A Usability Checklist for the Heuristic Evaluation of Mobile Phone User Interface

Jun Ho Park[1], Youngbo Suh[2], Cheol Lee[3], Yong Gu Ji[4], Myung Hwan Yun[5]
Dept. of Industrial Engineering, Seoul National University[1 2 3 5]
Dept. of Information & Industrial Engineering, Yonsei University[4]
{juncrom0[1], suh0816[2], iehis[3], mhy[5]}@snu.ac.kr, yongguji@yonsei.ac.kr[4]

Abstract

In the last decade, the research of the usability of mobile phones has been a newly evolving area with few established methodologies and realistic practices that ensure capturing usability in evaluation. Thus, there exists growing demand to explore appropriate evaluation methodologies that evaluate the usability of mobile phones quickly as well as comprehensively. This study aims to develop a task-based usability checklist based on heuristic evaluations in views of mobile phone UI practitioners. A hierarchical structure of UI design elements and usability principles related to mobile phones were developed and then utilized to develop the checklist. In order to demonstrate the practical effectiveness of the proposed checklist, comparative experiments were conducted on the usability checklist and usability testing. Majority of usability problems found by usability testing and additional problems were discovered by the proposed checklist. It is expected that the usability checklist proposed in this study could be used quickly and efficiently by usability practitioners to evaluate the mobile phone UI in the middle of the mobile phone development process.

Keyword : Usability, User interface, Checklist, Mobile phone, Heuristic evaluation

1. Introduction

In the last decade, the mobile phone has rapidly evolved from a simple device to make or receive phone calls to a complex multimedia interactive system. Due to the rapid development of mobile technologies and memory devices, in recent years, a mobile phone has expanded its functionality to include an mp3 player, digital camera, and digital multimedia broadcasting (DMB) system. Consequently, the mobile phone has been regarded as a common consumer appliance (Klockar et al., 2003).

The research of usability in mobile phones is a newly evolving area. Mobile phones have only been introduced to general consumers since 1990s and, thus, the related research community has had a relatively short history (Kjeldskov & Graham, 2003). Therefore, there is very little knowledge such as established methodologies and realistic practices that capture usability in evaluating mobile phones (Klockar et al., 2003). Result of a recent literature survey (Beck et al., 2003) indicated that 44 out of 114 papers, which were extracted from recent publications in the area of mobile HCI, utilized traditional usability evaluation techniques such as heuristic evaluation and that only 6 of these papers employ somehow new methods in realistic mobile use situation. However, most of the studies fall short of evaluating comprehensive aspects of usability issues on mobile phones. For instance, none of these studies considered the newly evolving mobile technologies and added functions, such as mp3 players and digital cameras. For this reason, further study is still needed to explore appropriate user interface (UI) evaluation techniques to evaluate the usability of mobile phones quickly yet comprehensively in order to enhance user satisfaction.

The purpose of this study is to develop a usability checklist that enables mobile phone UI practitioners to evaluate software UI design using the enhanced traditional UI evaluation approach and to explore potential usability problems comprehensively from a holistic perspective. To verify the effectiveness of the proposed usability checklist, a case study with advanced mobile phones has been conducted. In addition,

conventional usability testing has been conducted and its result has been compared with that of the proposed usability checklist.

2. Backgrounds

According to the literature that reviewed mobile HCI research methods (Scholtz, 2004; Kjeldskov & Graham, 2003), various usability evaluation methods (UEMs) are being developed to assess and improve usability of interactive systems. Interactive systems are usually designed through an iterative process involving design, evaluation, and redesign as shown in Figure 1 (Kies et al., 1998).

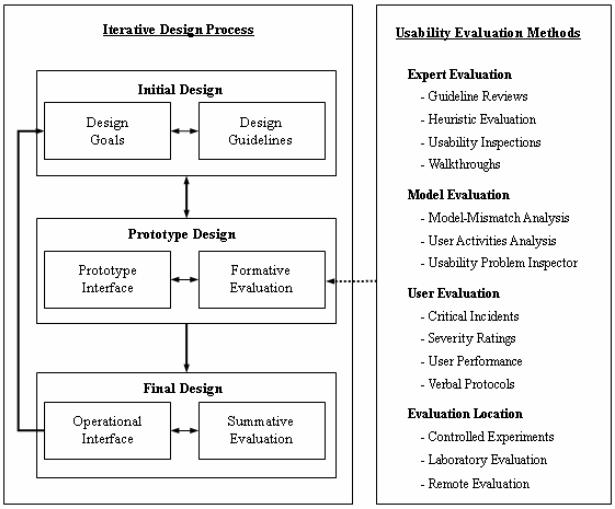

Figure 1. Usability evaluation methods in formative evaluation (Adapted from Kies et al., 1998)

During the initial design phase, goals and guidelines are iterated to finalize the design specifications that ultimately lead to a prototype design. Formative evaluation focuses on usability problems that need to be solved during the prototype design stage, before a final design can be accepted for release. Summative evaluation is then conducted to evaluate the efficacy of the final design or to compare competing design alternatives in terms of usability. Usability evaluation methods are used primarily for formative evaluations during the prototype design stage. These formative evaluations are focused on efficient and effective techniques to determine usability problems that need to be eliminated through redesign (Kies et al., 1998; Scholtz, 2004).

According to the literature on current practices of mobile evaluations, there is no consensus as to whether the usability of mobile devices should be assessed by heuristic evaluation or usability testing since both of these methods have their own strengths and limitations (Law & Hvannberg, 2002).

Heuristic evaluation is an established usability evaluation technique originally proposed by Nielsen (1994) as a cost-effective usability technique. It is an inspection method in which a panel of experts formally assesses an interface design with respect to a set of heuristics or rules of thumb (Nielson, 1994). The main strength of heuristic evaluation is its speed and affordability (Jeffries & Desurvire, 1992). It also provides conciseness, memorability, meaningfulness, and insight (Paddison & Englefield, 2003). Nielsen also suggested that the best result of heuristic evaluation can be gained if evaluators are double experts in usability engineering and in the domain of interest (Law & Hvannberg, 2004).

The nature of mobile phone development inherently requires a fast and inexpensive design process. Since the development phase of mobile phones is fairly short, heuristic evaluation can be performed without real end-users or a working prototype in the early stage of design process. In addition, if heuristics such as checklist are documented appropriately, they are easy to learn and thus can be used by non-usability experts (Nielsen, 1994). Hence, heuristic evaluation methods are recommended to evaluate software UI of mobile phones.

However, there are some limitations of heuristic evaluation. A limitation is that heuristic evaluation does not approximate the conditions under which real users would use the mobile system. In addition, heuristic evaluation provides little information about the magnitude of the usability problems that are detected. (Simeral & Branaghan, 1997). Moreover, the skills and experience of the usability experts can significantly influence the quality of heuristic evaluation. For this reason, heuristic evaluation should supplement usability

testing rather than replacing it (Law & Hvannberg, 2002). Subsequently, the focus of this study is to identify an opportunity for extending and improving the traditional heuristic evaluation technique.

3. Development of a Usability Checklist

The procedure to develop a usability checklist for the mobile phone UI of this study is illustrated in Figure 2.

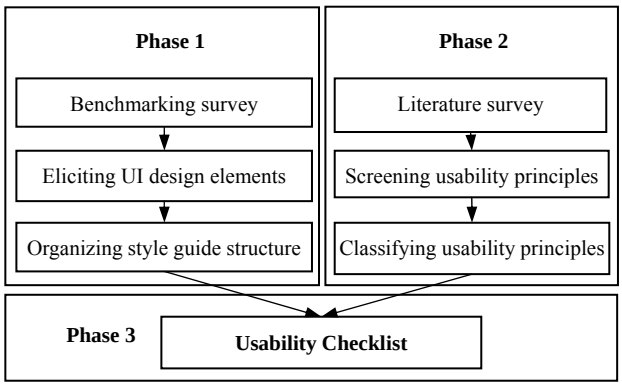

Figure 2. Development procedure of usability checklist

The above procedure can be divided into three phases. In Phase 1, previous research materials about the 'style guide', which was developed by several mobile phone companies, are collected and analyzed to obtain 'UI elements' that compose the mobile phone UI. Once UI elements are elicited, a hierarchical structure containing those elements is organized. This hierarchy forms the basis of the evaluation checklist, which will help to improve the effectiveness of heuristic evaluation. During Phase 2, a literature survey about usability issues is conducted. Survey results are arranged and are regarded as important usability principles that must be considered in the software UI design process of mobile phone. Arranged principles are carefully selected, deleted, and integrated into a structure ensuring classification of usability principles. This manipulation is surely performed under specific criteria. In Phase 3, through the pair-wise comparison, UI elements from Phase 1 and usability principles from Phase 2 are matched up. Then, for one UI element, questionnaire sentences relevant to matched usability principles are developed. The questionnaire was designed to be comprehensible and practical for usability experts as well as developers.

3-1. Style Guide Structure of Mobile Phone UI

To elicit UI elements related to the mobile phone UI, many style guides from mobile phone companies such as Nokia, Verizon, Sprint, and SK Telecom were collected and analyzed. The analyzed data was then classified by the keywords appearing in the style guide documents. The number of total elicited elements was 86. The style guide structure was specified through the element classifications and groupings of the relationship analysis. The elements were compared one by one, and the relationship inherent in each element pair was assessed by a focus group consisting of usability experts. Based on the assessed relationship, those elements were grouped, and each group was named to indicate its own representative characteristic; (1) UI Policies, (2) UI Screens, (3) UI interactions, and (4) UI components. As a result, a hierarchical structure of UI elements was developed as shown in Table 1.

Table 1. Hierarchical structure of style guide

1st Level	2nd Level	3rd Level
UI Policies	Menu Policies	Main menu, Menu grouping, Menu labeling
	Navigation Policies	Main menu navigation, List navigation, Photograph/Video file navigation
	Soft Key Policies	Soft key arrangement, Soft key allocation
	Pop-up Policies	Notification pop-up. Caution pop-up, Selection pop-up, Input pop-up
	Icon Policies	Static icon, Dynamic icon
UI Screens	Menu Screens	Idle screen, Main menu screen, 2nd depth menu screen, List menu screen, Check box menu screen, Radio button menu screen
	Status Screens	Preview screen, Animation screen, Multi-setup screen
	Function-based Screens	Calling screen, Search screen, MP3/Video playing screen, Photographing screen, Multimedia contents management screen
	External Screens	Idle screen, Message screen, Alarm/Morning call/Schedule screen, Photographing screen, MP3 playing screen
UI Interactions	Interaction Task Type	Confirmation, Input, Termination, Backward/Cancel, Search
	Interaction Feedback Type	Send, Task confirmation, Move, Lock, Save, Modify, Delete, Download, Initiate, Load, Connect
UI Components	Body Area	Title area, Subtitle area, Contents area
	List Types	Menu list, Single selection list, Multi-selection list, Markable list, View state form
	Widget	Roll-up box, Scroll bar, Radio button, Check box, Spin, Progress bar, Slider
	Text Field	Multi line text field, Single line text field
	Indicator	Status indicator, Stage indicator, Proceeding indicator
	Tab	Tab

3-2. Usability Principles of Mobile Phone UI

In this phase, as stated above, important usability principles of the mobile phone design were arranged according to the result of literature survey on usability issues. Consequently, a total of 43 usability principles were collected and arranged as Table 2.

Collected principles were examined by the following three criteria; a) Selection: Is there a practical impact on performance?, b) Integration: Is there any redundancy or similarity relevant to other principles?, and c) Deletion: Are users' subjective feelings include? (or have the mobile phone UI considerations not been considered?). Examined principles were also screened in terms of degree of comprehension, causality or correlation, and hierarchical relationships.

Table 2. Collected usability principles

Reference	Usability Principle
Constantine (1994)	Structure principle, Simplicity principle, Visibility principle, Feedback principle, Tolerance principle, Reuse principle
Nielson (1994)	Visibility of system status, Match between system and the real world, User control and freedom, Help users recognize, diagnose, and recover from errors, Recognition rather than recall, Aesthetic and minimalist design, Pleasurable and respectful interaction with the user, Consistency and standard
Treu (1994)	Effort
Dix et al. (1998)	Learnability, Predictability, Synthesizability, Familiarity, Generalizability, Consistency, Flexibility, Dialog initiative, Multi-threading, Customizability, Task migratability, Subsitutivity, Robustness, Observability, Recoverability, Responsiveness, Task conformance
Lauessen & Younessi (1998)	Ease of learning, Task Efficiency, Ease of remembering, Understandability, Subjective Satisfaction
Preece et al. (2002)	Effectiveness, Efficiency, Safety, Utility, Learnability, Memorability

Through the above screening process, 21 usability principles were selected and by comparing each principle one by one, relationships within all pairs of principles were specified as follows; '2' indicated a strong relationship, '1' indicated a moderate relationship, while '0' was used to indicate no relevant relationship. A total of 10 usability experts took part in this analysis. A table was constructed with the relationship indicating numbers (0, 1, 2) and was employed as an input dataset for Principal Component Analysis (PCA) with varimax rotation. The result of PCA is shown as Table 3.

Table 3. The result of PCA with Varimax rotation

Principle	Factor				
	1	2	3	4	5
Predictability	0.6329				
Learnability	0.9235				
Structure Principle	0.8875				
Consistency	0.9049				
Memorability	0.8497				
Familiarity	0.7475				
Recognition		0.6180			
Visibility		0.9041			
Simplicity		0.7014			
Subsitutivity		0.7838			
Feedback			0.8011		
Error Indication			0.4853		
Synthesizability			0.7405		
Responsiveness			0.4161		
Recoverability				0.8068	
Flexibility				0.2816*	
User Control				0.0390*	
Customizability				0.1650*	
Effectiveness					0.7006
Efficiency					0.6364
Effort					0.5474
Cronbach's alpha	0.9261	0.8791	0.8298	0.8141	0.8834

Note: * remained principles after adjustment

Each factor includes usability principles which have factor loadings greater than 0.4. However, flexibility, user control, and customizability (factor loadings < 0.4) were still included in the factor due to their importance in supporting mobile phone users.

From the result of the PCA, five groups were classified and functionally defined as follows: (1) Cognition support: relates to cognitive aspects of users, (2) Information support: relates to characteristics of mobile phone display and information, (3) Interaction support: relates to the interaction between user and mobile phone, (4) Performance support: relates to performance of the intended task of the user-mobile phone system, and (5) User support: relates to the degree of intervention of user. Table 4 shows the definition of each principle and grouped structure.

Table 4. The definitions and structure of usability principles

Principle		Definition
Cognition Support	Predictability	The User Interface must produce results that are in accord with previous commands and states.
	Learnability	The User Interface must be designed for user to learn easily the use of Mobile Phone.
	Structure Principle	The User Interface must be organized purposefully, in meaningful and useful ways that put related things together and separate unrelated things based on clear, consistent models that are apparent and recognizable to others.
	Consistency	The User Interface must be designed consistently.
	Memorability	The User Interface must be easy for users to remember how to use the mobile phone.
	Familiarity	The User Interface must be familiar to users.
Information Support	Recognition	The User Interface must be easy for users to recognize the status of systems or the use of Mobile Phone.
	Visibility	The User Interface should always keep users informed about what is going on, through appropriate feedback within reasonable time.

	Simplicity	Make simple, common tasks simple to do, communicate simply in the user's own language and provide good shortcuts that are meaningfully related to longer procedures.
	Subsitutivity	The information about numerical values must be easily understood by users.
Interaction Support	Feedback	The User Interface must keep users informed of actions or interpretations, changes of state or condition using clear, concise, and unambiguous language familiar to users.
	Error Indication	The representation of errors must be clear to users.
	Synthesizability	The user must be able to construct the proper model of the system. The system must display the correct clues to construct a proper model.
	Responsiveness	The system must respond in an appropriate time.
Performance Support	Effectiveness	The required range of tasks must be accomplished at better than some required level of performance by some required percentage of the specified target range of users within some required proportion of the range of usage environment.
	Efficiency	The system should be efficient to use so that once the user has learned the system, a high level of productivity is possible.
	Effort	The User Interface should be designed to minimize the user's effort for using the system.
User Support	Recoverability	If the user makes a mistake or the application fails, the user must be able to recover the work.
	Flexibility	The User Interface must be flexible so that adapts to various environments and users.
	User Control	The users must be able to control the system by their own decisions.
	Customizability	The user must be able to modify the interface in order to improve efficiency. The customizing features must be easily accessible.

3-3. Usability Checklist of Mobile Phone UI

Style guide structure and the arranged usability principles were utilized to develop a checklist. This checklist was intended to support the assessments of design alternative in terms of usability. Through the pair-wise comparison, each style guide structure component was matched with relevant usability principles (see Figure 3). For each component, the usability principle which must be met in the design of the mobile phone UI was elicited by usability experts' discussion.

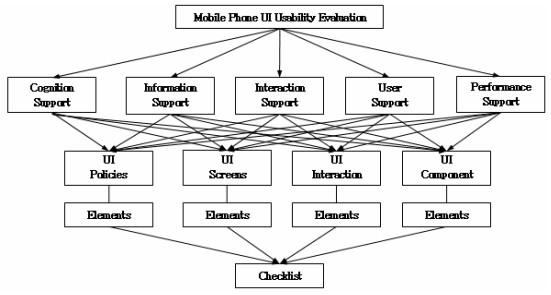

Figure 3. Conceptual structure of usability checklist

Based on the relevant principles, evaluation questionnaires for each element of style guide structure were derived from the focus group's unanimous decision which consisted of an equal number of usability experts and mobile phone developers. With the checklist, evaluators will be asked to rate their degree of satisfaction for each evaluation item using 7-point likert-type rating scale and also to describe suggestions if necessary. A part of the proposed evaluation checklist is exemplified in the Appendix.

4. Case Study

4-1. Usability Evaluation of Mobile Phone UI

To measure and demonstrate the practical effectiveness of the proposed usability checklist, comparative experiments on both usability checklist and usability testing (UT) were conducted, respectively, with simulated task scenarios in laboratory.

4-1-1. Mobile phones for usability evaluation

Samples of mobile phones for usability evaluation of the mobile phone UI were chosen considering state-of-the-art functions such as an mp3 player, QVGA displays (resolution: 320*240), and digital camera. As shown and described in Table 5, three mobile phones in the Korean telecommunication market were selected to be utilized in the experiment.

Table 5. Mobile phones for evaluation

Specifications	Sample 1	Sample 2	Sample 3
Manufacturer	LGT	Samsung	Ever
Model No.	HS8000	SPH-V6050	KTF-X6000
Released date	05/2005	07/2005	12/2004
LCD resolution	320*240	320*240	320*240
MP3 function	support	support	support
Camera	2 mega pixels CCD	1.3 mega pixels CCD	2 mega pixels CCD

4-1-2. Task-based usability checklist evaluation

The three mobile phone samples were evaluated using the proposed usability checklist developed in session 3.3. Ten graduate students (5 male and 5 female) having at least 2 years mobile phone experience were recruited from Seoul National University. In addition, their experiences on using mobile phones with mp3 player, QVGA displays, and camera were at least 1 year. Their average age was 24.5. At the beginning of the experiment,

each participant was provided with an informed consent form and a brief description of the goal and procedure of the experiment. Between questionnaires for each sample phone, participants were given a five-minute resting break. The mobile phones were evaluated in random order while the questionnaire for evaluation was identical across all three sample mobile phones.

4-1-3. Usability testing evaluation

Usability testing refers to a process that employs participants who are representative of the target population to evaluate the degree to which a product meets specific usability criteria. Among various handbooks of usability testing, Rubin's handbook (Rubin, 1994) was used to conduct usability testing in this study.

In this study, usability testing was performed as a laboratory experiment. A total of ten participants (5 male and 5 female) having at least 2 years mobile phone experience were recruited from student population of Seoul National University. Their average age was 23. The same mobile phones were employed in this usability experiment while different participants were employed to prevent the learning effects from the experience of checklist evaluation. UT scripts, which were utilized for usability testing, consisted of questionnaires for personal backgrounds, instructions of the experiment, task scenarios, and subjective satisfaction questionnaires.

During the experiment, every task of each participant for three sample mobile phones was recorded using a digital camcorder (DCR-PC105, SONY). The recordings were to be further analyzed in terms of error frequencies, task completion time, and task success rate. After finishing each task of the 19 task scenarios for each sample mobile phone, an interview was held to note remarks on errors or problems related to the task. Between experimentation for each sample phone, participants were given a five-minute resting break. The sequence of 19 task scenarios for usability testing was randomized for each sample phone while the task itself was identical across three sample mobile phones.

Based on QUIS (Questionnaire for User Interface Satisfaction; Chin et al., 1988), questionnaires for evaluating subjective satisfaction were given upon completing the entire task scenarios for each sample phone. QUIS was designed to measure users' subjective satisfaction about the product interface. It was divided into one section for overall satisfaction and sections for satisfaction for four specific interfaces. Participants were asked to rate their degree of satisfaction for each evaluation item using 7-point likert-type rating scale. As exemplified in Table 6, a task scenario consisted of frequently used function. Overall 19 task scenarios were classified into 8 top-levels and 18 sub-levels.

Table 6. Example of task scenarios for usability testing

Top-level	Sub-level	Description of Task Scenarios
Phone Book	Phone number registration	1. Starting at the idle screen, register following phone number.(Phone number: 012-3456-7890, Name: ㄱㄱㄱ, Address: 20, Category: mobile phone, Group: family)
		2. Starting at the idle screen, Enter the phone book menu by menu navigation and then register following phone number.(Phone number: 111-1111-1111, Name: ㄴㄴㄴ, Address: 30, Category: mobile phone, Group: family)
	Phone number searching & modification	3. Starting at the idle screen, press the short key and search the number of 'ㄱㄱㄱ'. Then, modify the number to '098-7654-3210'.
		4. Starting at the idle screen, enter the phone book menu by menu navigation and search the number of 'ㄱㄱㄱ'. Then, modify the number to '012-3456-7890'.

4-2. Results of Case Study

The results of usability evaluation for sample mobile phones using the checklist and usability testing were summarized in Table 7.

According to the result of usability checklist evaluation, the overall rating on sample 1 was the highest (5.14), sample 2 ranked the second (5.07), and sample 3 received the lowest score (4.49). With the exception of UI interactions, users' ratings on other parts were similar with the overall score while the mean rating of sample 2 was the highest among the three phones in terms of UI interactions. The result of the usability testing was similar to that of the checklist. The overall satisfaction on sample 1 was the highest (4.93), sample 2 ranked the second (4.71), and sample 3 ranked the lowest (3.15). Also, the mean score of all dimensions showed similar tendency.

Table 7. Summary of results

Type	Dimension	Mean of Rating Score (SD)		
		Sample 1	Sample 2	Sample 3
Usability Checklist	1. UI Policies**	5.15 (0.655)	5.04 (0.422)	4.36 (0.514)
	2. UI Screens*	5.17 (0.482)	5.12 (0.424)	4.59 (0.351)
	3. UI Interactions*	4.88 (0.618)	4.90 (0.585)	4.15 (0.541)
	4. UI Components	5.36 (0.671)	5.20 (0.558)	4.85 (0.590)
	Total*	5.14 (0.573)	5.07 (0.389)	4.49 (0.439)
Usability Testing	1. Overall reaction***	4.86 (0.618)	4.68 (0.271)	2.66 (0.162)
	2. Screen**	4.88 (0.311)	4.51 (0.182)	3.03 (0.109)
	3. Terminology & system information**	4.88 (0.358)	4.78 (0.163)	3.42 (0.515)
	4. Learning**	5.08 (0.259)	4.83 (0.192)	3.50 (0.308)
	5. System Capabilities**	4.98 (0.349)	4.73 (0.238)	3.45 (0.456)
	Total***	4.93 (0.081)	4.71 (0.170)	3.15 (0.323)

Note: significant at P<0.05(*), P<0.01(**), and P<0.001(***)

To investigate the difference in satisfaction among the sample phones, a series of ANOVAs were conducted. Regarding the usability checklist, the results of ANOVA showed that the effect of sample phones on overall satisfaction was significantly different ($F_{2, 27}$ = 4.99, P<0.05). Satisfaction in UI policies ($F_{2, 27}$= 5.55, P<0.01), UI screens ($F_{2, 27}$= 5.15, P<0.05) and UI interactions ($F_{2, 27}$= 4.74, P<0.05) were also significantly different. But the satisfaction of UI components were not significantly different ($F_{2, 27}$= 1.63, P=0.216) since all sample phones indicate the UI components appropriately.

In the case of usability testing, the result of ANOVA indicated that overall satisfaction across sample phones was also significantly different ($F_{2, 27}$= 12.41, P<0.001). In addition, ratings of overall reaction ($F_{2, 27}$ = 10.17, P<0.001), screen ($F_{2, 27}$= 7.04, P<0.01), terminology and system information ($F_{2, 27}$= 8.00, P<0.01), learning ($F_{2, 27}$ = 7.97, P<0.01), and system capabilities ($F_{2, 27}$ = 7.16, P<0.01) were significant different, respectively.

Post-hoc ANOVA analysis was conducted to further investigate the difference of mean scores between sample phones using Tukey's test. The results of Tukey's test revealed that there were significant differences of overall scores from the usability checklist between sample 1 and sample 3 (P<0.05) as well as between sample 2 and sample 3 (P<0.05). However, mean scores on sample 1 and sample 2 did not show a significant difference (P= 0.947). Regarding the result of usability testing, there were also significant differences of overall mean scores between sample 1 and sample 3 (P<0.001) and sample 2 and sample 3 (P<0.01). The mean scores on sample 1 and sample 2 did not show a significant difference (P=0.236).

According to the evaluation results from both usability checklist and usability testing, a greater number of problems were found through the checklist evaluation compared to the usability test. In usability test, usability problems were identified by the analysis of user observation and user interviews. As a result, a greater number of problems were found through the checklist evaluation compared to the usability test. Furthermore, usability problems that identified frequently in the usability testing were also co-discovered by the checklist. The checklist evaluation recognized 28 additional problems for sample 1, 10 for sample 2, and 27 for sample 3 as shown in Table 8.

Table 8. Comparison of checklist and usability testing

Phone	Sample 1		Sample 2		Sample 3	
Method	Checklist	UT	Checklist	UT	Checklist	UT
Number of problems	76 (28)[1]	54 (6)[2]	59 (10)[1]	54 (5)[2]	92 (27)[1]	73 (8)[2]

Note: [1]Number of additional usability problems found by checklist
[2]Number of usability problems not found by checklist

These problems were closely related to practical usage of mobile phones. In the case of sample 2, participants hardly perceived the speed of fast forward when using the mp3 play function, but this fact was not discovered when performing a checklist evaluation. However, problems on icon for caution pop-up or consistency about soft-key labels were discovered through the checklist evaluation, although the usability testing failed to recognize such problems.

In the result of experiments, the characteristics of usability testing and checklist evaluation could be compared and contrasted. The usability testing could discover the serious problems of interaction occurring in the mobile phone usage while the proposed checklist evaluation could found various problems of specific UI elements. Compared to the traditional heuristic evaluation, it was revealed that the checklist enabled evaluators to find majority of interaction problems found in the usability testing.

5. Discussion

The usability checklist developed in this study is mainly based on heuristic evaluation methods that are the most popular usability evaluation methods. According to Law & Hvannberg (2004), the effectiveness of heuristic evaluation closely depends on the importance of selecting usability guidelines. Accordingly, we have developed 21 usability principles that are crucial in the mobile phone UI design.

Beck et al. (2003) asserted that most of the previous studies on usability evaluation fall short of evaluating comprehensive aspects of usability issues on mobile phones. Due to the limited life cycle and rapid change in mobile technology, a usability checklist for heuristic evaluation should be updated quickly and easily by including the additional UI elements and appropriate usability principles into the evaluation framework. In this study, the mobile phone UI style guide structure was developed to evaluate the comprehensive aspects including state-of-the-art functions such as an mp3 player and digital camera.

In the proposed usability checklist, there exists a promising methodological benefit to enhance the effectiveness and efficiency of heuristic evaluation as follows. The result of comparative experiments on both usability checklist and usability testing revealed that about 90% of usability problems identified by usability testing were covered by the proposed usability checklist. According to the literature (Law & Hvannberg, 2002), heuristic evaluation reports typically do not predict 30% to 50% of usability problems found by usability testing. Given the result in this study, it can be asserted that the proposed checklist evaluation might improve the traditional heuristic evaluation technique. In addition, the checklist developed in this study could supplement usability testing since it could cover the majority of usability problems found in usability testing.

6. Conclusion

This study developed a suitable usability checklist systematically, which enables mobile phone UI developers to quickly and easily evaluate the usability of software UI. The structure of style guide including comprehensive mobile phone UI elements was specified and relevant usability principles were collected and then classified into five meaningful groups. In addition, the usability checklist was developed through the pair-wise comparison of UI elements and usability principles. To improve the practical effectiveness of the developed checklist, various usability principles were utilized to clearly specify evaluation items, which were relevant to the software UI of mobile phones, and items themselves were expressed concisely to facilitate the developer's interpretation.

The results of the case study successfully demonstrated the practical effectiveness of the usability checklist. In addition to the commonly recognized usability problems, the checklist of this study also enabled evaluators to identify the serious usability problems that could be identified by the usability testing. It may suggest a possibility that the proposed checklist evaluation could supplement the usability testing. The scores of subjective satisfaction from the checklist and usability testing showed similar tendencies and it also justified the effectiveness of the proposed checklist.

It is expected that the usability checklist developed in this study can be utilized by usability practitioners to evaluate the mobile phone UI concurrently during the process of development because it is intended to evaluate each UI element without time pressure and excessive experimental efforts. Finally, it should be mentioned that further continuous efforts are needed in order to update the checklist as new mobile technologies emerge.

Reference

Beck, E. T., Christiansen, M. K., Kjeldskov, J., Kolbe, N., & Stage, J. (2003). Experimental Evaluation of Techniques for Usability Testing of Mobile Systems in a Laboratory Setting. *Proceedings of OZCHI 2003 Conference*.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an Instrument Measuring User Satisfaction of The Human-Computer Interface. *Proceedings of the SIGCHI conference on Human*

factors in computing systems (pp. 213–218). Washington, D.C.: ACM Press.

Constantine, L.L. (1994). Collaborative Usability Inspections for Software. *Proceedings of the Software Development'94*. San Francisco: Miller Freeman.

Dix, A., Finlay, J., Abowd, G., & Beale, R. (1998). *Human-Computer Interaction*, Upper Saddle River, NJ, USA: Prentice Hall.

Jeffries, R., & Desurvire, H. (1992). Usability Testing vs. Heuristic Evaluation: Was there a contest? *SIGCHI Bulletin* (pp.39–41). New York, NY: ACM Press.

Kies, J. K., Williges, R. C., & Rosson, M. B. (1998). Coordinating computer-supported cooperative work: A review of research issues and strategies. *Journal of the American Society for Information Science*, *49(9)*, 776–779.

Kjeldskov, J., & Graham, C. (2003). A Review of Mobile HCI Research Methods. *Proceedings of the Mobile Human-Computer Interaction* (pp. 317–335). Springer-Verlag..

Klockar, T., Carr D. A., Hedman, A., Johansson, T., & Bengtsson, F. (2003). Usability of mobile phones. *Proceedings of the 19th International Symposium on Human Factors in Telecommunication*.

Lauessen, S., & Younessi H. (1998). Six Styles for Usability Requirements. In Dubois, P. & Opdahl (Eds.), *Proceedings of 4th Intl. Workshop on Requirements Engineering: Foundations of Software Quality* (pp. 1–12).

Law, E.L., & Hvannberg, E. T. (2002). Complementarity and Convergence of Heuristic Evaluation and Usability Test: A Case Study of UNIVERSAL Brokerage Platform. *Proceedings of the second Nordic conference on Human-computer interaction* (pp. 71–80). New York, NY: ACM Press.

Law, E.L., & Hvannberg, E. T. (2004). Analysis of Strategies for Improving and Estimating the Effectiveness of Heuristic Evaluation. *Proceedings of the third Nordic conference on Human-computer interaction* (pp. 241–250). New York, NY: ACM Press.

Nielsen, J. (1994). Heuristics Evaluation. In J. Nielson & R.L. Mack (Eds.), *Usability Inspection Methods* (pp.25–62). New York: John Wiley and Sons.

Paddison, C., & Englefield, P. (2003). Applying Heuristics to Perform a Rigorous Accessibility Inspection in a Commercial Context. *Proceedings of the 2003 conference on Universal usability* (pp. 126–133). New York: ACM Press.

Preece, J., Rogers, Y., & Sharp, H. (2002) *Interaction Design*. UK: Wiley.

Rubin, J. (1994). *Handbook of Usability Testing*. New York, NY: John Wiley & Sons.

Scholtz, J. (2004). *Usability Evaluation*. Retrieved Oct. 13, 2005, from http://www.itl.nist.gov/iad/ IADpapers /2004/ Usability%20Evaluation_rev1.pdf.

Treu, S. (1994). *User Interface Evaluation: A Structured Approach.* New York, NY: Plenum Press.

Appendix: Usability Checklist (Part 1: UI Policies)

1st Level Element	2nd Level Element	Evaluation Items	Evaluation (Most Negative←Negative←Normal→Positive→Most Positive) ① ② ③ ④ ⑤ ⑥ ⑦	Suggestion
Menu Policies	Main Menu	Is the menu type easy to recognize?		
		Is the main menu visually simple and clear?		
		Can users change the main menu type as they desire?	Yes / No	
	Menu Grouping	Is the main menu composed of related sub-menus?		
	Menu Labeling	Does the menu labeling follow the screen guideline?		
		Is the menu labeling familiar and easy to remember?		
Navigation Policies	Main Menu Navigation	When users press the direction keys to move to other top menus, is the response immediate?		
	List Navigation	Is the method of page movement or scroll consistent in every menu navigation?		
		Is the list selectable with the number key?	Yes / No	
	Photograph/ Video File Navigation	Is the navigation method between folders or files easily predictable?		
Softkey Policies	Soft Key Arrangement and Allocation	Is the soft key which performs the same function always in same location?		
		Is the method of soft key naming always consistent?		
		Is the naming of soft key consistent with actually performed function?		
		Is the confirmation key allocated in the middle?	Yes / No	
		Is the button on the keypad which performs soft key function correctly allocated?	Yes / No	
Icon Policies	Static Icon and Dynamic Icon	Is the icon interpreted equally by users?		
		Is the icon familiar to users?		