

Comparative Usability Evaluation of Mobile Map Applications

Patarada Thanachan⁺ and Arisara Jiamsanguanwong

Industrial Engineering Department, Faculty of Engineering, Chulalongkorn University, Bangkok 10330,
Thailand

Abstract. The purpose of this research was to evaluate usability of mobile map application. Comparative usability test was performed on map application by using NOSTRA map and Google map. Usability test is not only to identify the problem but also separate the high concerned problem. If the design does not follow usability guidelines or user-center design methodology during development and implementation, the application will create the complexness. In order to improve the quality of mobile application, usability test, the effective technique for product evaluation was focused in this research. The results showed several usability issues which mainly caused by the icon design or difficult vocabulary used in menu.

Keywords: Usability test, mobile application, user interface, user experiences.

1. Introduction

Technology becomes an essential part of lives, especially smartphones which has been dramatically used and developed during past few years [1]. Most of the smartphones were embedded with sensors and devices such as GPS, accelerometer, microphone, camera, and internal memory [2], [3], which allowed it to be compatible with a number of software to response for user needs. At the moment, more than three million mobile app, 24 categories were developed [4] despite the limitations of display size of smartphones [5].

Number of users downloaded the app was often used as the first criteria of successful mobile apps, although it would not give the comprehensive view of accomplishment since the apps might not be used or were uninstalled in short period. According to Localytics research, there were the number of mobile apps downloaded and used just once and uninstalled, also one fourth of them were downloaded only one time and never been used again [6]. Most of the complaints related to the design such as difficulties or bad user experience [7]. As a result, users may uninstall and give the negative feedback [7], especially for complex app with a large number of functions such as map app [8].

There were some evidences that user-friendly and ease of access of Graphical User Interface (GUI) could be used to overcome the difficulties in presenting the information on mobile app with the smartphone display size [9]. GUI is a user interface (UI) that uses computing graphics such as icons and menus [10]. One third of successful mobile apps were reported to have a good GUI design, which was a crucial part of the mobile app quality rating [11]. Moreover, improvement of UI would increase the product competitiveness [7]. The effective well-known technique used for improving UI was the usability test [7].

Usability test is an evaluation method to identify error and user experiences from their design [7]. The ease of human-computer interaction were quantitatively measured [12] and illustrated problems with its severity [1]. Without usability test, the application would have a complexity [13]. Thus, purpose of this study was to evaluate usability of map application of NOSTRA map comparing to the world best practice in navigation app market, Google map [14], [15]. A comparative usability test was conducted to illustrate UI issues and provided some improvement guidelines for designers to improve the usability of mobile map apps.

⁺ Corresponding author. Tel.: + 662-218-6814; fax: + 662-218-6813.
E-mail address: patarada.t@student.chula.ac.th.

2. Methodology

2.1. Participants

The usability test were conducted with one participant at a time with total five participants, two are males. Their age ranged from 23 to 35 years ($M = 28.8$, $SD = 4.38$). All participants had the smart phone and always use it. All participants were novice for both NOSTRA and Google map app.

2.2. Apparatus

The iPhone 5 with iOS 8.4.1 was used as a smartphone in experiments. NOSTRA map version 3.3.1 and Google maps version 4.10.0 were used as mobile map app for comparative usability test. Internet speed was set the minimum of 2mbps during experiments. The experiments were video recorded with Sony NEX-6 camera in private room with the facilities.

2.3. Measurement

2.3.1 Usability attributes

In present study, five usability attributes of learnability, efficiency, effectiveness, memorability, and satisfaction were used regarding to Nielson [12] and ISO 9241-11 [16]. Attribute of an ability to solve error from Nielson were analysed along with effectiveness attribute from ISO 9241-11. These usability attributes were measured as follows:

- Learnability - the time duration (second) to work successfully for the first time
- Efficiency - the time duration (second) to work successfully for experienced user
- Effectiveness - the task success ratio (TSR) by Completion ratio multiplies with Accuracy ratio of Completion ratio (number of application pages / expected number of number of application pages) and Accuracy ratio (actual number of click / expected number of click) [17]
- Memorability - the time duration (second) to work successfully after avoid using system for 5 days
- Satisfaction - Post-Study System Usability Questionnaire (PSSUQ) at the end of usability test [18]

2.3.2 Tasks

Ten tasks were set as scenario for participants to complete usability test for both NOSTRA and Google map application: Register, identify the current location, use hybrid map/ satellite map, searching for place (Chulalongkorn University), save the favourite place, finding the route from current location to domestic airport by not using an express way, searching the nearest LPG gas station, open traffic layer, open favourite place from bookmark, and share location by copy link, respectively. The standard paths and time of each task in NOSTRA map app were collected from three of their designers (Cronbach's alpha 0.988), average time were calculated and used in the analysis (Table 1).

2.3.3 Post-study system usability questionnaire

The Post-Study System Usability Questionnaire (PSSUQ) was used as measurement for the satisfaction attribute of usability in present study [18]. PSSUQ consists of 19 questions with 7-point Likert scales, with strongly agree: 1, and strongly disagree: 7 [18]. Average scores were calculated and used in analysis.

2.4. Procedure

Prior to beginning, participants were informed the purpose of the research and asked to sign a consent form for their participation. Then, they were given a brief instruction that the experiments would be conducted three rounds with 10 tasks for both NOSTRA and Google map application. It would take 30 minutes per a round of an experiment. Instructions, consent form, and tasks given were proved to be clear and quickly understandable to all participants. Time and activities were collected via video recorded.

For the first round, participants performed tasks without any instruction until complete in order to evaluate learnability and effectiveness of the app. Then, participants were instructed and practiced for all tasks before second round of experiment. The objective of the second round was to evaluate efficiency of the application. For the final round, participants performed the same tasks without any instructions after avoided

using both applications for 5 days. The objective of the final round was to evaluate memorability of the app. Participants answered PSSUQ questionnaire before dismissed.

Table 1. Mean and standard deviation of time each rounds of NOSTRA (N) and Google map app (G) and standard time of NOSTRA map app (second)

| | | Task 1 | | Task 2 | | Task 3 | | Task 4 | | Task 5 | |
|--------------------|-------------|---------|---------|---------|----------|---------|---------|---------|---------|---------|----------|
| Std. Time (NOSTRA) | | 39.97 | | 1.82 | | 5.87 | | 14.35 | | 5.44 | |
| Round | Application | N | G | N | G | N | G | N | G | N | G |
| 1 | Mean | 131.70 | 84.31 | 43.84 | 31.26 | 56.38 | 12.94 | 33.75 | 14.73 | 55.41 | 157.91 |
| | (SD) | (64.20) | (44.24) | (43.33) | (57.72) | (48.48) | (9.84) | (14.80) | (5.57) | (32.00) | (129.83) |
| 2 | Mean | - | - | 2.32 | 2.52 | 8.66 | 4.77 | 17.13 | 10.60 | 29.33 | 8.23 |
| | (SD) | - | - | (0.32) | (0.63) | (4.21) | (1.87) | (5.32) | (4.47) | (8.29) | (4.18) |
| 3 | Mean | - | - | 4.39 | 2.46 | 14.83 | 5.93 | 24.05 | 9.78 | 38.13 | 6.64 |
| | (SD) | - | - | (3.32) | (0.33) | (10.19) | (2.50) | (6.30) | (2.57) | (34.22) | (2.87) |
| | | Task 6 | | Task 7 | | Task 8 | | Task 9 | | Task 10 | |
| Std. Time (NOSTRA) | | 32.56 | | 10.82 | | 6.64 | | 5.43 | | 3.46 | |
| Round | Application | N | G | N | G | N | G | N | G | N | G |
| 1 | Mean | 85.29 | 127.20 | 175.99 | 237.51 | 76.70 | 47.28 | 19.00 | 78.37 | 11.75 | 146.98 |
| | (SD) | (17.48) | (91.24) | (55.05) | (184.09) | (57.56) | (92.59) | (16.18) | (48.43) | (8.29) | (164.04) |
| 2 | Mean | 46.79 | 33.60 | 15.87 | 34.71 | 10.89 | 6.84 | 6.84 | 9.81 | 4.75 | 14.70 |
| | (SD) | (32.52) | (5.81) | (4.02) | (20.56) | (1.74) | (4.23) | (2.91) | (3.87) | (1.09) | (14.02) |
| 3 | Mean | 82.63 | 87.51 | 48.17 | 115.05 | 11.07 | 6.71 | 14.88 | 13.71 | 4.45 | 38.25 |
| | (SD) | (51.37) | (51.65) | (42.14) | (110.72) | (1.00) | (1.61) | (10.89) | (7.58) | (0.93) | (61.30) |

Table 2. Task success ratio (TSR) (1st round) of NOSTRA (N) and Google map app (G)

| | | Task 1 | | Task 2 | | Task 3 | | Task 4 | | Task 5 | |
|-------------|--|--------|-------|--------|--------|--------|--------|--------|-------|---------|--------|
| Application | | N | G | N | G | N | G | N | G | N | G |
| U1 | | 1.29 | 3.17 | 1.00 | 1.00 | 26.25 | 4.17 | 5.33 | 1.00 | 33.00 | 46.75 |
| U2 | | 5.44 | 2.67 | 15.00 | 660.00 | 17.33 | 4.17 | 1.00 | 1.00 | 5.00 | 617.50 |
| U3 | | 2.14 | 2.83 | 9.00 | 1.00 | 4.08 | 1.00 | 1.00 | 1.00 | 13.75 | 201.25 |
| U4 | | 1.14 | 1.33 | 825.00 | 9.00 | 18.42 | 1.00 | 2.50 | 3.33 | 1.33 | 1.00 |
| U5 | | 1.14 | 1.50 | 759.00 | 1.00 | 1.00 | 3.33 | 1.00 | 1.00 | 17.50 | 803.25 |
| | | Task 6 | | Task 7 | | Task 8 | | Task 9 | | Task 10 | |
| Application | | N | G | N | G | N | G | N | G | N | G |
| U1 | | - | - | 54.69 | 43.13 | 4.69 | 1.00 | 14.00 | 52.42 | 1.00 | 214.65 |
| U2 | | 2.00 | 58.51 | 19.69 | 4.38 | 177.94 | 1.50 | 14.00 | 3.00 | 4.00 | 45.00 |
| U3 | | 2.10 | 4.86 | 48.75 | 24.44 | 28.50 | 320.67 | 1.00 | 16.33 | 1.00 | 2.50 |
| U4 | | - | 2.02 | 28.00 | 294.56 | 2.50 | 1.00 | 1.00 | 33.33 | 1.00 | 1.00 |
| U5 | | 4.80 | 16.43 | 67.06 | 162.00 | 10.63 | 1.00 | 1.50 | 58.50 | 1.00 | 7.20 |

3. Results

3.1. Initial step

Mean and standard deviations of time each round for all tasks from NOSTRA and Google map app are shown in Table 1. A paired-samples t-test were conducted to compare time between apps each task (1st round). The results revealed that there were significant difference ($p < 0.05$) in time used between apps for task 3, 4, and 9. Participants achieved tasks 3, 4 by using Google map app significantly faster than NOSTRA map app, while they achieved tasks 9 by using NOSTRA map app significantly faster than Google map app for the first time used. These primary results showed that Google map app had a better design in term of learnability than NOSTRA map app for task 3, 4 while NOSTRA map app had a better design in term of learnability than Google map app for task 9. Percentage of time difference between first round and standard time each task of NOSTRA map app were calculated and used as learnability percentage.

For efficiency (2nd round) of NOSTRA map app, percentage of time relative to its standard time each task were calculated. The least three efficiency percentage of NOSTRA map app were Task 5 ($M = 19.96$,

$SD = 6.30$), 8 ($M = 62.09$, $SD = 8.76$), and 7 ($M = 71.33$, $SD = 16.19$). TSRs were calculated (1st round) to illustrate the problems in design of effectiveness of the apps (Table 2). Results showed that there were some misleads in design for task 2, 8 of NOSTRA app, and for task 2, 5, 7, 8, 10 of Google app. For memorability (3rd round), percentage of time difference between third round and second round for each task were calculated and used as memorability percentage in next session. For satisfaction, results revealed no significant difference in PSSUQ scores between NOSTRA ($M = 2.94$, $SD = 0.81$) and Google map app ($M = 2.98$, $SD = 0.51$).

3.2. Defining design problems

For NOSTRA map app, learnability percentage were plotted as Pareto histogram and were listed out major issue tasks as following, task 2, 5, 7, as well as memorability percentage were also plotted and were listed out major issue tasks as 3, 6, 7 and 9. These tasks were created path analysis of users and compare to their standard path in order to identify problems and its root cause as shown in Table 3.

Nine problems were identified from path analysis of 6 major tasks as shown in Table 3. Problems were analysed for its root cause. Most of the problems were found in task 7, 5 as same as other tasks that the icon used in app did not convey the information or were placed in an inappropriate location of the screen. Another issue was that words used on interface were misinterpreted by users.

Table 3. Problems Encountered of NOSTRA Map Application from Path Analysis

| Task | Problem description | Root Cause |
|-------|---|--|
| 7 | Cannot find category, so user typed in order to search. | Keyboard instantly popped up in search page, blinding category provided. |
| 7 | Did not choose a sub category Icon. | Icon sub-category was not easily noticeable. |
| 7 | Did not know how to use a searching location function. | Not obvious what 'searching location function' means. Words used in searching location function were incomprehensible. |
| 5 | Cannot save the favorite places. | List of favorite places button was not in the appropriate display area and was the same icon as the save favorite places. The edit name tab popped up when saved place, participants then had to save location again to record it in favorite. |
| 6 | Get lost to other function. | Function finding route was complicated so that user cannot memorable. |
| 9 | Unable to open the list of favorite places. | List of favorite places button was not in the appropriate display area and was the same icon as the save favorite places. |
| 5 | Unable to show detail result page. | Show detail results bar was not deemed to be press-able. |
| 2 | Did not see current location button. | Not clear icon and color used, and the inappropriate display area. |
| 3 | Cannot chose to Hybrid Map. | Not obvious what 'Hybrid Map' means. Participants always go to 'More' menu. |
| 2,3,7 | Got lost into Measurement Tools function. | Measurement Tools button name was Map Tools, which mislead user. |

4. Discussion and Conclusion

The purpose of this research was to evaluate usability of map apps of NOSTRA comparing to Google. The results showed that several usability issues such as difficulty of use or complex design were found. Most of the problems founded in this study are related to the design of icons and their location in both apps which were inappropriately presented. The redesign of icons and change their location on screen should be considered to improve its usability. For example, the touchable area to open up the menu in task 5 of both NOSTRA and Google map app were hardly noticeable for participants. Some icons may be needed to indicate this functions.

Also, words use in apps should be minimized and confirmed their comprehension concerning individual difference of users. For example, participants were not able to understand the term *Hybrid Map* that made them lost in task 3 of NOSTRA map app. Moreover, when participants lost during any tasks in NOSTRA map app, they often selected *more* menu instead of main menu. Thus, the *more* menu should be integrated into main menu to reduce user confusions. In case of Google map app, they had only the main menu, search box, current location icon, and find route.

Mobile map app is quite specific and complex application which should be carefully designed to provide an intuitive human-centric navigation experience and avoid user experience problems. The improvement in usability design will not only improve user experience, but also enhance the app competitiveness. Although this study illustrated some design problems with its causes which users encountered during completing tasks, but the designers should be aware of these UI problems before launching their product since this is one of key success factor of the app in market. There were also some limitations in this study, as the tasks used in present study were only some part of the app, future study should concerns the whole activities of the app in order to identify possible UI problems.

5. References

- [1] W. Prata, C. R. Mont’Alvão, and M. Quaresma. Usability testing of mobile applications store: Purchase, search and reviews. A. Marcus (Ed.): *DUXU/HCI*. 2013, Part IV, LNCS 8015, pp. 714–722.
- [2] T. Chikaraishi, T. Minato, and H. Ishiguro. Development of an android system integrated with sensor networks. *IEEEERSJ International Conference on Intelligent Robots and Systems*. 2008, **349**: 22–26.
- [3] A. Kushwaha and V. Kushwaha. Location based services using android mobile operating system. *International Journal of Advances in Engineering & Technology*. 2011, **1**(1): 14–20.
- [4] Apple store. Available from <https://itunes.apple.com/en/genre/ios/id36?mt=8>. Retrieved: 15th January, 2016.
- [5] A. P. Chhetri, K. Zhang and E. Jain. A mobile interface for navigating hierarchical information space. *Journal of Visual Languages and Computing*. 2015, **31**: 48–69.
- [6] Global mobile statistics 2013 Section E: Mobile apps, app stores, pricing and failure rates. Available from <https://mobiforge.com/research-analysis/global-mobile-statistics-2013-section-e-mobile-apps-app-stores-pricing-and-failure-rates#mobile-app-flops>. Retrieved: 5th December, 2015.
- [7] C. Tang and X. Yang. Research of Usability Testing Method of Mobile Phone. *Advanced Materials Research*. 2011, **308-310**: 299-302.
- [8] G. Majore and M. Kepka. Taxonomy and Application of EM for Functional Design of Web Map Applications. *Procedia Computer Science of ICTE in Regional Development*. 2015, **77**: 158 – 166.
- [9] M. Łobaziewicz. The design of B2B system user interface for mobile systems. *International Conference on Procedia Computer Science of Communication, Management and Information Technology (ICCMIT)*. 2015, **65**: 1124 – 1133.
- [10] Ó. S. Ramón, J. S. Cuadrado, J. G. Molina and J. Vanderdonckt. A layout inference algorithm for Graphical User Interfaces. *Information and Software Technology*. 2016, **70**: 155–175.
- [11] C.-M. Karat. Cost-Justifying Human Factors Support in Development Projects. *Human Factors Society Bulletin*. 1992, **35**(8).
- [12] J. Nielson. *Usability engineering*. Morgan Kaufmann Publishers Inc. San Francisco, 1993.
- [13] J. van der Linden and C. van de Leemput. Observatory of students’ uses of computer-based tools. *Psychologie Française*. 2015, **60**(2): 145-157.
- [14] Rank History of NOSTRA Map Application. Available from <https://www.appannie.com/apps/ios/app/nostra-map-thailand/rank-history/#vtype=day&countries=TH&start=2012-08-17&end=2015-12-28&device=iphone&view=rank&lm=1>. Retrieved: 5th December, 2015.
- [15] Rank History of Google Map Application. Available from <https://www.appannie.com/apps/ios/app/google-maps/rank-history/#vtype=day&countries=TH&start=2012-12-13&end=2015-12-28&device=iphone&view=rank&lm=1>. Retrieved: 5th December, 2015.
- [16] ISO 9241-11. *Ergonomic requirements for office work with visual display terminals (VDTs)*. Part 11 Guidance on usability, 1998.
- [17] C.-C. Lin. Exploring the relationship between technology acceptance model and usability test. *Springer Science Business Media New York*. 2013, **14**: 243–255.
- [18] J. R. Lewis. IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. Technical Report 54.786. *International Journal of Human-Computer Interaction*. 1995, **7**(1): 57-78.