seguintes temas: heurísticas, problema de bin-packing, problema de rotulação cartográfica de pontos, busca tabu, GRASP, POPMUSIC e algoritmos paralelos.

Ney Wagner Freitas Cavalcante cursa Doutorado no Instituto COPPEAD de Administração da Universidade Federal do Rio de Janeiro (UFRJ). Mestre em Informática na Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Graduado em Administração de empresas. Tem experiência em Administração nas áreas comercial, vendas e negócios.

Aline da Silva Alves cursa Doutorado no Programa de Pós Graduação em Informática na Universidade Federal do Estado do Rio de Janeiro (UNIRIO). Mestre em Informática pela mesma universidade. Graduada em Informática - Análise de Sistemas pela Universidade Estácio de Sá. É Tecnologista em Saúde Pública na Fundação Oswaldo Cruz (Fiocruz). Possui experiência na área de Ciência da Computação, com ênfase em Interação Humano-Computador: Acessibildade web, Usabilidade web, Sistemas de Informação, Comunicação e Informação em Saúde e Recursos Educacionais Abertos.

(esta página par está propositadamente em branco)