

Usability Evaluation of 4-Direction Keys for Ladder Menu Operation

Takeshi Nagami¹(✉), Yoshikazu Seki¹, Hidenori Sakai²,
and Hiroaki Ikeda³

¹ National Institute of Advanced Industrial Science and Technology (AIST),
Tsukuba, Japan

{nagami-takeshi,yoshikazu-seki}@aist.go.jp

² Ricoh Company, Ltd., Tokyo, Japan

hidenori.sakai@nts.ricoh.co.jp

³ Chiba University, Chiba, Japan

iked@faculty.chiba-u.jp

Abstract. In this research, the usability of three typical menu selection operation patterns using 4-direction keys was evaluated. These patterns are often used to set parameters and call functions in information equipment. In this experiment, the independent variables were the key operation methods, menu depth, number of menu items, and user age range. The dependent variables were task execution time and number of key presses. Sixty-two people participated in the experiment. The results show that the effect of menu depth and number of items varies depending on the operation methods. These results will be used to help establish ISO/IEC 17549-21.

Keywords: Menu operation · 4-direction key · Usability · Standardization

1 Introduction

When using information equipment such as mobile phones or digital cameras, a 4-direction key (as shown in Fig. 1) is usually used to operate the menu selections shown in the display area. However, there are no standards established for these operation methods. Because new equipment now includes sophisticated and complex functionalities, operation methods have become complicated. In this situation, the many kinds of operation methods may confuse users. In addition, it is difficult to learn how to operate such equipment. To address these problems, ISO/IEC 17549-2 [6] was published in May 2015.

This paper experimentally evaluates how differences in menu patterns and 4-direction key operations affect usability. This experiment was carried out to develop the international standard described above. The goal was to define operation method guidelines for selection menus that use a 4-direction key.

Menu operations have often been considered a simple form of information equipment interaction. Even in the 1970 s, it was pointed out that the relative merits of menu design also led to operability problems. There has since been much research on menu layout methods and the systematization of menu items [10, 13]. Some research

has investigated the role of menu depth and breadth in information system user interfaces [9, 12]. In addition, factors that affect the movement time and accuracy of menu selection using a mouse have also been investigated [11].

However, as mobile technologies have continued to develop, applications have diversified, and mobile functions have increased. In addition, most people now use mobile devices, and the elderly in particular desire an easy-to-use interface. We examine menu operation using the cross key, which is often implemented in information equipment in today.

Usability is generally represented by terms such as ease of use, user-friendliness, and high learning efficiency. In ISO 9126-1 [7], usability is defined as consisting of Understandability, Learnability, and Operability. ISO 9241-11 [8] 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.” Based on these documents, the usability characteristics of the 4-direction key operation methods are listed in Table 1.

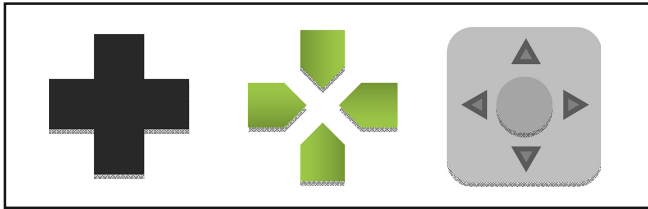


Fig. 1. Examples of 4-direction keys

2 Experiment for Usability of Menu Operation

The Information Technology Research and Standardization Center within the Japanese Standard Association (JSA) performed a field survey of navigation methods of off-the-shelf electronic equipment [5]. The survey included the following products: cellular phones, digital cameras, digital video cameras, music players, personal digital assistants (PDAs), game machines, printers, multiple function copiers, televisions, and projectors. It concluded that navigation methods could be classified into six types, as shown in Fig. 2.

- Type 1: The focus of operation moves up or down endlessly by an up-key or down-key, respectively. It moves through a menu of the hierarchy by a right-key or left-key, respectively.
- Type 2: The operation of a right-key or left-key switches the top menu endlessly. The focus of operation moves up or down endlessly by an up-key or down-key, respectively.
- Type 3: The focus of operation moves up or down and stops at the top or bottom of the menu by an up-key or down-key, respectively. It moves among the hierarchy of ladder menus by a right-key or left-key, respectively.

- Type 4: The focus of operation moves up or down and stops at the top or bottom of the menu by an up-key or down-key, respectively. A right-key or left-key are used set parameters, or to enable or disable the selected feature.
- Type 5: Key operation is the same as Type 3.
- Type 6: Key operation is the same as Type 2.

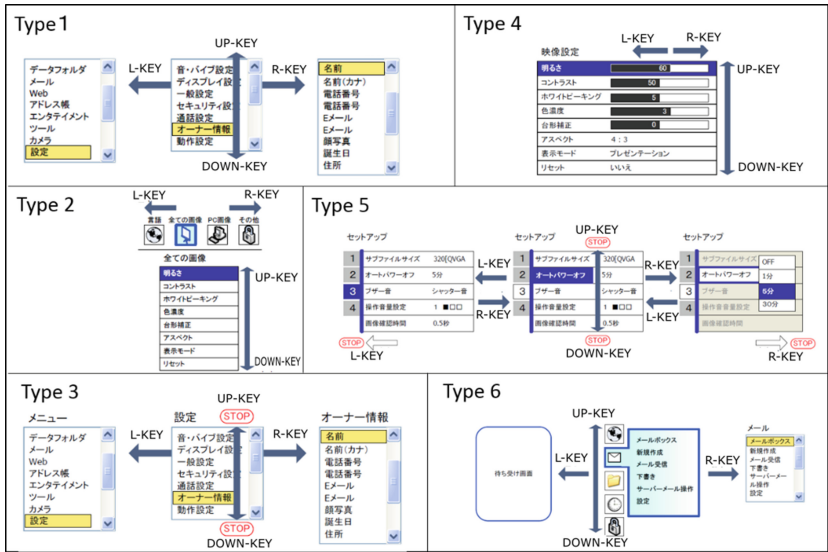


Fig. 2. Overview of navigation methods

In addition, many products were found to use three of the six types.

The purpose of this experiment was to evaluate the usability of the Types 1–3 operation methods described above. The quantitative usability metrics used in this study are listed in Table 1.

Table 1. Characteristics of usability and its metrics

Characteristics	Indices	Metrics
Effectiveness	Clarity of the menu structure	Percentage of correct answers Retry frequency
	and key operations	
Efficiency	Operation speed	Number of key presses
	Ease of learning	
Satisfaction	Preference	Impressions
	Affinity	Change in task performance
	Tiredness	time

2.1 Experimental Design

Each participant was given the task of setting up a piece of equipment by selecting menu items, as shown in Fig. 3. The display screen consisted of an issue statement, ladder menu, and 4-direction key. A different scene appeared depending on the factor of the experiment. The factors in this experiment were age range, type of operation method, number of menu items, and depth of menu hierarchy.

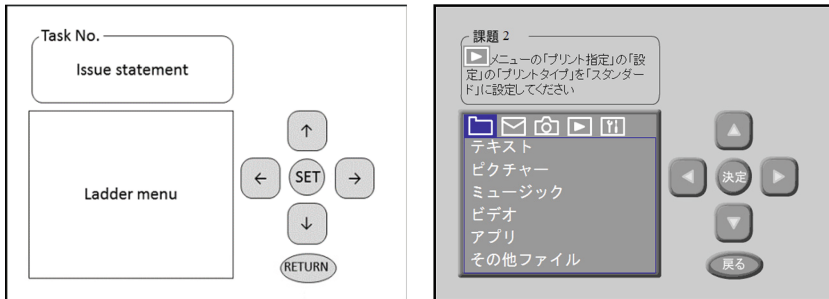


Fig. 3. Design of simulator screen and an example of the task of Type 3

- Two age ranges were used: 22–59 years and 60 + years.
- Three types of operation methods were selected from the six types listed above: Types 1, 2, and 3.
- The number of items in a menu. Six items could be displayed at once on the menu without scrolling. When nine items were displayed, the participant had to scroll the menu, otherwise the target did not appear.
- Two levels of number of layers were tested, 3 and 5 layers, including the target item.

In this experiment, factors b, c, and d were within-subject factors.

2.2 Participants

A total of 62 people participated in this experiment. The sample consisted of 40 men and 22 women. The number of participants for each operation condition of the device and their age ranges are shown in Table 2.

Table 2. The number of participants

Operation tool	Age range(Years)	
	Age22-59	Age60+
Touch Panel	15	15
Mouse	17	15

Thirty of the participants came to a sound-insulated room that was prepared for the experiment. They used a touch panel to operate the menus.

The other 32 participated using their own PC from their home or office over the Internet. They probably operated the menus with a mouse.

2.3 Apparatus

For the experiments in the laboratory, the experimental arrangement comprised a 6-inch touch panel display and a PC. In contrast, the experiments over the Internet comprised a mouse, display monitor, and PC. In addition, the server communicating with them was placed on the Internet.

We built a simulator for the tasks that worked in a Web browser, as shown in Fig. 3. The simulator received a set of stimulation data from the management server and transmitted a log of a participant’s operations. When the key operation event occurred, the event was logged in milliseconds.

2.4 Stimulus

The questions and menu items were generated from data based on a survey of products such as digital cameras. Fictitious menus relating to digital cameras, PDAs, and mobile phones were created, as shown in Fig. 4.

—「撮影」メニューの「ISO感度」を「100」に設定してください
(Please select "Camera", select "ISO speed", and set to 100 the ISO speed number.)

—「再生」メニューの「スライドショー」の「設定」の「再生間隔」を「5秒」に設定してください
(Please select "Play", select "Slideshow", select "Setting", select "Interval", set to "5s" the interval time.)

撮影	(Camera)
ムービー	(Movie)
再生	(Play)
画面表示	(Screen info.)
設定	(Setting)

Fig. 4. Examples of questions and menu items

A question sentence was presented that imitated a typical sentence in the operating manual of a piece of equipment. Each sentence was constructed along the order of the menu items until the target was reached. Ten questions for each of the fictional menus were prepared.

The number of levels and menu items were controlled depending on the experimental conditions. However, the top-level menu of each condition has only five items. The length of shortest path to each target was controlled by the almost conditions. A condition of the number of level of menu layer has a case where the length is different (Fig. 5).

To measure user satisfaction, which is one of the characteristics defining usability, a questionnaire was used. A 22-item questionnaire was prepared after reviewing the Software Usability Measurement Inventory [1, 3, 4].

1. I have been able to quickly understand how to operate the equipment.
2. I think that there are any useless way in the operating procedure.
3. I have been able to quickly learn how to operate the equipment.
4. I think the operations are unified.
5. I thought the operations were not as hard as expected.
6. I have the frustration on the behavior of the system.
7. I was often get lost what to do.
8. I thought it was easy to understand the operation procedure.
9. I thought it was easy to find the target item I need.
10. I thought the menu structures were easy to understand.
11. I had not been able to know my position during the operation.
12. I think the operation of the equipment is friendly.
13. I had been tired while operating.
14. I became familiar with the operating procedure.
15. I have confidence in the menu operation of such equipment usually.
16. I thought my operations were excellent.
17. I thought the placement of menu and the size of character were clear.
18. I thought the display of screen was easy to see.
19. I thought the size of the screen was appropriate.
20. I had eye fatigue while this experiment.
21. I left comfortable using the system.
22. I thought the display speed was appropriate.

Fig. 5. Questionnaire for participants

The questionnaire was created based on some web usability tests, such as SUS. Each question item is classified into four terms of the following.

Q1 to 6: Understandability of the operation system

Q7 to 11: Clarity of the menu structure

Q12 to 16: Familiarity

Q17 to 22: Visibility

2.5 Procedure

Equal numbers of participants used the three navigation methods. Each participant was assigned 30 tasks. The order of the tasks was randomized.

First, participants were to ask practice a task twice. They then performed the actual set of assigned tasks. The time limit for each task was set to 10 m. After all of the tasks were complete, the participants responded to the questionnaire on the screen.

3 Results

Task execution time is the number of seconds taken to reach the goal menu item from the start of the task. Table 3 shows the mean and standard deviation of the task execution time for each condition.

A t-test of these mean values did not show a significant difference ($t(60) = .752$, $p = .45$).

Table 3. Average and standard deviation of task execution time of all tasks

	Operation tool	
	Touch Panel	Mouse
Average(second)	17.63	18.06
SD	13.37	10.08

3.1 Task Execution Time for Each Condition

In the following analysis, results were statistically processed by mixing the output data of the two types of operating device: the touch panel and the PC.

Tables 4 and 5 show the average task execution time for each condition.

Table 4. Average task execution time in 22–59 years

Menu Hierarchies	Menu Items	Types of navigation methods			Average
		Type 1	Type 2	Type 3	
3	6	11.2	11.3	12.2	11.6
	9	10.4	12.3	12.0	11.5
5	6	17.5	20.9	19.1	19.0
	9	14.8	16.6	16.7	16.0
Average		13.3	15.1	14.8	14.3

An ANOVA was used to test the main effects and interactions for significance. Significant main effects were observed for all factors ($p < .01$ for each). There were significant interactions between the age ranges and menu levels ($p < .01$) and between the menu levels and menu items ($p < .01$).

Table 5. Average task execution time in 60 + years

Menu Hierarchies	Menu Items	Types of navigation methods			Average
		Type 1	Type 2	Type 3	
3	6	20.9	22.1	19.4	20.9
	9	20.4	22.8	22.6	21.7
5	6	31.7	34.9	34.2	33.3
	9	26.3	28.7	29.7	27.3
Average		24.5	26.8	26.1	25.6

The main effect of the operation methods, according to a post-hoc test by Tukey's HSD, was the difference between Type 1 and Type 2 ($p < .05$).

In each interaction between the age range and menu levels, as well as the menu levels and menu items, the simple main effects were tested. The difference between age ranges 22–59 and 60 + was greater when the number of menu levels was five. In the interaction of the menu levels and menu items, while the number of menu items was not significant for a menu consisting of three levels, the task execution time of a menu with nine items was significantly shorter than one with six items ($p < .01$) for five levels.

3.2 Number of Key Presses for Each Condition

Tables 6 and 7 show the average number of key presses from start to finish for each task.

Significant main effects were observed for three factors: the type of operation method, age range, and number of menu levels. No main effect was observed for number of menu items. There was a significant interaction between the menu levels and menu items ($p < .01$).

Table 6. Average number of key presses in 22–59 years

Menu Hierarchies	Menu Items	Types of navigation methods			Average
		Type 1	Type 2	Type 3	
3	6	12.7	11.7	12.6	12.3
	9	13.4	16.1	16.3	15.2
5	6	22.2	23.6	25.2	23.6
	9	20.0	22.5	21.3	21.2
Average		17.0	18.6	18.8	18.1

According to a post-hoc Tukey's HSD multiple comparison, in the main effect of operation methods, there were significant differences between Type 1 and Type 2 ($p < .01$) and Type 1 and Type 3 ($p < .05$).

In the interaction between the menu levels and menu items, a simple main effect was tested. A menu with six items had significantly fewer key presses than one with nine items ($p < .01$) in a menu consisting of three levels. However, in a menu consisting of five levels, a menu with nine items had significantly fewer key presses than one with six items ($p < .01$).

Table 7. Average number of key presses in 60 + years

Menu Hierarchies	Menu Items	Types of navigation methods			Average
		Type 1	Type 2	Type 3	
3	6	14.3	14.7	13.9	14.3
	9	15.2	18.3	18.3	17.0
5	6	24.8	25.2	27.9	25.8
	9	21.2	23.3	24.4	22.7
Average		18.7	20.4	21.2	19.9

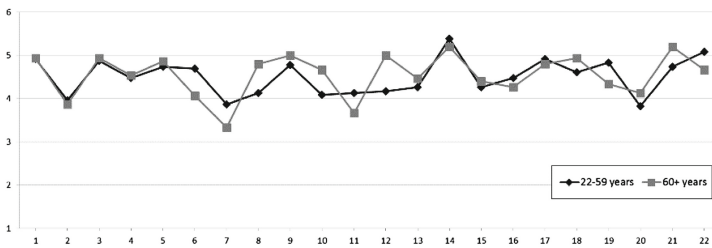
3.3 Retry Frequency

A retry is when a user returns to a previous state when he/she determines that an incorrect key operation has been made. Retry frequency is often used as a metric to assess the clarity of the menu structure and its associated key operations.

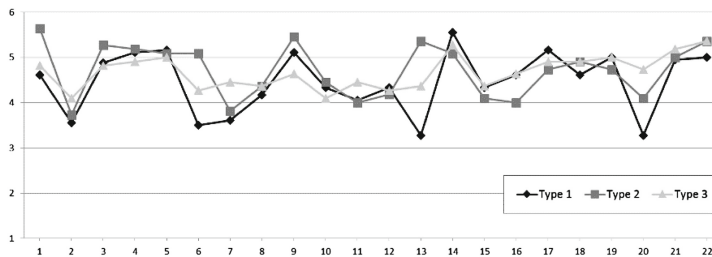
Two definitions for a retry were possible in this experiment. One definition of retry is that the operator backtracks through the menu hierarchy. The other definition is the act of returning from too many of up or down operations in the menu. When navigating up or down in a menu that appears on the screen, however, it was not possible to distinguish whether a retry had occurred or the operator was using a particular strategy in the case of an endless scroll. Hence, we examined only the case of returning to the menu hierarchy. Table 8 shows the retry frequency.

Table 8. Retry frequency for each condition

Age ranges	Types of navigation method			Total
	Type 1	Type 2	Type 3	
20-59	27	20	71	118
60+	35	73	79	187
Total	62	93	150	305



(a) Aggregating by age ranges



(b) Aggregating by types of navigation method

Fig. 6. Plot of aggregating questionnaire

3.4 Questionnaire

The results of the questionnaire are summarized in Fig. 6. About the questions that is a negative form in Japanese, the value of the answers were inverted. The graphs of the age ranges indicate the same tendencies (Fig. 6(a)), but that of the operation method types indicate some partial differences (Fig. 6(b)). Figure 7 shows question items for which the difference of each category by types of navigation method.

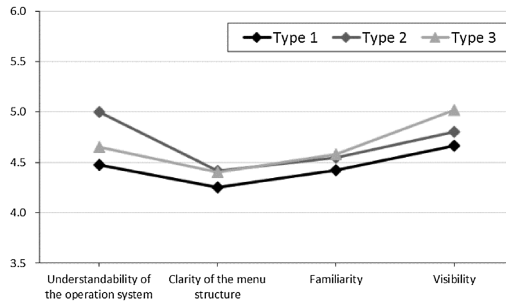


Fig. 7. Difference of each category by types of navigation method

4 Conclusions

Given the task execution time, number of key presses, and retry frequency results, it is appropriate to recommend the operation method of Type 1. Equivalently, in terms of effectiveness and efficiency, Type 1 is the best option. In addition, it is recommended to reduce the number of menu levels. In this study, in the same way as the results of previous studies on the depth and breadth of the menus [9, 10, 12], the number of menu items is less impact on the effectiveness and efficiency. For older users, the menu levels should be made as shallow as possible because there was an interaction between the age range and the menu hierarchies. In the age range condition, even though there is no difference in the number of key presses, the task execution time of 60+ years participants was significantly slower than that of the 22–59 years. However, the degree to which these parameters must be reduced varied depending on the operation method. If the number of items is small, Type 2 is also recommended.

About retry frequency, for the age range 60+, Type 1 had the smallest. For the age range 22–59, Type 2 had the smallest. At any age ranges, the retry frequency of Type 3 was the largest. In contrast, Type 1 was small. However, Type 2 in age ranges was large difference of the retry frequency. Even though the retry frequency of young user was small, that of the elderly was close to the highest value.

According to the results of questionnaire as shown Fig. 7, Type 1 has the most advantage about the understandability of the operation system. About the Visibility, Type 3 is seemed to be the most advantage.

References

1. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: adding an adjective rating scale. *J. Usability Stud.* **4**(3), 114–123 (2009)
2. Allen, R.B.: Cognitive factors in the use of menus and trees: An experiment. *IEEE J. Sel. Areas Commun.* **1**, 333–336 (1983)
3. Frøkjær, E., Hertzum, M., Hornbæk, K.: Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated?. In: *Proceedings of the ACM CHI 2000*, pp. 345–352 (2000)
4. The de facto industry standard evaluation questionnaire for assessing quality of use of software by end users. <http://sumi.ucc.ie/>
5. Information Technology Research and Standardization Center (INSTAC), JSA. The standardization of 4-direction key interaction of mobile information devices (2010)
6. ISO/IEC 17549-2: User interface guidelines on menu navigation – Part 2: Navigation with 4-direction device
7. ISO 9126-1:2001: Software engineering – Product quality – Part 1: Quality model
8. ISO 9241-11:1998: Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability
9. Kiger, J.I.: The depth/breadth trade-off in the design of menu-driven user interfaces. *Int. J. Man Mach. Stud.* **20**(2), 201–313 (1984)
10. Norman, K.L.: *The Psychology of Menu Selection: Designing Cognitive Control at the Human/Computer Interface*. Ablex Publishing Corporation, Norwood (1991)
11. Walker, N., Smelcer, J.B., Nilsen, E.: Optimizing speed and accuracy of menu selection: a comparison of walking and pull-down menus. *Int. J. Man Mach. Stud.* **35**(6), 871–890 (1991)
12. Sisson, N., Parkinson, S.R., Snowberry, K.: Considerations of menu structure and communication rate for the design of computer menu displays. *Int. J. Man Mach. Stud.* **25**(5), 479–489 (1986)
13. Teitelbaum, R.C., Granda, R.E.: The effects of positional constancy on searching menus for information. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1983*, pp. 150–153 (1983)