A new data collection method for usability testing:
NEM : Novice Expert ratio Method

ABSTRACT

This poster describes a new quantitative measure for usability testing. Various aspects of this method are described here, i.e. the basic concept, the measurement procedure, an interpretative technique, some examples applied to various products, an alternative technique using model-based evaluation, a tentative procedure for setting the criterion for NEM value, and an automatic measurement program "GAP FINDER".

1 INTRODUCTION

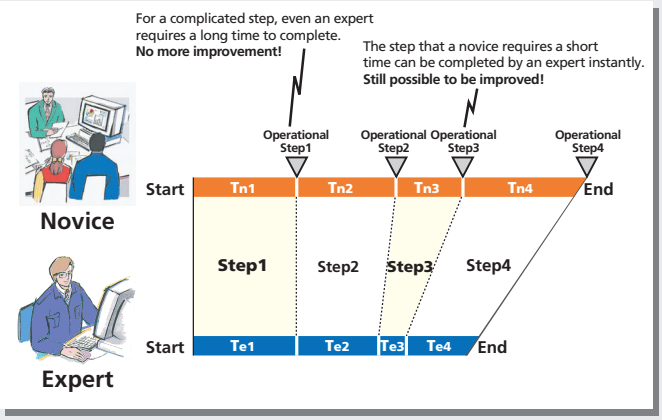


Fig. 1

・Is it enough to consider steps that resulted in the error?

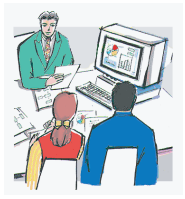
・Should a step that required a long time to complete be improved and be completed in a shorter time?

・Is it true a step that required a short time does not have a problem to be improved?

In this poster, you will find the answer and why we had an idea to develop a new method to differentiate the step that has a problem from the step that has no problem based on the time measure.

Tn: the average time in seconds taken by the novice users.
Te: the average time in seconds taken by the expert users.

2 THE METHOD



time required by **Novice**

VS

time required by **Expert**

$$\text{NE ratio (Ri)} = \frac{\text{Novice (Tn)}}{\text{Expert (Te)}}$$

Fig. 2

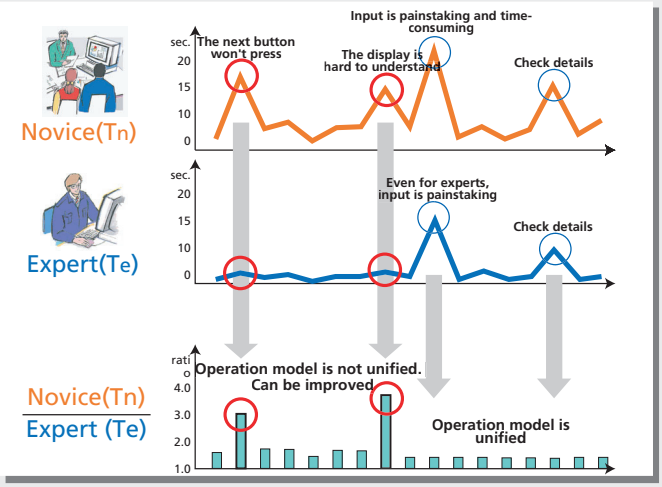


Fig. 3

The basic idea of this new method, NEM, which is an acronym for Novice Expert ratio Method, is to compare the time required by a novice user to the time required by an expert user(Urokohara et al. 2000).
It is conceivable that the time required by the expert will be the minimum because the expert user must have a knowledge on the correct procedure, thus comparing the time required by the novice user and the time required by the expert user will bring the ratio on the degree of the difficulty of that specific procedural step.

Thus, NE ratio(R) can be calculated by the following equation:
Ri = Tn / Te (1)

Where Ri stands for NE ratio for step i where i=1 to n for a certain task procedure, Tn is the average time in seconds taken by the novice users and Te is the average time in seconds taken by the expert users.
When the calculation of NE ratio for the whole procedure was finished, steps that showed higher NE ratio will be inspected deliberately to clarify why the novice took such a long time than the expert.

$$\text{Task Performance} = \frac{S - NEh}{S}$$

NEh=Num. of task steps with high NE ratio
S=Num. of total task steps

Fig. 4

3 APPLICATION EXAMPLES

Case 1: Car Navigation System (Subjects: 30 novices and 4 experts)

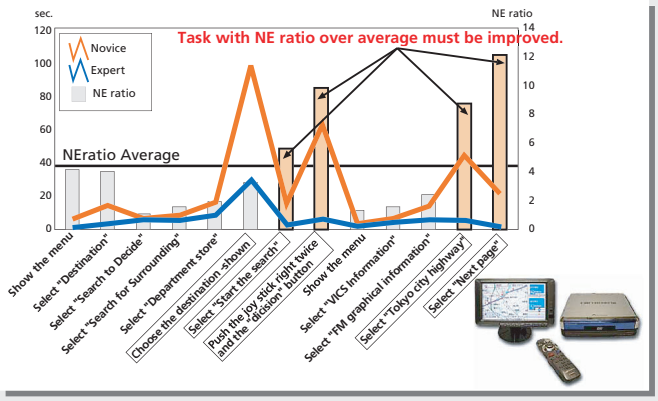


Fig. 5

It this case, NEM was applied to the evaluation of the car navigation system.
The boxed items are the operational steps that were diagnosed to have the usability problems by the inspection of usability specialists, most of which have high NE ratio.

It was confirmed that NEM is quite useful in identifying operational steps with a usability problem.

- Orange line: average time required by the novice users
- Blue line: average time required by the expert users
- Bar graph: NE ratio

Case 2: Set up Screen Design for a Scanner (Subjects: 5 novices and 2 experts)

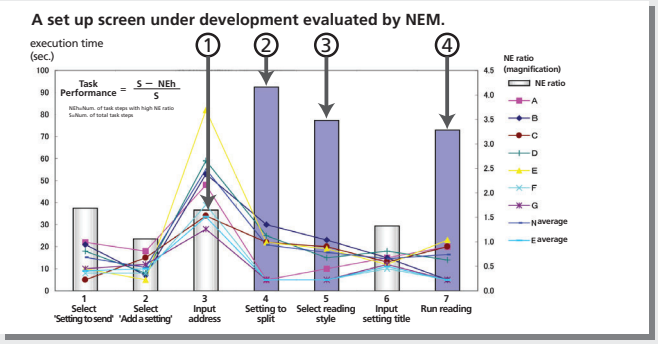


Fig. 6

A QC department of a company picked up an idea of NEM from our web site and applied it to their product.

One of the most practical characteristics of NEM is that even a non-specialist can find usability problems without difficulty, was proved to be true for this case.

Operational steps with usability problems:
① Users took time to input address. (Step 3)
② Users were confused to select 1 out of 3 alternatives. (Step 4)
③ Users were confused to select 1 out of 4 alternatives. (Step 5)
④ Users were confused because the scanning was run without confirmation of registration. (Step 7)

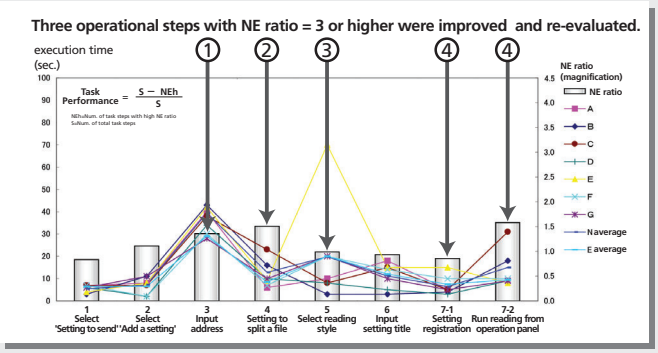


Fig. 7

After improved:
① Turned out to be a time-consuming step even for accustomed users. No improvement was possible. (Step 3)
② Reduced alternatives from 3 to 2 and used the user-friendly terms,which let users recognize the difference easily. (Step 4)
③ Used the user-friendly terms without extra explanation, which let users recognize the difference easily. (Step 5)
④ Divided the step into 2 groups to turn off the users' confusion. (Step 7-1/7-2)

NE ratio for these operational steps became lower than before the improvement, which proved the validity of NEM.

4 MODEL-BASED TECHNIQUE

Comparison of NE ratios based on the real measurement and the estimation by KLM and MODAPTS

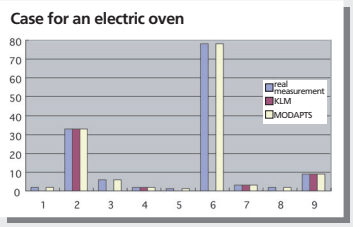


Fig. 8

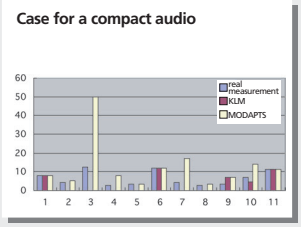


Fig. 9

Is it always possible to find the qualified expert users for targetted product?
- It is not always true regarding the designers who designed the product as expert users.
- It is sometimes difficult to regard the user with a certain amount of training as an expert.

We decided to adopt a model-based approach that will generate the time estimates instead of human experts.
- Key Stroke Level model proposed by Card et al. (1983)
- MODAPTS or the Modular Arrangement of Predetermined Time Standards proposed by Komatsubara et al. (1995)

Case studies for an electric oven and a compact audio are described here.

5 CRITERION FOR NE RATIO

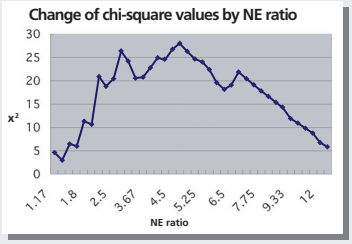


Fig. 10

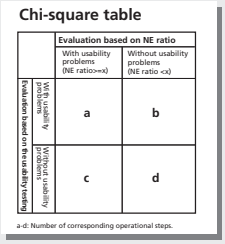


Fig. 11

For practical purposes, it is convenient if there is a certain criterion for NE ratio to differentiate the procedural steps that have the usability problem from the steps that have not.

We tried to determine the criterion by applying chi-square method. Based on our analysis, NE ratio around 4.5 was found to be critical or the threshold.

6 AUTOMATIC CALCULATION BY "GAP FINDER"

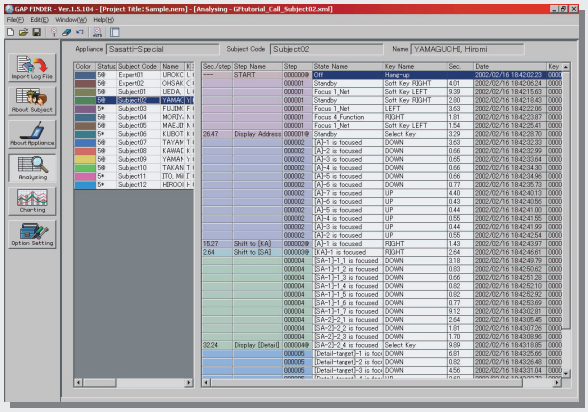


Fig. 12

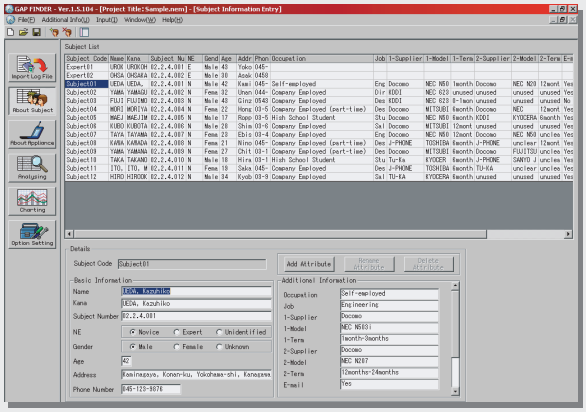


Fig. 14

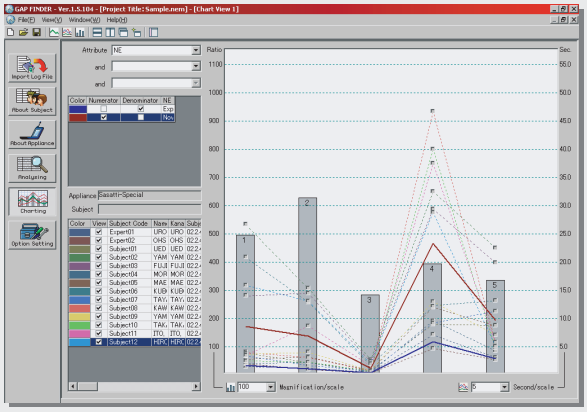


Fig. 13

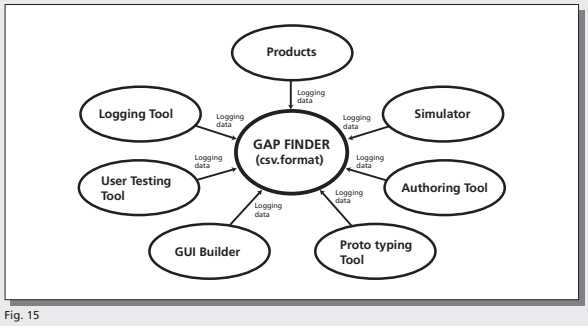


Fig. 15

Because it requires some amount of time to calculate NE ratio, there were many requests by the initial users of this method for a tool that will conveniently support the calculation job. GAP FINDER is the software that will facilitate this job.

This Windows-based software will be demonstrated at the session.

7 FUTURE PLAN

Based on the feedback from the user, we will further improve GAP FINDER.

Masaaki Kurosu
National Institute of Multimedia Education
Tokyo JAPAN
+81-53-478-1481
PFD00343@nifty.ne.jp

Haruhiko Urokohara, Daisuke Sato
Usability R&D Dept. U'eyes novas, Inc.
8th fl. Fujita INZX Building
9-5 Shinsen-cho, Shibuya-ku, Tokyo 150-0045
JAPAN
+81-3-5489-7470
urokohara@novas.co.jp

Keywords
Usability evaluation, task performance measure, novice user, expert user, keystroke level model

GapFinder