RESEARCH ON DESIGN SATISFACTION OF STUDENTS ON THE USE OF MOBILE INSTANT MESSENGER INTERFACE DESIGN

Ching-Han Kao [1], I-Chun Lin [2]

[1-2] Department of Industrial Design, Ming Chi University of Technology, TAIWAN.

[1] kaoch@mail.mcut.edu.tw, [2]fiiiiishlalala@gmail.com

ABSTRACT

Portable, wearable mobile products have become increasingly popular in recent years as conduits of interpersonal interaction and communication. However, their immediacy, interactivity, and reduced size have led to new usability problems not found in the previous usage model of PCs. Instant messengers in smart phones have replaced the chat and message functions of traditional tools, becoming mainstream products in social network communications. This study explored the interface usability and user satisfaction of mobile phones, using the most popular social networking apps WhatsApp, LINE and WeChat, in addition to Taiwan's M+ Messenger, as case studies. Our results demonstrate that most subjects were able to carry out tasks in LINE the fastest, followed by M+ Messenger, WeChat and WhatsApp. Subjects generally required a longer period of time to update their personal status. LINE scored the highest on the System Usability Scale, while WhatsApp scored the lowest and was not used to subjects.

Keywords: Mobile Instant Messenger, Interface Design Quality, Smart Phone, Usability Evaluation

INTRODUCTION

T Consumers around the world are irresistibly drawn to mobile devices. There are already five billion mobile phone users worldwide, one billion of whom use 3G mobile services, according to a 2010 survey by Wireless Intelligence, a U.S. mobile communications survey organization. In the third quarter of 2010, up to 19.12 million users in Taiwan accessed the internet through mobile devices, according to data from the National Communications Commission (NCC). Many people have more than one mobile device; for instance, a person may own a mobile phone, network interface controller (NIC) and laptop, indicating that the mobile internet population is growing rapidly. A survey of mobile internet users in Taiwan conducted by Google & Ipsos in October 2012 indicated that smart phones have already become the center of everyday life. Consumers are increasingly dependent on smart phones as their constant companions, with 64% of respondents saying they use smart phones every day and 43% saying they would definitely take their smart phones with them when leaving the house. Another 16% would rather give up their televisions than be without their smart phones. These indicators all clearly demonstrate that advancements in mobile technology have facilitated a giant leap towards a mobile lifestyle. Approximately 79 million mobile phones had been sold worldwide by the third quarter of 2010, an increase of up to 96% over the previous year, according to a Gartner survey. Driven by the mass production volumes of Samsung, LG and HTC, the market share of Google Android, an open platform, has already surpassed iPhone iOs, becoming the world's second largest mobile phone system, topped only by Symbian. Apart from smart phones, the mobile communications market also includes other mobile devices. Driven by the diverse services of mobile technology, by 2015 there will be 10 million mobile device users in Taiwan and the development of the mobile market will

be even more robust, according to estimates by the Market Intelligence & Consulting Institute. The attraction of mobile devices is largely based on the accompanying range of applications (apps); users can download apps to a wide range of needs from app stores offered by mobile platforms; for instance, iOs (Apple)'s platform is termed App Store; Android's is Google Play, and Windows Mobile (Microsoft)'s is Windows Marketplace. The social communications app market owes its existence to continued sales of smart phones, which not only drive the downloading of mobile instant messengers (MIMs) but have also led to MIM software becoming a new arena of social networking. This study researched free MIMs, which are more popular with users, and explored how satisfied users are with interface configuration and usability. The results could serve as reference for businesses involved in designing related apps.

LITERATURE REVIEW

Instant Messenger

Instant Messenger (IM) software was first used in military radio communications systems as a tool for emergency response and instant communication. In 1988, a Finnish man named Jarkko Oikarinen developed internet Relay Chat (IRC), a service that enabled users to transmit instant messages, which quickly became popular with young people. The first widely used IM software was ICQ, launched in 1996 by Mirablis, a company established by four young Israeli engineers. Taking the phonetic sound of "I Seek You", ICQ provided the Worldwide Web of the time with a new method of communication. When their computers were connected to the internet, users could chat with others through the ICQ window, which also offered functions such as file transfer, chat rooms and storage of dialogue records; ICQ had more global users than any other IM software of the time (Leung, 2001; Cunningham, 2003; Marquez, 2003).

Mobile devices gradually emerged as internet technology progressed, and IM moved from PCs to mobile phones. Mobile phones, being portable devices, made interpersonal communication more convenient. The greatest difference between PC and mobile versions of IM was that the former required both users to be at their computers; if a user left the computer he/she would not be able to communicate with the other user instantly. With MIM installed on mobile phones, however, communication was much easier as most users take their mobile phones everywhere. The earliest MIM software for mobile devices was WhatsApp, which was developed in 2009 and in its early stages enjoyed a near global monopoly. Next, the LINE app was developed, attracting users with cute, quirky stickers; WeChat also has a strong following in China.

Small Screen Design

Mobile devices are much smaller than PCs and other fixed devices. Due to continued technological advancement and the emphasis on mobility, previous key-controlled interfaces have gradually been replaced by touchscreens, which have in recent years become the dominant type of interface for increasingly smaller, slimmer, and lighter mobile devices. Mobile phones are mobile devices with smaller screens, approximately 2-5 inches in size; larger mobile devices, known as tablets, e-book readers etc., have screens as large as 9-10 inches and are gradually replacing computers.

Due to the small screens on smart phones, particular attention must be paid to usability in their design. Marcus (2001) called this type of small screen 'babyface' and indicated that small screen design could lead to limited spatial and color resolution, limited font choice, and limited information visualization. The visual browsability of small screens must be

Leena and Luna International, Chikusei, Japan.
(株) リナアンドルナインターナショナル, 筑西市,日本
ISSN: 2186-845X ISSN: 2186-8441 Print
www.ajmse. leena-luna.co.jp

considered in their design, as they must be able to display sufficient information in a limited space. Currently, smart phone tools and apps are displayed as icons, as image and symbols transcend cultural and linguistic limitations, making them easily read and understood by everyone. The operational use of smart phones is very different from PCs; even in a mobile state they must facilitate operations such as quick browsing or single-handed input. Marcus (2001) indicated that interface design must include the following five elements: Metaphor, Mental Model, Navigation, Interaction, and Appearance. Currently, the main type of interface used in smart phones, a touchscreen, is operated intuitively and is more in line with our usage environment. Due to special features of this type of interface, design principles applicable to a variety of interfaces have been developed.

Different concepts are used in interface design based on the size of the equipment. Design of small screens is more limited and requires greater care to provide the user with an optimal experience.

User's Mental Model

One's ability to operate unknown objects is limited, but a good mental model can facilitate successful engagement. Mental models are influenced by the cognitive experience of users. Cognition can be classified as experiential or reflective: the former represents the extent to which we perceive and respond to our surroundings, while the latter deals with thought and comparison (Preece, 1998). Craik (1943) defined a mental model as an internal representation of reality, suggesting that these symbolic representations are manipulated for purposes of reasoning, action and cognition, to understand the link between symbolic representations and external reality. Norman (1988) indicted that a mental model is built on links established between analogies, prior knowledge, and new situations. According to Norman (1988), designers must note the following four principles in the design of a user interface: (1) Conceptual model – combine the viewpoints of the designer and users, and develop an accurate mental model to reduce error; (2) Visibility – Ensure that there is a good inter-relationship between the interface and functions, and ensure that feedback is provided to facilitate a positive user experience; (3) Mapping – Use correct controls to create good interaction between the system and its users, and provide users with appropriate feedback; (4) Feedback – Users must have good operational feedback in order to know whether the action has been completed; poor feedback may lead to errors. Good interface metaphors can guide the user to build the right mental model and enjoy a positive experience; metaphors use concrete methods to represent abstract concepts that are more difficult to understand. The mobile phone and computer images in Fig. 1 show common metaphoric concepts, which can provide clues to help users understand the symbolic meaning of icons. Metaphors are even more important in smart phones due to reading limitations, as shown in Fig. 2.

Fig. 1. Common computer and mobile phone icons

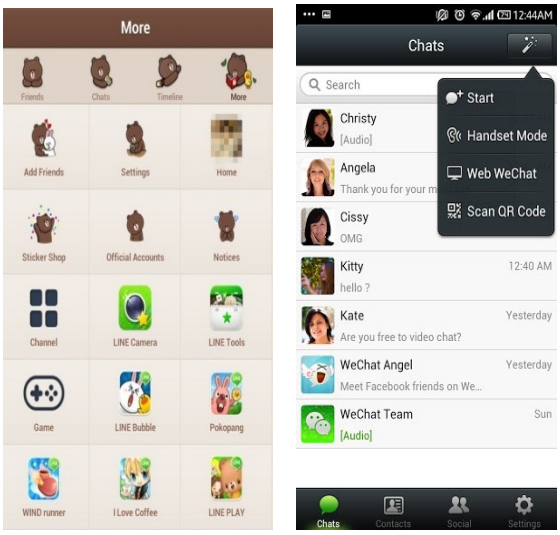


Fig. 2. LINE & WeChat interfaces

Interface usability

Due to intense competition among the wide range of software available, users have gradually begun to focus on the usability of the interface. Usability is a user-centered concept; the goal of design is to match product design with the habits and needs of the user. When evaluating the usability of a product, we must also consider its usage background. ISO 9241 (1988) defines usability 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". Grundin (1992) defined usefulness as a measure of how well a system can be learned, indicating that usefulness can be divided into utility and usability. Preece (1998) claimed that usability can improve human-computer interaction, enabling users to safely, realistically and efficiently operate computers. According to Nielsen (1994), useful design in human-computer interaction must cover the following five attributes, which are used as evaluative criteria: (1) Learnability: How easy is it for users to accomplish basic tasks the first time they encounter the design? (2) Efficiency: Once users have learned the design, how quickly can they perform tasks? (3) Memorability: When users return to the design after a period of not using it, how easily can they reestablish proficiency? (4) Errors: How many errors do users make, how severe are these errors, and how easily can they recover from the errors? (5) Satisfaction: How pleasant is it to use the design? Users are not necessarily in a static environment when using smart phones and may be operating these devices on the go; therefore, usability issues are even more important when applied to a smart phone interface.

RESEARCH METHODOLOGY

This study was divided into two phases: the first involved the evaluation of usability through the completion of four operational tasks, in which the amount of time required to complete these tasks was used to evaluate the objective performance of users. In the second phase a questionnaire survey was conducted to understand the subjective assessments of respondents, who were also interviewed about whether they had encountered problems while performing tasks. The questionnaire was the 5-interval Likert System Usability Scale (SUS), which was used to evaluate the usability of four MIMs.

Leena and Luna International, Chikusei, Japan.
(株) リナアンドルナインターナショナル, 筑西市,日本
ISSN: 2186-845X ISSN: 2186-8441 Print
www.ajmse. leena-luna.co.jp

RESEARCH METHODS AND PROCEDURES

Introduction to software

Because there are many MIMs currently on the market, this study selected four relatively representative MIMs as case studies, to evaluate their interface usability. These four software models are outlined below:

(1) LINE: Developed by Japanese company NHN, Line was launched in 2011 and immediately became popular. LINE, unlike other software, allows users to use virtual stickers in addition to inputting text when chatting. The wide range of diverse stickers allows users to instantly express their feelings without inputting any text. These fun and quirky stickers are emotive and entertaining. The idea of sharing one's feelings using LINE has also been aggressively promoted through TV ads, which increased awareness of the app. As of July 2013, LINE had been downloaded more than 200 million times and is currently the most popular MIM. Its main features are free voice and video calls, ability to send videos and voice messages, social networking services, sticker shop, official account, multiple ways to add friends, etc.

(2) WhatsApp: Created by Jan Koum (formerly a senior employee with 20 years of experience at Yahoo), WhatsApp was inspired by the English expression 'what's up?' and in its early stages had a monopoly on the communications app market. Its main functions are messaging and dialogue. Currently, iOS users must pay to download WhatsApp, while Android users can use the app free of charge for one year but are then required to pay for continued use. This arrangement initially caused considerable loss of users. Its main features are multimedia, ability to send photos and videos, group chats, ability to say no to pins and usernames, no need to add buddies, and ability to leave offline messages.

(3) WeChat: Developed by Tencent, WeChat has been downloaded over 30 million times in China. Its main features are ability to send photos and videos, video chat, group chats, emoticons, and connection to Facebook. It is also considered to be a find-a-friend platform and uses plug-ins to expand its functions. For example, the purpose of 'shake' and 'look around' is to enable users to meet new friends.

(4) M+ Messenger: Developed by Taiwan Mobile, M+ Messenger originally supported only basic photo, text and map functions; however, voice messaging and free call functions were added later. One of its unique functions is 'pre-arranged messaging', which allows users to input a message and then designate a date on which the message is to be transmitted. This function is particularly useful for holidays. The telecom tagging function in the phonebook has also been well-received. Currently, M+ Messenger has been downloaded over five million times in Taiwan. Users can interact with their friends using not only basic text and voice messaging but also dynamic stickers and phone gift functions.

Research process

Subjects performed the operations shown in Table 1 using the four types of MIM software. We recorded the time required to complete each task and observed the subjects during the operation. The experimental process is outlined below:

(1) Subjects were permitted to spend a total of 10 min operating each type of software for 1-2 minutes prior to testing. This enabled subjects to have a basic understanding of the four types of MIM and an opportunity to ask any questions they may have.

(2) Following the experimental procedures and specific tasks were explained to subjects, they tested each of the four types of software (see Table 1); the order of tasks was randomly arranged.
(3) Code tables were used to calculate the time required for subjects to complete each task.
(4) After each MIM experiment was completed, we interviewed subjects about their thoughts on the experimental process.
(5) Subjects were asked to complete the SUS; all procedures (3)~(5) were repeated four times.
(6) We analyzed the time required to complete tasks, qualitative interview data and subjective levels of satisfaction.

Table 1. Operational tasks

Task	Objective of experiment
Send text message	Explore whether subjects are influenced by existing mental models when operating different interfaces.
Send graphic message	Understand how different icon designs for the same software functions influence operation.
Update status	Explore how subjects are affected if the same functions, such as personalized signatures and status, have different names and locations.
Update profile image	Explore how subjects are affected if the same functions, such as photos and personal images, have different names and locations.

Measurement items

This system measured experimental (operational performance of users) and subjective (system usability) scales.

Experimental: Time required for completion refers to the length of time required for subjects to complete each task; we counted time in seconds, starting from after instructions had been given and finishing when the task was completed. Greater length of time implied poorer operational performance.

Subjective: After subjects had completed the experiments, we used SUS to collect data on their subjective assessments, asking them to describe any difficulties they encountered during the operations and provide specific suggestions.

Questionnaire

The questionnaire consisted of a basic demographics section and the SUS, both of which are explained in detail below.

The System Usability Scale

Developed by John Brooke in 1996 while employed at Digital Equipments Co. Ltd in the UK, the SUS is a subjective scale frequently used in product usability research. This is a 5-interval Likert scale consisting of 10 questions for which the full score is 100 points, employing positive/negative cross-questioning techniques to measure the subjective assessments of respondents. The results are then quantified into usability satisfaction scores.

The scale is often employed by researchers or corporations to assess system usability because it provides a simple method of collecting and analyzing data. The calculation method is to subtract 1 from the original scores of items 1, 3, 5, 7, and 9, subtract 5 from the original

Leena and Luna International, Chikusei, Japan.
(株) リナアンドルナインターナショナル, 筑西市,日本
ISSN: 2186-845X ISSN: 2186-8441 Print
www.ajmse. leena-luna.co.jp

scores of items 2, 4, 6, 8 and 10, and then multiply the sum of scores by 2.5 to obtain the final SUS score. Table 2 introduces the SUS.

Table 2. System Usability Scale

No.	Question
1	I think that I would like to use this system frequently.
2	I found the system unnecessarily complex.
3	I thought the system was easy to use.
4	I would need the support of a technical person to be able to use this system.
5	I found the categorized result in this system were well sensible.
6	I thought there was too much inconsistency in this system.
7	I would imagine that most people would learn to use this system very quickly.
8	I found the system very cumbersome to use.
9	I felt very confident using the system.
10	I needed to learn a lot of things before I could get going with this system.

RESEARCH RESULT

Analysis of sample structure

Our sample group consisted of 30 students, with equal numbers of males and females. In the 21-25 age bracket were 21 subjects; in the 20 or younger bracket, 8; and in the 26-30 bracket, 1. Fourteen subjects were undergraduates and sixteen were postgraduates. Fig. 3 presents the MIMs that subjects had previously downloaded. As indicated by the figure, LINE was the most frequently downloaded MIM in this study, followed by Facebook and then Skype, WeChat and WhatsApp. The most frequently used software was also LINE, which indicated that subjects were more familiar with LINE and Facebook (see Fig. 4).

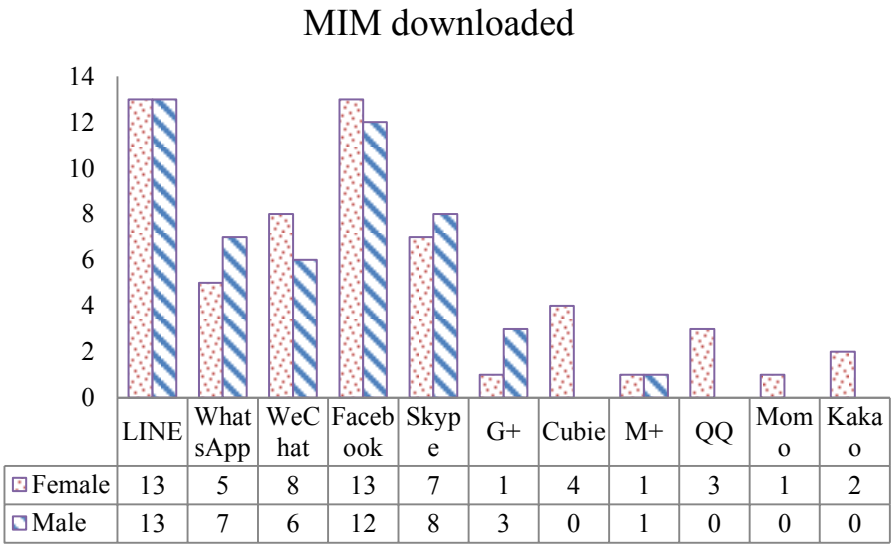
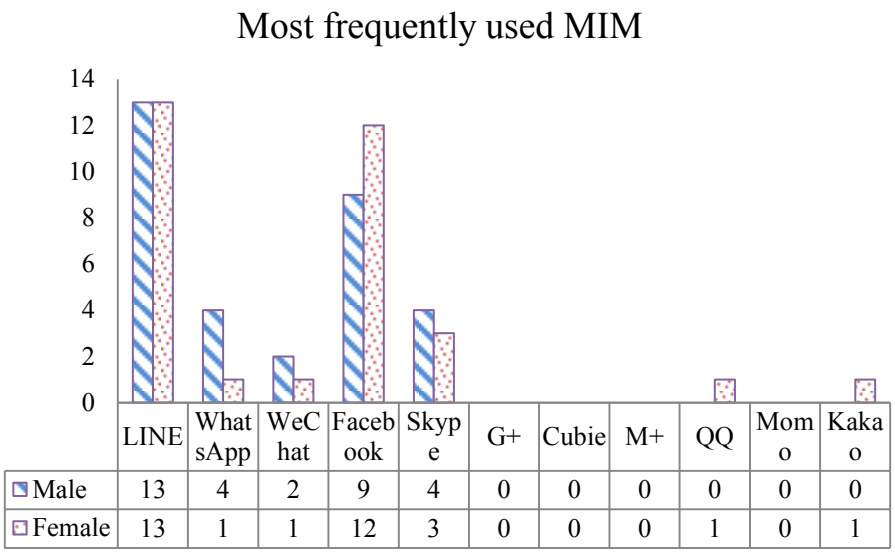


Fig.3. MIM downloaded



	LINE	WhatsApp	WeChat	Facebook	Skype	G+	Cubie	M+	QQ	Momo	Kakao
Male	13	4	2	9	4	0	0	0	0	0	0
Female	13	1	1	12	3	0	0	0	1	0	1

Fig. 4. Most frequently used MIM

Usability Test Results

(1) Efficiency:

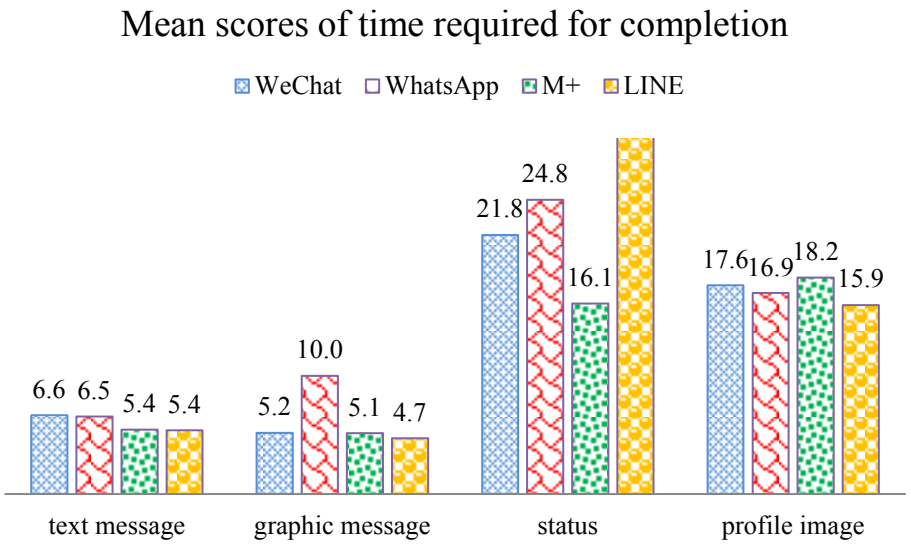


Fig. 5. Mean scores of time required for completion

The test scores (Fig. 5) demonstrate the performance of users in operating these four types of software. The time required to send text messages and update profile images did not differ significantly, although subjects were fastest in sending text messages. Subjects generally took longer to update statuses, possibly because they did not often use this function or the interface options were unclear, according to post-test interviews.

(2) SUS analysis

The mean SUS scores for each of the four MIMs were as follows: LINE-67.8, WhatsApp-53.6, WeChat-64.5 and M+ Messenger-59.4; LINE had the highest overall mean score, while WhatsApp had the lowest, although the differences among the four MIMs overall were not significant.

Leena and Luna International, Chikusei, Japan.
(株) リナアンドルナインターナショナル, 筑西市,日本
ISSN: 2186-845X ISSN: 2186-8441 Print
www.ajmse. leena-luna.co.jp

We then separately analyzed each item of the SUS. For Q1 "I think that I would like to use this system frequently", LINE scored the highest with an average of 3.3, followed in descending order by WhatsApp1.9, WeChat2.6 and M+ Messenger2.4. This shows that subjects favored LINE's interface. The mean scores for Q2, "I found the system unnecessarily complex", were 1.9, 2.3, 2.1 and 1.7, respectively. These relatively low scores indicate that subjects did not feel the interface design of the software was complex. The mean scores for Q3, "I thought the system was easy to use", were 3, 2.1, 2.6 and 2.3, respectively, which were not significantly different.

The mean scores for Q4, "I would need the support of a technical person to be able to use this system", were 2.3, 1.6, 2.5 and 2.4; only WhatsApp had a relatively low mean score, indicating that subjects felt the WhatsApp interface was easier to operate compared to the other three. The mean scores for Q5, "I found the categorized results in this system were sensible", were 3, 2, 2.6 and 2.7, indicating that subjects felt the interfaces of all MIMs except WhatsApp were relatively well-integrated. The mean scores for Q6, "I thought there was too much inconsistency in this system", were 2.1, 1.8, 2.4 and 2.2, which were not significantly different, indicating that the consistency of the interface design of these four MIMs was acceptable to subjects. The mean scores of Q7, "I would imagine that most people would learn to use this system very quickly", were in order as follows: 3.1, 2.7, 2.9 and 2.8, indicating that subjects were generally satisfied with the learnability of these four types of software. The mean scores for Q8, "I found the system very cumbersome to use", were 3.1, 2.7, 2.8 and 2.6, respectively, indicating that subjects were somewhat confused when using the interfaces. The mean scores for Q9, "I felt very confident using the system", were 2.8, 2.3, 2.8 and 2.5, respectively; these scores did not differ significantly, indicating that most subjects were capable of correctly operating these four types of software. The mean scores for Q10, "I needed to learn a lot of things before I could get going with this system", were 2.5, 2.2, 2.4 and 2.1, respectively. These scores approached the median and did not differ significantly, indicating that subjects did not have strong feelings on the issue of requiring additional learning to operate the software.

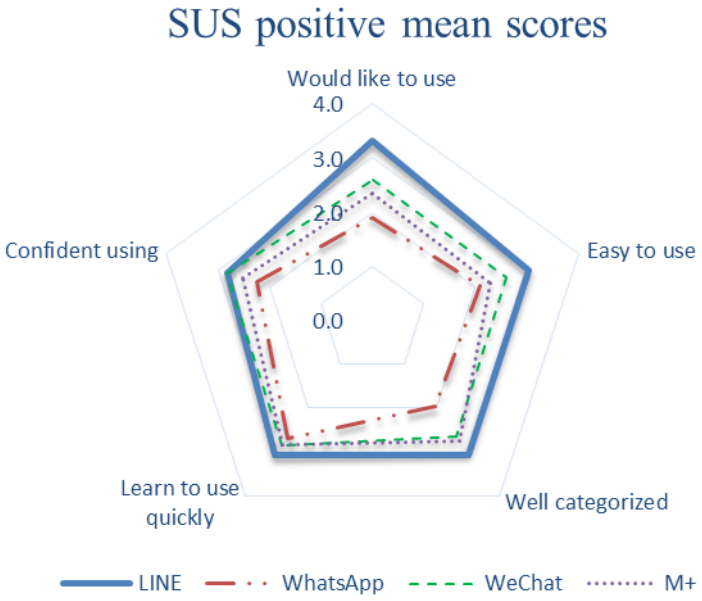


Fig. 6. SUS positive mean scores

The further away from the center, the higher the positive SUS mean score obtained by the software in question. Fig. 6 illustrates that LINE, as the software more frequently used by subjects, obtained the highest score, while WhatsApp scored the lowest. WhatsApp was one of the earliest MIM, and its interface configuration is different from the other three; therefore, subjects may have been less familiar with WhatsApp, resulting in a lower score.

(3) Interviews

According to interview data, subjects found that updating their personal status (personalized signature) was the most difficult of the four tasks, because it involves more steps and different types of software provide somewhat different narratives, which increase the difficulty of recognition. Because most subjects were more familiar with LINE and the interface of WhatsApp displayed functions at the bottom, which was a different design from the other three, respondents generally expressed confusion about the WhatsApp interface and were unable to operate it based on existing mental models, requiring more time to learn its operation.

DISCUSSION & CONCLUSION

The objective of this study was to explore the interface usability of MIMs and test the usability of four currently relatively well-known MIMs. Results show that subjects were faster and more efficient when completing tasks using LINE, because they were more familiar with this software. The resulting SUS scores for LINE were also considerably higher than the other three types of software tested. Although WhatsApp has a niche in the mobile communications industry, its interface is different from the other three and it obtained the lowest scores, requiring relatively longer amounts of time for tasks to be completed.

Because the age of MIM users is increasing, and users from different backgrounds have different viewpoints on usability, future studies could sample subjects of more diverse age and background. This study also did not discuss levels of satisfaction or learning constructs, which will be incorporated for the sake of completeness into future research.

Leena and Luna International, Chikusei, Japan.
(株) リナアンドルナインターナショナル, 筑西市,日本
ISSN: 2186-845X ISSN: 2186-8441 Print
www.ajmse. leena-luna.co.jp

REFERENCES

[1] Craik, K. (1943). *The Nature of Explanation.* Cambridge: Cambridge University Press.

[2] Cunningham, P. J. (2003). IM: Invaluable New Business Tool or Records Management Nightmare? *Information Management Journal-Prairie Village*, *37*(6), 27-33.

[3] Grudin, J. (1992). Utility and usability: Research issues and development contexts. *Interacting with computers*, *4*(2), 209-217.

[4] International Organization for Standardization. (1998). *ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Part 11: Guidance on Usability.*

[5] Jennifer, P., Yvonne, R., & Helen, S. (2002). Interaction design: beyond human-computer interaction. *NY: Wiley*.

[6] Leung, L. (2001). College student motives for chatting on ICQ. *New Media & Society*, *3*(4), 483-500.

[7] Marcus, A. (2001). Babyface design for mobile device and the web. In S. J. Michael & S. Gavriel (Eds.), *Human-Computer Interface internet (HCII) 2:*514-518. Mahwah, NJ: Lawrence Erlbaum Associates.

[8] Marquez, J. (2003). The effect of instant messaging on the social lives of students within a college dorm. Retrieved December 12, 2006, from http://www.stanford.edu/class/pwr3-25/group2/projects/marquez.html

[9] Nielsen, J. (1994). Enhancing the Explanatory Power of Usability Heuristics. *CHI'94 Conference Proceedings*, 152-158.

[10] Norman, D. A. (1988). *The Design of Everyday Things.* New York, NY: Basic Book.

[11] Preece, J., Rogers, Y., Benyon, D., Simon, H., & Carey, T. (1998). *Human-Computer Interaction*. Harlow, UK: Addison-Wesley.