The Usability of Metaphors with Different Degree of Abstract in Interface Design

Ming-Chuen Chuang and Inglen Lo

National Chiao-Tung University, Institute of Applied Arts,
Taiwan
cming@cc.nctu.edu.tw, inglenlo.aa89g@.nctu.edu.tw

Abstract. Recently, more and more devices everywhere are getting "smarter" with a multi-modal hierarchical menu and form interface. One of the main points of the menu or interface design is to provide users with ease-to-use operation environment. This make them not only learn efficiently but also feel fun (interested) in the process of learning. However, there is no one concept of design suit everyone because the needs and purposes of users are much different from individuals. To satisfy them, the varied design concepts have been suggested to fit for their distinct perceptions and experiences. Consequently, new assessment, called usability, is also required to estimate whether the design concepts are good or not. Therefore, this study attempted to investigate into the usability of 3D interface design. For that, 3 types of main menu of the mobile phone's interface metaphor design were developed as stimuli with different degree of abstract in this study. Then, a four-phase experiment was conducted to explore the usability evaluation of 3 types of metaphorical interface design with different degree of abstract, including: (1) to investigate users' opinions on a mobile phone's interface design; (2) to verify whether the simulated graphics and interactions corresponding to the metaphors intended (pilot study); (3) to measure the usability of 3 types of metaphorical interface design simulated in this study; (4) to compare the preference for any one of the 3 types of metaphorical interface design. The experimental procedures and the results of the analysis would be interpreted respectively according to different phases. Additionally, the degree of abstract in the metaphorical interface design was defined by the average ratings in phase 1: metaphor 3 were regarded as abstract interface design and metaphor 1 and metaphor 2 were regarded as concrete interface designs, but the degree of concrete in metaphor 1 was stronger than in metaphor 2.

Keywords: Metaphor, interface design, usability, SUS score, preference.

1 Introduction

This is a feasibility study focused on 3D interface metaphors of mobile devices. There were three reasons explained why the cellular phone's interface was decided as typical stimuli in this study: (1) the cellular phone is the most popular one of mobile product; (2) varied but short lifestyle of mobile phone product is caused by users'

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 343–352, 2007.

fickle in affection; (3) there is a tendency towards 3D interface of products but limited in applying to mobile phones. Also, despite plenty of design researches into mobile phone products, more of them focus on examining the essentials of exterior design. Consequently, the reformation of products' appearance design is becoming faster and faster except their interface design. Later, some researchers found that and made studies of the icon design towards Kansei impressions but limited in present mobile phone products.

Thereore, there are three purposes of this study: (1) to investigate users' opinions on a mobile phone's interface design; (2) to verify whether the simulated graphics and interactions corresponding to the metaphors intended (pilot study); (3) to measure the usability of 3 types of metaphorical interface design simulated in this study; (4) to compare the preference for any one of the 3 types of metaphorical interface design.

1.1 Design Principles of User Interface Design

The user interface is a medium for communication between users and computational devices (systems). As a result, the appraisal of usability is directly depended on whether the user interface design is good or bad. Therefore, there are various principles generalized from lots of researches into the user interface design so far. For examples, Microsoft (1995) defined 5 bases of interface design, including consistent, learnable, intuitive, extensible and attractive; Sutcliffe (1983) suggested 6 fundamentals to the software interface designs, information structure, consistency, compatibility, adaptability, economy and guidance not control included; Norman (2000) addressed 5 principles of good interface design as follows: affordance, good conceptual model, good mapping, feedback and visibility. Based on the opinion of user's experience, Veryzer (1999) recognized four essential design properties as the characteristics of interface designs: operation, understanding, structure and value. Furthermore, 5 frameworks of interface designs towards usability problems were advised by Nielsen (1993), including easy to use, efficient to use, few error, subjectively pleasing, easy to remember. To conclude, that ease-to-learn of operational environment to make users learn effectively is the essential principle of user interface design.

Recently, graphical user interface (GUI) is praised for ideal interactive environment in common. It mainly contains five elements: screen, windows, icons, menus and point devices. Sanders and McCormick (2002) defined them as : display, operational environment, control panel, integration of control and display, and information structure. However, there are also potential problems with GUI, such as lack of real touch and immediate responses, misunderstandings of symbolic cognition from the different contexts of cultures, overloading derived from too many invisible multi-layers, less interests of learning resulted from trivial operating sequences, etc. Therefore, designers need some guidelines, procedural advice and computational support in using concepts for their design problems (Carroll & Rosson, 1994). In response to the needs, metaphors are widely applied to visualize and specify the structure of GUI design (Kobara, 1991).

1.2 Metaphor

In the early 80s, Lakoff and Johnson (1980) offered a new vision of cognitive processes as the capacity of projecting from a given (well known) domain to a new (less known) domain. This capacity of mappings between domains is immediately recognized as "metaphor". Later, they (1987) developed their ideas of metaphor into a modern approach to cognition, known as experientialism. One of the main concepts of experientialism is the image-schema which is abstract pattern derived from our bodily experience and other (everyday) interactions with the external world. Briefly, the essence of user interface design is an appropriate metaphor which is intuitively related to their cognition, experience and knowledge (Erickson, 1990); the core of a metaphor is the understanding and experiencing one kind of thing in terms of another. After that, metaphors as a design source are often applied in designing the user interfaces to provide the underlying images, terms and concepts that make communication possible at all (Marcus, 1993). Also, the metaphors should be an intuitive access to all functionality. That means the user should be not only familiar with the metaphor domain but also able to perform the mapping between the metaphor domain and the application domain.

In conclusion, for designers, a good metaphor is necessarily to look after both sides of functionality (information structure) and visualization (aesthetics and consistency, etc.); for users, a good metaphor of interface design is necessarily able to help themselves get applicable instructions in procedures to reduce the chances of mistakes happened.

1.3 Usability Measurements: System Usability Scale (SUS) and Interview

Initially, usability was defined as the degree of efficiency and effectiveness of use within a specified range of users, tasks, tools, and environment (Bennet, 1984; Shackel, 1984), which results in over 87% of usability researches, especially in HCI domain, focus more on measuring objective estimations of effect on goals (Nielsen and Lavy, 1994). To a certain extent, the product performances based on the subjective measurements are actually better than others but uncertainly make users be satisfied with the product itself. In fact, user's satisfaction is considered as one of emotional expression.

To sum up, the emotional reactions accompany objective evaluations but are unconsciously operated to influence the degree of satisfaction (Zajonc, 1980). So, the subjective emotional issues should no be excluded from usability problems. In other words, usability should simultaneously take both objective measurements and subjective assessments into account (Norman, 2002, Fig 1). Therefore, the ISO9241-11 (1998) re-describes usability as "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". Further, "effectiveness" means the ability of users to complete tasks using the system, and the quality of the output of those tasks; "efficiency" means the level of resource consumed in performing tasks; "satisfaction" means users' subjective reactions to using the system.

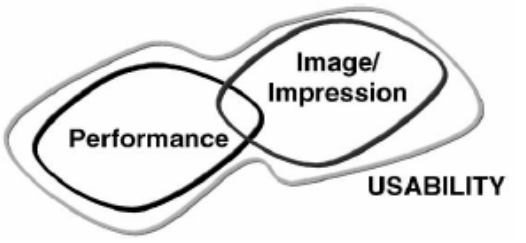

Fig. 1. The concept of Norman's usability

In general, the demands of evaluating usability of systems within an industrial context are often neither cost-effective nor practical to perform a full-blown context analysis and selection of suitable questions. But, users could be very frustrated if they were presented with a long questionnaire. It was very possible that they would not complete it and there would be insufficient data to assess subjective reactions to system usability. Nevertheless, what we need is a general indication of the overall level of usability of a system compared to its competitors or its predecessors. That means the measure had not only to be capable of being administered "quickly" and "simply" but also to be "reliable" enough to be used to make comparisons of user performance changes from version to version of a product.

In response to these requirements, the System Usability Scale (SUS) was developed. It is a simple, ten-item Likert scale giving a global view of subjective assessments of usability (cited from John Brooke). Generally, it is used after the respondent using the system (interface) being evaluated but before any discussions takes place. Due to the numbers for individual items are not meaningful on their own, SUS scores should be calculated according to following steps: first sum the score contributions from each item (each item's score contribution will range from 0 to 4) and then multiply the sum of the scores by 2.5 to obtain the overall value of SU. That means SUS scores have a range of 0 to 100.

Sometimes, interviews are also regarded as usability evaluation approaches. In Rubin's model (Rubin, 1994), interviews are not only used in the beginning of development stage to design the questionnaire but also used in the last stage of the evaluation stage of usability testing to clarify user responses and to collect additional information. Interviews are of two types: Structured and Open-ended (USINACTS Usability, 2004). Generally, structured interviews can be designed rigorously with a predefined set of questions to avoid biases and usually provide more reliable and quantifiable data than open-ended interviews. Later, the structured interviews were also applied in the last phase of the usability evaluation in this study.

2 Method

Initially, this was a cooperative project to propose applicable metaphors based on 3-D graphic to the interface of mobile phone product. After a succession of discussions and improvements, 3 types of metaphorical interface design with different degree of abstract were selected, simulated and constructed as Fig 2, Fig 3 and Fig 4 below. In response to the purposes of this study above, three phases of empirical experiments were proceeded as follows:

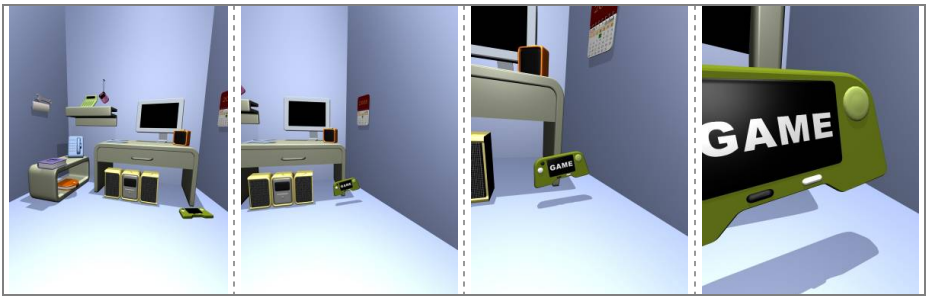

Fig. 2. Metaphor 1 - mobile phone's interface as personal room

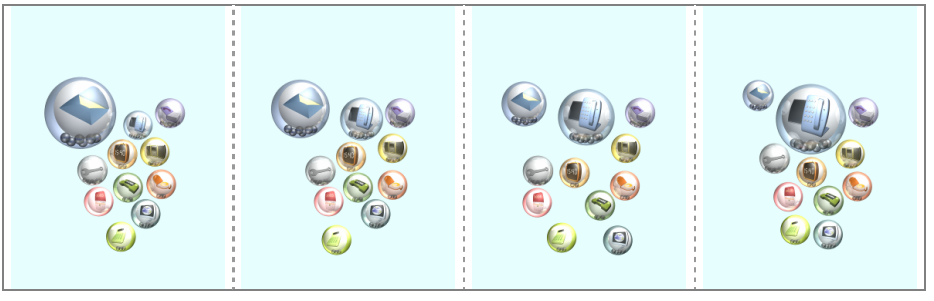

Fig. 3. Metaphor 2 - information structures as transparent bubbles

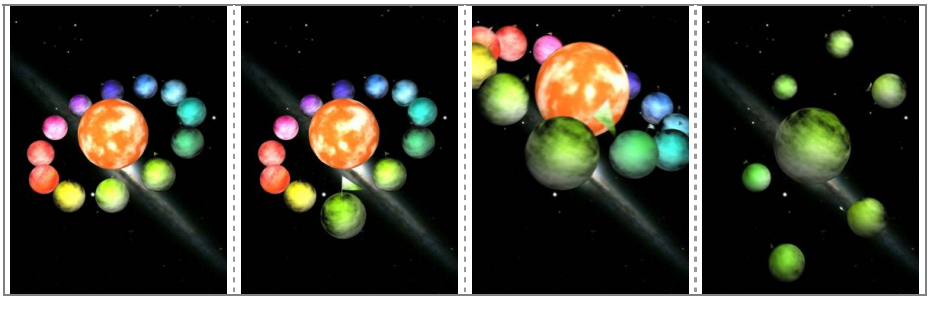

Fig. 4. Metaphor 3 – search for functionalities as an exploration of interplanetary

2.1 Phase 1: User's Response to Mobile Phone Products

Seventy-three people participated in this questionnaire survey, 43 of whom were students, 25 of whom were workers, and 5 of whom were on active duty of military service. There were 27 female and 46 male participants ranging in age from 19 to 35. The average age was 25.05 (SD = 3.01). On average, they have had 3.66 personal mobile phone products (SD = 1.25), ranging from 2 to 8.

In order to gather user's opinion on mobile phone products, participants were asked to respond to 3 similar questions with 14 possible answers prepared in advance

(13 of them were specific and 1 of them were any of other possibilities, shown in Table 1). The 3 questions were describe as: (1) which 3 of factors were the most important elements to evaluate whether a product of mobile phone is good or bad; (2) which 3 of factors were attractive essentials to individual to want to have a product on an impulse; (3) which 3 of factors were the most important issues if individual wanted to buy a new one. Also, participants were instructed to mark the factors in order according to the extent of importance (1 meant the most important factor, 2 was next, and 3 was next again).

Table 1. 14 possible answers to individual's opinions on mobile phone product

01. appearance	02. functionality	03. interface manipulation	04. graphic design of interface	05. size
06. weight	07. brand	08. price	09. quality	10. place of production
11. assessment (from others)	12. product name (or slogan)	13. limit of quantity	14. others ……	

2.2 Phase 2: (Pilot study) Examinations of Correspondence Between the Simulated Interface Designs and Metaphorical Concepts

Four volunteers were recruited to examine whether the simulated graphics and interactions corresponding to the metaphors intended, 3 of whom were graduated students and 1 of whom was worker. There were 2 female and 2 male volunteers ranging in age from 24 to 31. On average, they have had 4.25 personal mobile phone products, ranging from 3 to 6. To verify the correspondence between the visualization of 3 types of interface design and original metaphor concepts, open-ended interviews were applied in this phase. This aimed to estimate how much time volunteers could correctly guess the objects of metaphors. They were also asked to interpret the helpful cues to guess. Here, volunteers were allowed only to see the static graphics of the metaphorical interface design in response to all the questions.

2.3 Phase 3: The Usability Evaluations of 3 Types of Metaphorical Interface Design

Eight subjects were recruited to measure the usability of 3 types of metaphorical interface design with SUS questionnaire, 6 of whom were students and 2 of whom were workers. There were 5 female and 3 male subjects ranging in age from 15 to 27. On average, they have had 3.38 personal mobile phone products, ranging from 1 to 7. In this phase, without restrictions on time, subjects were permitted to manipulate the constructed interfaces by order and instructed to perform 2 tasks in each interface respectively. After operating and performing the tasks on an interface, they were requested to evaluate the usability of three different metaphorical interface designs with SUS questionnaire. Then, the SUS scores were calculated to indicate whether it was a good metaphorical interface design.

2.4 Phase 4: User's Preference for the Metaphorical Interface Design

Twelve subjects were enrolled in this phase, 9 of whom were students and 3 of whom were workers. There were 7 female and 5 male subjects ranging in age from 15 to 31. On average, they have had 3.67 personal mobile phone products, ranging from 1 to 7. Here, after operating and performing the tasks on each metaphorical interface design, subjects were requested to evaluate personal impressions on the 3 types of metaphorical interface design with SD questionnaire and compare individual's preference for any one of them.

3 Results and Discussion

The results of different experimental analysis in each phase were described respectively as follows:

3.1 Phase 1: User's Response to Mobile Phone Products

To interpret the results more systematically and easily, the observations were analyzed through the frequency of descriptive statistics. Then, a score (a cumulative number, abbreviated to "sum"; the highest score = 219 = 73 x 3) was obtained by adding across the answers (frequencies) in the same question but different orders.

In response to question 1 (to evaluate whether a product of mobile phone is good or bad), "appearance" was considered as the most important factor (sum = 54), next was "functionality" (sum = 39), next was "quality" (sum = 27), next was "interface manipulation" (sum = 25), and next were "brand" and "price" (the same as sum = 24). Effects on other factors were ignored because their sum total was less than 20.

In response to question 2 (to find out the attractive essentials of a mobile phone product), "appearance" was obviously regarded as the most important factor (sum = 65), next was "functionality" (sum = 27), and next was "price" (sum = 22). Effects on other factors were ignored because their sum total was less than 20.

In response to question 3 (to clarify which factors really determine a user's decision to buy), "appearance" was still regarded as the most important factor (sum = 57), next was "price" (sum = 48), and next was "functionality" (sum = 36), next were "brand" (sum = 21), and next was "interface manipulation" (sum = 20). Effects on other factors were ignored because their sum total was less than 20.

Furthermore, the average rating of questionnaire items on a 9 point scale ranging from "strongly agree to abstract" to "strongly agree to concrete" reflected the degree of abstract in metaphorical interface designs. Consequently, that the metaphorical interface designs scored 3.5, 0.75, and -2.25 in turn meant metaphor 3 were regarded as abstract interface design and metaphor 1 and metaphor 2 were regarded as concrete interface designs, but the degree of concrete in metaphor 1 was stronger than in metaphor 2.

3.2 Phase 2: (Pilot study) Examinations of Correspondence Between the Simulated Interface Designs and Metaphorical Concepts

There were no significant differences in making a guess at the metaphor of visualization and in performing completion times found among different metaphorical

interface designs. However, large differences in surprise and interest ratings were evidently resulted from volunteer's personal attitude towards mobile phone products. Therefore, further results and discussions might be appropriately illustrated with the outcomes in next phases.

3.3 Phase 3: The Usability Evaluations of 3 Types of Metaphorical Interface Design

In order to appraise whether it was good metaphor at interface design, the usability evaluations of 3 metaphorical interface designs were performed by means of SUS questionnaire. Then, SUS scores were calculated and illustrated with Fig 5 below.

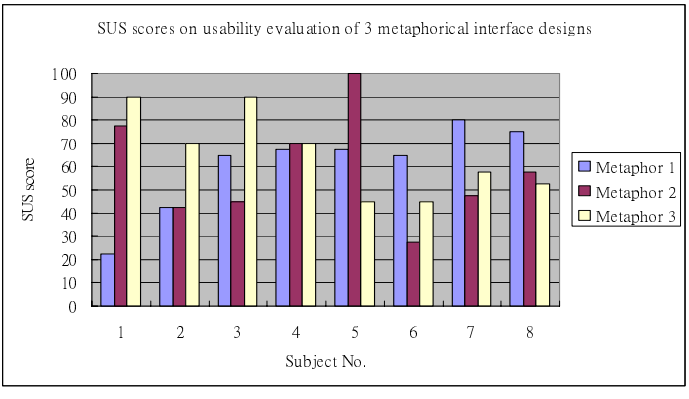

Fig. 5. SUS scores on usability evaluation of 3 metaphorical interface designs

As shown in Fig 5, it was indicated that there were large dissimilarity of cognitions between different participants (especially in the results of metaphor 2). That meant what was someone's favor might be disliked by another. Fortunately, the interview method was simultaneously combined in this phase. The discrepancies in the results of the scores were easily ascribed to subject's attitude towards mobile phone products. Despite some of them treat the products as only communicated tools, for example, another ones regarded them as representations of personality. Besides, the SUS scores were deeply affected by personal preferences, too. So, it was certainly needed to compare the metaphorical interface designs based on user's subjective responses.

3.4 Phase 4: User's Preference for the Metaphorical Interface Design

In this phase, subjects were requested to make a favorite choice between coupled selections of metaphorical interface design. Consequently, in a couple of metaphor 1 and metaphor 2, 8 subjects chosen metaphor 1 as their preference; in a couple of metaphor 2 and metaphor 3, 7subjects chosen metaphor 2 as their preference; in a couple of metaphor 3 and metaphor 1, 10 subjects chosen metaphor 1 as their preference. To conclude, metaphor 1 was more popular than other.

In addition, there might be a good idea to analyze and gain further useful information on the results of the coupled comparison by means of quantity analysis methodologies, such as through MDS analysis to indicate the potential mental space of human beings. Within time limitation in this study, only the result of metaphor 1 was taken for example and shown in Fig 6 below.

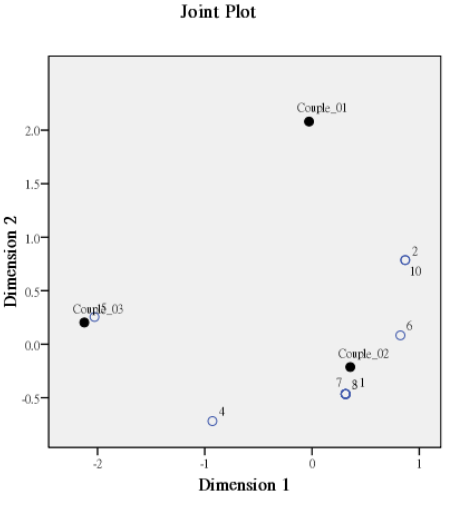

Fig. 6. The result of MDS analysis of metaphor 1

4 Conclusion

Although, the System Usability Scale (SUS) is a commonly used, freely distributed, and reliable questionnaire consisting of 10 items, it was also verified that some of subjects (non-native English speakers) failed to understand the word "cumbersome" in Item 8 of the SUS ("I found the system to be very cumbersome to use") without instructor's further explanation (Finstad, 2006). Also, in this study, it was found out that more of subjects might be not used to seek for further help to any questions spontaneously but appreciate if the instructor would voluntarily give them more information. So, it was suggested that the descriptions of the questionnaires should be confirmed with no doubt before being proceeded. Moreover, it was also noticed that subject's attitude towards mobile phone products brought large differences in the results of both usability evaluation and subjective impressions appraisal. Because of SUS estimations and interviews being substitute for quantity measurement in this study, only 4-12 subjects were recruited to perform the assessment tasks. In consequence, it was difficultly aimed at generalize any principle of metaphorical interface design in common. So, that might be a good point to improve this study in further works.

References

1. Benyon, D., Palanque, P.: Critical Issues in User Interface Systems Engineering. Springer, New York (1996)
2. Brooke, J.: SUS: A Quick and Dirty Usability Scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (eds.) Usability Evaluation in Industry, Taylor & Francis, London (1996) Also see http://www.usabilitynet.org/trump/documents/Suschapt.doc
3. Finstad, K.: The System Usability Scale and Non-Native English Speakers. Journal of usability studies 1(4), 185–188 (2006)
4. Imaz, M., Benyon, D.: How Stories Capture Interactions. In: Proceedings of Human-Computer Interaction – INTERACT '99, Edinburgh, Scotland, September 1-3, 1999, pp. 321–328. IOS Press, Amsterdam (1999)
5. Kissel, G.V.: The effect of computer experience on subjective and objective software usability measures. In: Conference companion on Human factors in computing systems 1995 (1995)
6. Nielsen, J.: Usability Engineering. Academic Press, Boston (1993)
7. Nielsen, J.: Why You Only Need to Test With 5 Users. Alertbox, (March 19, 2000) (accessed on 06 February 2007) http://www.useit.com/alertbox/20000319.html
8. Norman, D.A.: The design of everyday things. Doubleday, New York (1990)
9. Norman, D.A.: Emotion & design: attractive things work better. Interactions 9(4), 36–42 (2002)
10. Shneiderman, B., plaisant, C.: Designing the user interface: strategies for effective human-computer interaction, 4th edn. Addison Wesley, London (2005)
11. Vatrapu, R., Perez-Quinones, M.A.: Culture and Usability Evaluation: The Effects of Culture in Structured Interviews. Journal of usability studies 1(4), 156–170 (2006)