

# Usability evaluation of four mechanical pencils using usability task analysis

Toshiki Yamaoka

*\* Wakayama university, Faculty of Systems Engineering, tyamaoka6@gmail.com*

**Abstract:** Usability task analysis was introduced and four mechanical pencils were evaluated based on it in order to verify the effectiveness. It is a kind of task analysis and can get a lot of user requirements from viewpoint of good and bad point, and SCT. The user requirements and problems of each product and GUI are collected together. Each product and GUI can be compared using it simultaneously. It can analyze the data statistically using multiple regression, FCA and so on. Although it takes time to complete compared with protocol analysis, it seems natural to conclude that it is a useful method for developing and designing products or GUI.

**Key words:** *Usability Evaluation, Usability Task Analysis, FCA, multiple regression, SCT*

## 1. Introduction

The famous protocol analysis [1] has been used frequently in order to extract usability problems in product development. The user requirements for product development have been created by various interview methods and the direct observation. The problems and user requirements were used respectively in product development. The problems are improved and become the user requirements, while user requirements are collected separately. As these methods are qualitative, statistics are not used in order to examine relation between the problems, or user requirements deeply.

Usability task analysis is a kind of task analysis [2,3] and proposed in order to extract usability problems and user requirements together and efficiently. The usability task analysis also can analyze data quantitatively using DEMATEL, FCA [4,5,6], multiple regression and so on.

## 2. What is the usability task analysis?

### 2.1 The procedure

The procedure of collecting user requirements is as follows (table 1.).

(1) The test participants check and evaluate every task of products or GUI.

Every task of products or GUI is evaluated from view point of the good and bad points of the task. Test participants are asked to fill out appropriate evaluation words in the blanks in Sentence Completion Test (SCT) [7] after operating every task. SCT is aimed at collecting the causal relationship of usability of the task.

---

SCT

As the product or GUI in the task is [ (a) ], I feel that [ (b) ].

So, the task has [good or poor ] usability.

---

SCT example in case of alarm clock

As the figures on dial are [ big ], I feel that [ it is easy to see them ].

So, the task has [ good ] usability.

The every task of products or GUI is graded according to the Likert scaling: strongly agree(5), agree(4), neutral(3), disagree(2), strongly disagree(1). The good points become user requirements ,while the bad points are changed into good user requirements. For example, bad point “heavy” is changed into user requirement “light” which is an antonym of “heavy”. The user requirements are structured.

Table 1. Usability task analysis

	Product-1	Product-n	user requirements
<b>Task(1)</b>	Point [ 2 ]	Point [ 4 ]	
	good point	good point	good user requirements
	[ (a) ]→[ (b) ]	[ (r) ]→[ (s) ]	(a) , (r)and (d'),(m')(d,m is changed into d',m')
	bad point	bad point	bad user requirements
	[ (d) ]→[ (e) ]	[ (m) ]→[ (n) ]	(d) and (m)
<b>Task(n)</b>	Point [ ]	Point [ ]	
	good point	good point	good user requirements
	[ ]→[ ]	[ ]→[ ]	
	bad point	bad point	bad user requirements
	[ ]→[ ]	[ ]→[ ]	
<b>Synthetic evaluation</b>	Point [ ]	Point [ ]	
	good point	good point	good user requirements
	[ ]→[ ]	[ ]→[ ]	
	bad point	bad point	bad user requirements
	[ ]→[ ]	[ ]→[ ]	

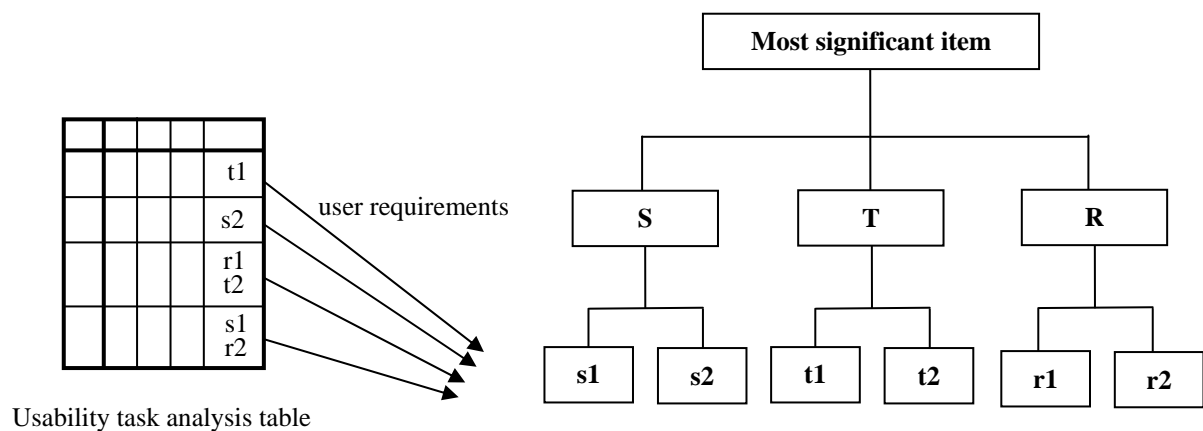


Figure 1. Structured user requirements

- (2) The test participants check all tasks as the synthetic evaluation and grade using the Likert scaling.
- (3) The collected user requirements are structured using card sorting (figure 1.).
- (4) The collected user requirements are analyzed using DEMATEL, FCA, multiple regression and so on.

## 2.2 Statistical analysis

DEMATEL, FCA, multiple regression and so on are used for analyzing user requirements. Especially as FCA (Formal Concept Analysis) is unfamiliar, the function of it is explained as follows. FCA can make the relationship between objects and the attributions clear. In case of SCT, participants evaluate a task from viewpoint of usability in the blank (a), and fill the feeling in the blank (b). As the data in the blank (a) and (b) means the objects and attributions respectively, these data can be analyzed by FCA in case of good or poor usability respectively.

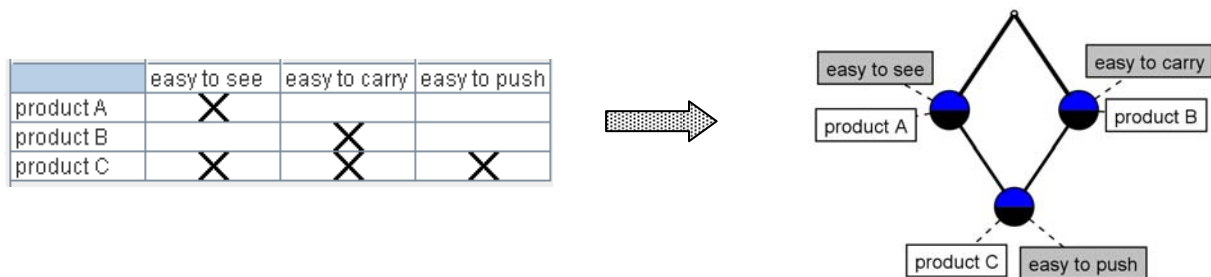


Figure 2. The Hasse diagram

When product A, B and C as objects were evaluated using “easy to see”, “easy to carry” and “easy to push” as attributions, the following diagram is drawn by FCA. As the two routes from the top of the diagram are linked to the product C, it is found to consist of “easy to see”, “easy to carry” and “easy to push”. The product A and B are also understood by the same way. Attributions counted frequently in the matrix are located at a higher position in a diagram. Two items, “easy to see”, “easy to carry”, in case of figure 2 are located at the higher position in a diagram.

## 3. Four mechanical pencils were evaluated by the usability task analysis

The purpose of the test is to confirm the effectiveness.

### 3.1 Method

Participants: 11 university students (21-24 years old, male:9, female:2)

**Tasks:** They evaluated four mechanical pencils using the usability task analysis

task1: “easy to take them”, task2: “hold them”, task3: “advance a lead”,

task4: “write”, task 5: “synthetic evaluation”

**Mechanical pencils** (figure3)

**A:** The mechanical pencil for drawing

**B:** The mechanical pencil is thin and small.

**C:** The mechanical pencil is thick with no slippery grip.

**D:** The mechanical pencil is ordinary type with no slippery grip.



Figure 3. Four mechanical pencils

### 3.2 Results and Discussion

#### (1) Structured user requirements (figure 4)

User requirements written in synthetic evaluation were collected and classified. Very important user requirements were “firm finishing”, “no slippery grip”, “moderate weight and size” and “thick diameter”. The most significant item “Writing comfortably and certainly” for mechanical pencils was structured by these items. Important user requirements to design mechanical pencils are “good shape for comfortable writing” and “grasping mechanical pencils firmly”.

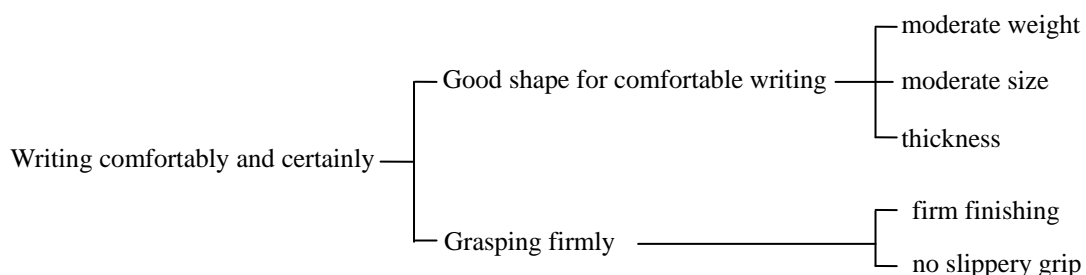


Figure 4. Structured user requirements written in synthetic evaluation

#### (2) Important tasks influenced synthetic evaluation were obtained by multiple regression

Response variable: “synthetic evaluation”

Explanatory variable: “easy to take them”, “hold them”, “advance a lead”, “easy to write”

The important tasks influenced synthetic evaluation obtained using AIC, Ru, adjusted multiple correlation coefficient were as follows.

The mechanical pencil A: “hold them”, “advance a lead”

The mechanical pencil B: “easy to take them”, “advance a lead”, “easy to write”

The mechanical pencil C: “advance a lead”, “easy to write”

The mechanical pencil D: “hold them”, “advance a lead”, “easy to write”

As product A is designed for drawing line, “hold them” and “advance a lead” seem to be weighted rather than “write”. Table 2 shows that “advance a lead” is very important requirement for four mechanical pencils.

The requirements in order to importance are as follows.

“advance a lead” > “easy to write” > “hold them” > “easy to take them”

Table 2. Results of the multiple regression

	A	B	C	D
easy to take them		V		
hold them	V			V
advance a lead	V	V	V	V
easy to write		V	V	V

Focused on “advanced a lead”, the requirements are structured. The bad user requirements are changed into the good user requirements. Very important design factors are found according to the structured requirements of advanced a lead (figure 5).

- force to push button: light force, pushing lightly and smoothly without slipping
- stroke to push button: optimum stroke for pushing
- feeling to push button: good click and minute feeling of pushing button
- sound to push button: good sound for clicking
- shape of button: good button shape and size for pushing

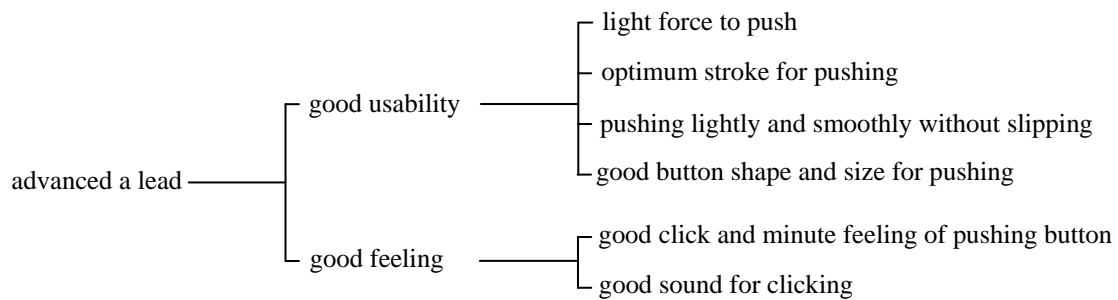


Figure 5. The requirements of advanced a lead structured

(3) The relation between four tasks and four mechanical pencils was found by FCA

The average data of the task for the mechanical pencil were changed into binary data (1,0). These data are shown below. According figure 6, the mechanical pencil D is evaluated highly, while B evaluation is low. As B is ordinary type with no slippery grip and has no weak points, D seems to be evaluated highly.

	easy to take them	hold them	advance a lead	easy to write
A	0	0	1	1
B	0	0	0	0
C	1	1	0	1
D	1	1	1	1

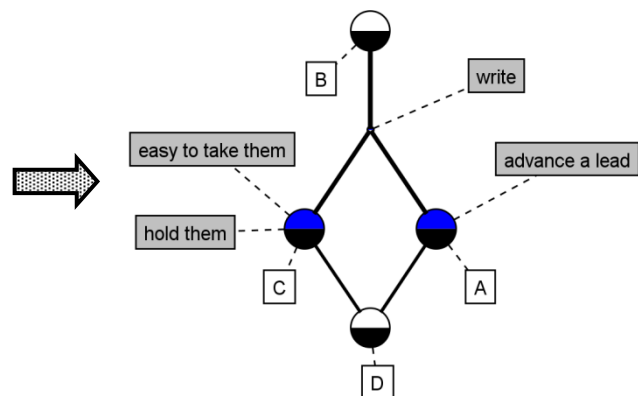


Figure 6. The relation between four tasks and four mechanical pencils

	moderate weight	firm finishing	no slippery grip	moderate size	thickness	easy to handle	easy to write	easy to see
A	1	1	1	0	0	0	1	1
B	0	0	0	0	0	1	0	0
C	1	1	1	1	1	0	0	0
D	0	1	1	1	0	1	0	0

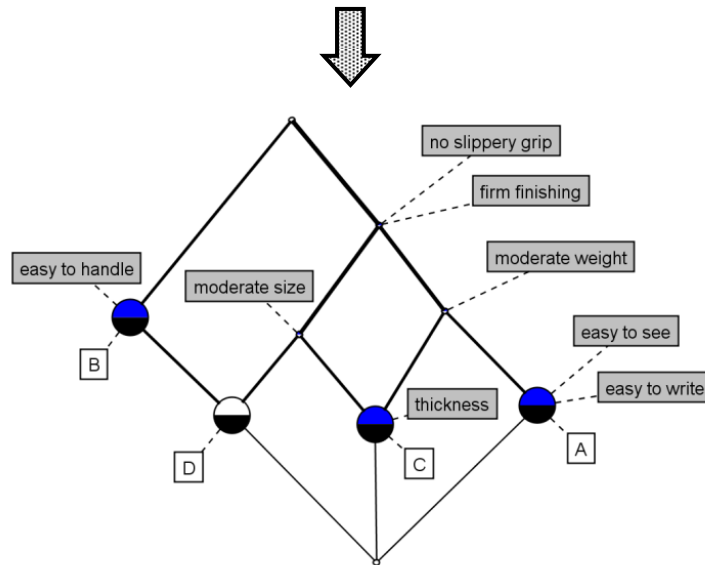


Figure 7. The relation between important user requirements and four mechanical pens

(4) The relation between important user requirements and four mechanical pencils were found by FCA (figure 7).

The important user requirements were extracted by the usability task analysis. If each pens are decided to be include these user requirements based on participants' comments, 1 is filled out in the cell of matrix. According to Figure7, "no slippery grip" and "firm finishing" are shown as important user requirements. However, participants' comments caused "easy to write" low evaluation differently from the results of multiple regression.

#### 4. The comparison with protocol analysis

The differences between the usability task analysis and protocol analysis based on the evaluation data of four mechanical pencils and participants' comments are shown below.

- (1)The protocol analysis is easy and does not take time, while the usability task analysis takes time.
- (2)The usability task analysis can get a lot of user requirements from viewpoint of good and bad point, SCT. User requirements and problems of each product and GUI can be collected together. And each product or GUI can be compared using it simultaneously.
- (3)The usability task analysis can analyze data statistically using multiple regression, FCA and so on.
- (4) The usability task analysis can be available at any product development stage.

## 5. Conclusions

The usability task analysis was introduced in the paper and four mechanical pencils were evaluated based on it. Although it takes time to complete, it gives designers and engineer a lot of user requirements to develop products or GUI. As analyzing data statistically, the valuable data can be used in order to construct product or GUI design. It seems natural to conclude that the usability task analysis is a useful method for developing or designing products or GUI.

## References

- [1] Bainbridge,L and Sanderson , P .(Wilson , J.R. and Corlett, E. N,ed) (1995) *Verbal protocol analysis*, Evaluation of human work, A practical ergonomics methodology, 2ndEd, pp169-201
- [2] Yamaoka, T. (2002.) *A universal design method using 3 point task analysis and 9 universal design items*, Korea,Journal of the Science of Emotion & Sensibility:Vol5,No2, pp63-72
- [3] Stammers, R.B.and Shephard, (Wilson , J.R. and Corlett, E. N,ed) (1995) *Task analysis*, Evaluation of human work, A practical ergonomics methodology, 2ndEd, pp144-168
- [4] Ganter,B and Wille,R.(1999) *Formal Concept Analysis*, Springer
- [5] Yamaoka,T.(2010) *A study on an analyzed method of the emotional images of product* , The proceedings of the Kansei engineering and emotional research international conference, pp 1995-2002
- [6] Yamaoka, T.(2012) *A Kansei Evaluation Method Based on Sentence Completion Test and Formal Concept Analysis*, Proceedings of the International Conference on Kansei Engineering and Emotion Research KEER 2012, pp145-150
- [7] Ikuma, J. and Inematsu, N.(Agari,I.,general ed)(2001) *Sentence Completion Test*, Handbook of psychological assessment 2<sup>nd</sup> edition, Nishimura Shoten, pp232-246