ICEB 2013 Proceedings

International Conference on Electronic Business (ICEB)

Winter 12-1-2013

Usability Evaluation And Correspondence Analysis Of Smartphone Operating Systems

Chi I Hsu

Chao Chang Chiu

Wen Lin Hsu

USABILITY EVALUATION AND CORRESPONDENCE ANALYSIS OF SMARTPHONE OPERATING SYSTEMS

Chi I Hsu, Kainan University, Taiwan, imchsu@mail.knu.edu.tw
Chao Chang Chiu, Yuan Ze University, Taiwan, 101chiu@gmail.com
Wen Lin Hsu, Yuan Ze University, Taiwan, wenlin1120@gmail.com

ABSTRACT

User-centered smartphone interface design is important so that consumers can easily learn about and begin to use newly purchased smartphones. This study first evaluates the effectiveness, efficiency, and user satisfaction of smartphone interfaces in Taiwan in terms of two representational operating systems: Android and iPhone OS. The usability evaluation includes observational experiments, user questionnaires, and the Wilcoxon sign-rank test. This study then conducts the correspondence analysis to summarize positive/ negative evaluations of usage and specification for smartphone OS. In observational experiments, a total of 48 participants with no previous smartphone experience on using smartphone with Android and iPhone OS are asked to perform five common phone tasks. All experiments are recorded and observed. Results and observations are discussed to attain a closer match between user needs and the performance of smartphone OS.

Keywords: Usability Evaluation, Correspondence Analysis, Smartphone, User Satisfaction

INTRODUCTION

Mobile phones are a ubiquitous part of everyday life in much of the developed world. The International Data Corporation (IDC) [2] says that "the growth of the worldwide converged mobile device market (commonly referred to as smartphones) more than doubled that of the overall mobile phone market" in the first quarter of 2010. According to the International Telecommunication Union (ITU), the number of cell phone subscriptions around the world will reach 5 billion in 2010 [3]. As consumers become more familiar with and comfortable using smartphones, the smartphone market continues to grow, and market competition continues to intensify.

Mobile phones have been and continue to be transformed into multiplex multimedia instruments. The functionality of smartphones has gradually approached that of handheld computers [11]. In addition to the basic calling and short message functions, smartphones also offer Internet access, personal information management, digital cameras, games, and multimedia [5]. As such, while mobile phones are now regarded as common consumer devices [8], smartphone users, especially inexperienced ones, often face difficulties with common smartphone functions pertaining to setting up, configuring and accessing data services [10]. Yamashita et al. [13] explored potential usability gaps confronted by consumers who switched from a familiar to an unfamiliar mobile phone interface. The results show that users with previous mobile phone experience often encounter problems when learning a new interface due to differences or complexities associated with the new system, including non-intuitive features. As such, the design of the smartphone user interface is vital—consumers must be able to easily understand, set up and use their new products.

According to ISO9241-11 [4], usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." A usability evaluation is based on measures of users performing tasks with the products [6]. The current study evaluates the effectiveness, efficiency, and user satisfaction of smartphone interfaces on two representational operation systems: Android and iPhone OS. This study identified six groups of users based on gender, age, and information technology background—in total, 48 subjects who have never used smartphones are asked to solve specific smartphone tasks. After basic instructions were given, participants were asked to perform specific tasks; all experiments were recorded and observed. Once the tasks were completed, participants were asked to fill out a questionnaire regarding their satisfaction with the interface functions. Statistical results are reported and compared. This study then conducts the correspondence analysis to summarize positive/negative evaluations of usage and specification for smartphone OS. The correspondence analysis displays the evaluations in two-dimensional graphical form. Results and observations are discussed to attain a closer match between user needs and the performance of smartphone OS.

RELATED WORK

Many researchers have investigated cell phone menu designs. Lee et al. [9] clearly represented navigation paths on the Navigation Path Diagram, and discovered that supportiveness was negatively affected by difficult to understand text labels, icons, or menus. Surprisingly, Ziefle and Bay [14] mentioned that adding graphics to the display can improve problematic traditional menus in terms of meeting consumer demands. Today, most smartphones use legible icons rather than descriptions, which solves many of these types of problems.

Research on the usability of mobile devices exists; however, various mobile device features make it difficult to examine usability, such as mobile context, connectivity, small screen space, and restrictive data entry methods [12]. Therefore, many researchers focus on simulation prototypes instead of actual mobile phones. Ziefle et al. [15] chose to investigate a software prototype of a palm-computing platform; participants worked with a mouse, which most of them were highly familiar with. The results were considered underestimated—using mobile phones in a real environment requires that users simultaneously manage complex demands including holding the phone, locating specific functions, and inputting data. This type of problem was also associated with the usability study by Huang et al. [1]; participants were asked to perform tasks on a paper prototype in the study, but admitted that they would likely perform poorly on real devices.

Keijzers et al. [7] conducted a usability benchmark study on three types of smartphones. All participants had to be familiar with computers and cell phones to negate differences in personal characteristics that could contribute to differences in usability. Measurement items included: the percentage of tasks solved as effectiveness; task completion time, the number of hierarchical levels in the menu used, and the number of detour steps as efficiency; and the questionnaire as satisfaction. The results showed significant differences in usability for the selected functions. However, all designed tasks were basic based on smartphone functions currently available. Therefore, our experiment focuses on available smartphone models and takes both newly evolving mobile technologies and basic functions into consideration in an attempt to shed light on the usability of smartphones.

In this study, we measure and analyze participants' navigational behaviors on two different phone interfaces: Android and iPhone OS. The results may assist mobile phone designers to create more intuitive, universal, and easy-to-use interfaces that can attract diverse users in terms of gender, age, and computer-related experience.

EXPERIMENT

Participants

To include users with various characteristics and backgrounds, this study identified eight groups of users based on gender, age, and the presence or lack of an information technology background—defined as having studied or worked in information technology related fields. Six people were invited to each group, for a total of 48 participants. In addition, to ensure that the experience would not influence the usability evaluation, only people who had never used a smartphone with Android and iPhone OS were invited to participate. Table 1 shows the characteristics of the grouped participants. To separate the participants into two groups based on age, this study defined that participants over 35 years old were classified as middle-aged, while those under 25 years old were classified as young adults.

Table 1. Characteristics of Test Participants

		IT background	Non IT background
Middle age (>35)	male	Group 1	Group 5
	female	Group 2	Group 6
Young (<25)	male	Group 3	Group 7
	female	Group 4	Group 8

Test Materials

In this study, two popular smartphones were used: iPhone 3G for the Apple iOS (Phone-I) and HTC magic for the Android OS (Phone-II). Fig. 1 shows the Physical appearance for the two smartphones. Both employ an application grid main menu presentation, and both were configured for use with the same mobile network provider (Far EasTone) in Taiwan. In order to ensure fairness, all smartphones were reset to the initial factory settings. Moreover, once subjects completed the tests, all data was erased and units were reset to an identical start screen for the next participant.

Phone-I Phone-II

Figure 1. Physical Appearance for the Two Smartphones

Experimental Tasks

As smartphones have a very diverse range of functions, this study selected several typical tasks for the experiment:

Task 1. Call a number stored in the contacts list.
Task 2. Send a new e-mail message.
Task 3. Enter an appointment in the calendar.
Task 4. Search for the string "yzu" in Google, and link to the Yuan Ze University website.
Task 5. Download software from either the Apple Store or the Android Market.

Procedure

At the beginning of the experiment, participants were given a five-minute basic instruction on button functions, how to type, and the location of common applications on the desktop. Participants randomly began tasks on either smartphone.; they were asked to perform the tasks in the predefined order and then fill-in the questionnaire after completing all tasks on both types of phone. The experiments were timed (a limit was set of 10 minutes for each task), recorded using video cameras, and observed. Participants were informed that they could withdraw from the study at any time if they felt frustrated or otherwise wished to stop.

Measures

Usability measures included effectiveness, efficiency, and user satisfaction regarding interface operations after actual use of various smartphone functions:

1. The percentage of tasks solved (effectiveness)
2. Task completion time (efficiency)
3. Number of detour steps required (efficiency)
4. Questionnaire survey (satisfaction).

In the questionnaire, participants were asked to rate their perceptions of each task on a scale of 1-5 regarding the five satisfaction items: a. ease-of-use, b. icon clarity, c. meet user needs, d. attractiveness, and e. overall satisfaction with the interface design.

EXPERIMENT RESULTS

Effectiveness

Table 2 shows the effectiveness results. Phone-I users showed relatively higher effectiveness in Task 3. Phone-II users showed relatively higher effectiveness in Task 1. For Task 1, after selecting an individual to call from the contact list, Phone-II units clearly indicated the dialing status with the “calling mobile phone” icon, while Phone-I units did not; as such, Phone-I users may not have known what icon to press to make the call. (In reality, they only needed to press on the number to make the call.) For Task 3, the Phone-I users were relatively more effective—no Phone-I user failed to complete the task. However, we found that many Phone-II users did not know how to save the new event in the calendar, or did not know they needed to find the save button; therefore, the success rate for the Phone-II in this task was the lowest in terms of all five tasks.

Efficiency

Table 3 shows the efficiency results. For Task 2 and Task 3, the time span required to complete the tasks was extraordinarily long as compared to other tasks. One reason may be connected to the long character strings participants needed to enter. For Task 2, participants were asked to type specific eight characters in the subject field, and specific two characters in the addressee field with no content in the message field before sending the mail. As required by Task 2, Task 3 asked participants to type specific characters in the subject line and set the time of the new event. Thus, the keyboard design significantly affected the time required to complete the task.

The Phone-I users had relatively lower efficiency for Task 5. We observed many participants could not determine how to begin the download when using the Phone-I: some were misled by the update icon and did not know how to return to the menu.

The Wilcoxon Sign-Rank Test was used to compare the two smartphones in terms of efficiency. The Wilcoxon Sign-Rank Test showed significant differences for efficiency on Task 1 (Phone-II better than Phone-I, $Z=-4.724$, $p<0.0001$) and Task 5 (Phone-II better than Phone-I, $Z=-4.313$, $p<0.0001$). One possible explanation for this result is that Phone-II units often include words to clarify the function, so that participants were able to clearly proceed through each step.

Satisfaction

Table 4 shows the satisfaction results. The questionnaire results show that participants felt pleased when using both smartphones (all scores > 3 for all tasks). In detail, the Phone-II scored slightly higher than the Phone-I in terms of ease-of-use and icon clarity (Tasks 1, 2 and 3), perhaps because the interface design of the Phone-I is intuition-based, users need some time learn and get used to intuition-based operations or icons. However, in terms of meeting users’ needs (Tasks 1, 2, 4 and 5), attractiveness (all tasks), as well as preferences for the overall interface design (Tasks 1, 2, 4 and 5), the Phone-I scored slightly higher.

Table 2. Effectiveness Results

	Task 1		Task 2		Task 3		Task 4		Task 5	
	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II
Success Rate [%]	91.67	100	93.75	95.83	100	89.58	97.92	100	95.83	100

Table 3. Efficiency Results

	Task 1		Task 2		Task 3		Task 4		Task 5	
	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II
Time [seconds]	58.11	33.67	121.44	139.85	131.35	131.72	64.53	56.04	88.22	48.46

Detour steps	3.43	1.5	3.22	2.1	3.67	3.27	1.8	3.5	2.75	2.5

Table 4. Satisfaction Results

	Task 1		Task 2		Task 3		Task 4		Task 5	
	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II	Phone-I	Phone-II
Ease of use	3.67	3.92	3.92	3.4	3.5	3.73	3.81	3.33	3.65	3.65
Icon-clarity	3.81	3.88	3.99	3.4	3.4	3.54	3.96	3.73	3.85	3.77
Meets needs	3.77	3.75	3.71	3.56	3.13	3.25	3.77	3.5	3.69	3.5
Attractiveness	3.85	3.48	3.88	3.5	3.63	3.52	4.06	3.4	3.85	3.42
Interface design	3.71	3.52	3.63	3.58	3.38	3.4	3.69	3.42	3.67	3.44
Average	3.76	3.71	3.82	3.69	3.40	3.49	3.86	3.48	3.74	3.55

CORRESPONDENCE ANALYSIS

This study then conducts the correspondence analysis to summarize positive/negative evaluations of usage and specification for smartphone OS. This study mainly compared the Apple iOS and the Android OS. In addition, this study also included the third smartphone OS, Symbian, for comparison purpose. This study collected Taiwan online user reviews regarding iOS, Android, and Symbian. There are 1718 reviews for both for iOS and Android. However, there are only 217 reviews for Symbian, because it is not as popular as iOS and Android in Taiwan.

The categorical data for usage evaluations include jailbreak, web access, games, applications, and performance. Fig. 2 shows the positive/negative evaluations of usage for smartphone OS. For Android, there are positive evaluations in jailbreak and performance. For iOS, there are positive evaluations in games and negative evaluations in applications. For Symbian, there are negative evaluations in games and web access.

The categorical data for specification evaluations include appearance, photograph, price, OS, specification, and battery endurance. Fig. 3 shows the positive/negative evaluations of specification for smartphone OS. For Android, there are positive evaluations in OS and specification. For iOS, there are positive evaluations in appearance. For Symbian, there are positive evaluations in photograph.

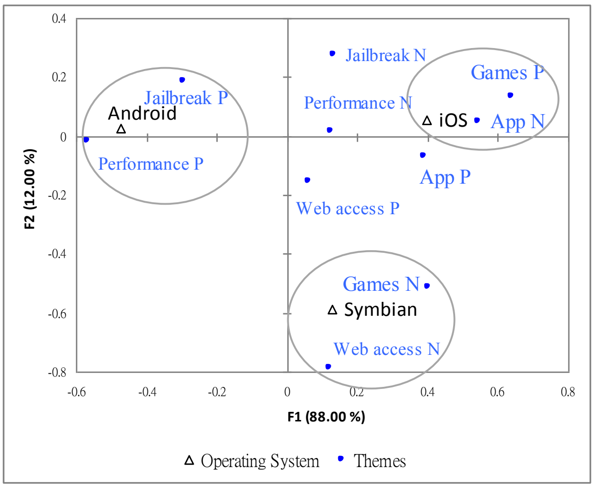

Figure 2. Usage Evaluations of Smartphone OS

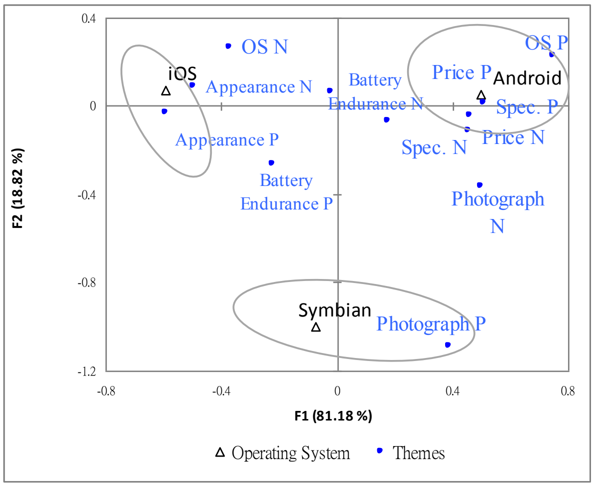

Figure 3. Specification Evaluations of Smartphone OS

CONCLUSIONS

This study evaluates investigates the effectiveness, efficiency, and user satisfaction of two smartphone representational operation system interfaces, Android and Phone OS, across five task contexts. Observations based on the experiment results are listed below in the order that the tasks were successfully performed.

First, Phone-II employs clear textual descriptions to indicate the dialing icon in Task 1, while the Phone-I does not. Although the Phone-I may have an aesthetically pleasing interface design, it is more difficult for users to intuitively select the correct dialing icon, and some participants quit this task for this reason.

Second, all users required significant periods of time to complete Task 2 and 3. The two tasks were very similar, and were the only tasks that required users to enter character strings. It concerns the ratio of a smartphone's virtual or physical keyboard with text entry speed. Text input is an old problem for mobile devices because of the space constraint. This study suggested that a convenient keypad design is required with larger touch-keypad for users to input character strings easier.

Third, the save icon of Phone-II is at the bottom of the rolling screen in Task 3 — some users forgot to save or thought they had already saved before returning to the home screen. Therefore, they failed to complete the task.

Fourth, this study observed that the download button on the Phone-I is difficult to locate in Task 5, and when users attempted to select the toolbar on the bottom of the screen to update or chose another page, they often could not return to the previous page; in turn, some quit the task.

Finally, the positive evaluation results of correspondence analysis are summarized as the following. For Android, there are positive evaluations in OS, jailbreak and performance. For iOS, there are positive evaluations in appearance and games. In addition, there are positive evaluations in photograph for Symbian.

In terms of limitations, the small sample size of participants makes generalizations difficult. Also, there are many smartphone vendors, and new models are frequently released. Thus, the results and observations in this study may not apply far into the future. Further research with expanded participant sets is needed to ensure that the results are truly representational.

ACKNOWLEDGEMENTS

Special thanks to E. Li, T. Li, H. Gee for their assistance in collecting the experimental data; and S. Fang, Y. Huang, Y. Lin for their assistance in the correspondence analysis.

REFERENCES

[1] Huang, S. C. (2006) 'Empirical Evaluation of a Popular Cellular Phone'S Menu System- Theory Meets Practice', *J of Usability Stud*. Vol. 1, No. 2, pp. 136-150.
[2] IDC - Press Release (2010) available at http://www.idc.com/getdoc.jsp?containerId=prUS22333410
[3] International Telecommunication Union (2010) available at http://www.itu.int/en/Pages/default.aspx
[4] ISO/IEC (1998) 'Ergonomic Requirements for Office Work with Visual Display Terminals (VDT)s—Part 11 Guidance on Usability', ISO/IEC 9241-11 (E)
[5] Jin, B. S. & Ji, Y. G. (2010) 'Usability Risk Level Evaluation for Physical User Interface of Mobile Phone', *Computers in Industry*, Vol. 61, No. 4, pp. 350-363.
[6] Jokela, T., Koivumaa, J., Pirkola, J., Salminen, P., & Kantola, N. (2006) 'Methods for Quantitative Usability Requirements: A Case Study on the Development of the User Interface of a Mobile Phone', *Pers and Ubiquitous Comput*, Vol. 10, No. 6, pp. 345-355.
[7] Keijzers, J., Ouden, E. Den, & Lu, Y. (2008) 'Usability Benchmark Study of Commercially Available Smart Phones- Cell Phone Type Platform, PDA Type Platform and PC Type Platform', *Proceedings of the 10th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 265-272.
[8] Klockar, T., Carr, A. D., Hedman, A., Johansson, T., & Bengtsson, F. (2003) 'Usability of Mobile Phones', *Proceeding of the 19th International Symposium on Human Factors in Telecommunications*, pp. 197-204.
[9] Lee, Y. S., Ryu, Y. S., Smith-Jackson, T. L., Shin, D. J., Nussbaum, M.A., & Tomioka, K. (2005) 'Usability Testing with Cultural Groups in Developing a Cell Phone Navigation System', *Proceedings of HCI International*
[10] Niman, B. von, Böcker, M., Schneider-Hufschmidt, M., Flygt, M., Parodi, P., Ketola, P., Williams, D., & Tate, M. (2006) 'Enabling and Improving the Use of Mobile e-Services', http://portal.etsi.org/stfs/STF_HomePages/STF285/STF285.asp
[11] Smartphone Comparison & Reviews (2010) available at http://www.broadbandgenie.co.uk/smartphones
[12] Weiss, S. (2002) 'Handheld Usability'. *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS*, p. 412.
[13] Yamashita, A. F., Barendregt, W., & Fjeld M. (2007) 'Exploring Potential Usability Gaps When Switching Mobile Phones: An Empirical Study', *Proceedings of the 21st British CHI Group Annual Conference on HCI 2007 People and Computers XXI*, Vol. 1.
[14] Ziefle, M. & Bay, S. (2004) 'Mental Models of a Cellular Phone Menu. Comparing Older and Younger Novice Users'. *Lect Notes in Comput Sci*. Vol. 3160, pp. 25-37.
[15] Ziefle, M., Schroeder, U., Strenk, J., & Michel, T. (2007) 'How Younger and Older Adults Master the Usage of Hyperlinks in Small Screen Devices', *Proceedings of the Conference on Human Factors in Computing Systems*, pp. 307-316.