Feasibility Study of Predictive Human Performance Modeling Technique in Field Activities

Naomi Yano[1], Toshiyuki Asahi[1], Shiníchi Fukuzumi[1], and Bonnie E. John[2]

[1] Human interface Technology group, Information and Media Processing Laboratories, NEC Corporation, 8916-47, Takayama-Cho, Ikoma, Nara, 630-0101, Japan
{n-yano@cb,t-asahi@bx,s-fukuzumi@aj}.jp.nec.com
[2] Human-Computer Interaction Institute, Carnegie Mellon University, 5000 Forbes Avenue Pittsburgh PA 15213, USA
bej@cs.cmu.edu

Abstract. This paper reports the results of usability evaluation based on the predictive human performance models applied to any products, and introduces application of a tool for predicting operational time to an IP phone and an electronic health record system UI consulting. We assessed effectiveness of this tool using efficiency estimation technique and extracted the practical the practical problems. In the human interface (HI) consultation process, as the tool predicts execution time of the current version and of the improved version, the improvement effect could be assessed. For an IP phone, we created the modified user interface designs from the point view of operational efficiency, so that we could indicate effectiveness of this tool by comparing task execution time. For an electronic health records, however, it is difficult to "directly" verify the effectiveness of the modified user interface designs from an efficiency standpoint. Through an evaluation scenario, the tool provided data that is necessary for assessment of improvement in this case.

Keywords: Usability Evaluation, Predictive human performance model, Consultation, Efficiency, Understandability.

1 Background and Objectives

Usability evaluation methods that use predictive human performance models are easy to deal with because they can evaluate without subject or prototype [1][2]. Such methods have several advantages over traditional usability testing methods: (1) low time and effort cost, (2) providing data about human performance early in the development life-cycle, (3) being used to generate quantitative usability metrics for proposed designs.

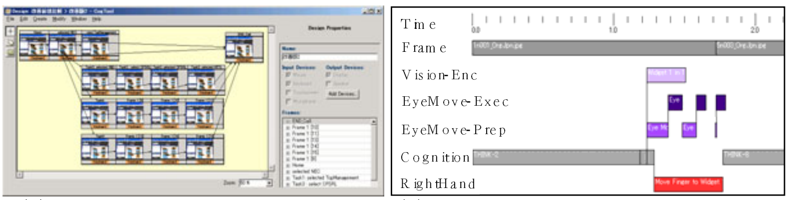
Predictive human performance modeling techniques have been investigated for many years and their feasibility regarding such factors as prediction accuracy or data input cost has been validated through many studies [3]. However, unexpected problems may still remain when they are applied to actual field activities, e.g., HI evaluation/improvement. Thus, the authors tried to apply a predictive human

M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2011, LNCS 6771, pp. 180–189, 2011.
© Springer-Verlag Berlin Heidelberg 2011

performance modeling technique to software products and to verify its feasibility in order to develop a usability guideline.

2 Tool for Predicting Operational Time

We selected CogTool as an example of predictive human performance modeling techniques. It is a kind of HI prototyping tool that can predict the operation time of a skilled expert in relation to a target task on the order of seconds [4]. Because HI operation times can be predicted using this tool, the operational efficiency of multiple design proposals can be compared. For this reason, it can be applied for benchmarking design proposals in the early product development stages. The design proposal is represented with a storyboard, and the correct operational procedures for executing a certain task are entered in the form of manipulating HI parts on the screen.



(a) Entry screen of storyboard (b) Visualization of the prediction

Fig. 1. Input and output of CogTool: (a) Users entry screen of storyboard on this screen. The storyboard is represented by screen images, widgets and screen transition. (b) This screen shows the computational process of the predicted time for human performance. Time runs from left to right and the widths of the boxes are proportional to the time they take to execute. The rows of boxes are different types of "operators (perceptual, cognitive or motor process, etc)" that happen in the course of performing these Tasks.

3 Methods

The tool is applied in HI consultation to assess the usefulness of a predictive human performance modeling technique. Because it can predict task execution time, we think the tool could be used in two major ways for consultation: (1) to assess the operational efficiency of multiple design proposals, and (2) to show the improvement effect of the modified user interface design. Here, we used the tool in the latter way.

3.1 Procedural Steps

We are commissioned to improve the user interface of a current system and propose a modified user interface design. This tool is applied in two consultation situations, one is virtually planned in the laboratory and the other is an actual one which has clients (developers) and their product. The procedural steps are as follows.

(1) Extract problems of the current system version by a heuristic evaluation.
(2) Focus on a problem.

(3) Write down a modified user interface design to solve the problem.
(4) Create a wire-frame image based on the hand-written modified user interface design.
(5) Apply CogTool to the current and improved system versions.
(6) Analyze the application results and assess the improvement effect.
(7) Present the results to clients. (for the real consultation case).

4 Apply CogTool to Systems

CogTool was applied to an IP phone system and an electronic health record system, which are target consulting systems. Former has efficiency problems and the latter has understandability problems. Modified user interface designs for solving them were created. Although the current implementation of CogTool only predicts performance time after a user becomes skilled (i.e., training or practice have eliminated understandability problems), and thus, we did not expect it to predict understandability issues, we applied it to both systems to see if it could show the efficacy of modified user interface designs. Analysis of these cases will enable testing of the practicality of the tool in two separate cases: assessing the improvement effect for efficiency and assessing the improvement effect for understandability. Described below are the cases of applying the tool to these systems.

4.1 Case (1) IP Phone System

This tool was applied in HI evaluation/improvement processes in which it is assumed that the tool is used in consultation for an IP phone system. Potential practical problems were discussed with two usability engineers. Because it was made clear that the operation of outgoing call needs many steps through the use of a heuristic evaluation, a modified user interface design was improved for efficiency. We set up four tasks and predicted task execution time by the tool. The predicted time of the improved version was 2.5 - 5.4 seconds shorter than the current version for three tasks and was 0.1 seconds longer than the current version for one task. We assumed the results should be enough to show the advantages of modified user interfaces objectively even in the actual consultation.

4.2 Case (2) Electronic Health Record System

The tool is applied in HI consultation for an electronic health record system. The details of the procedural steps are described in the following subsections.

Problems with Current System and Modified User Interface Designs

Consulting workers evaluated the current electronic health record system through the use of a heuristic evaluation and extracted problems. Described below are “Text box operations” and “Item selection”, which were a part of the problems and the modified user interface designs to solve them.

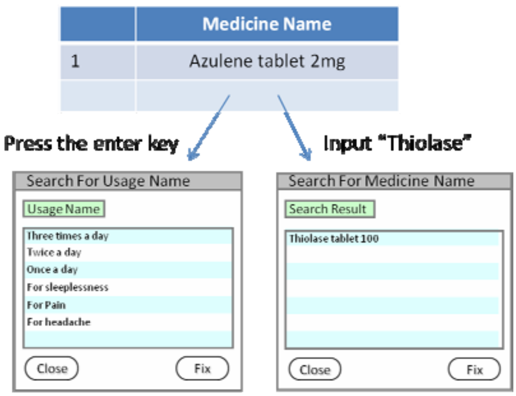
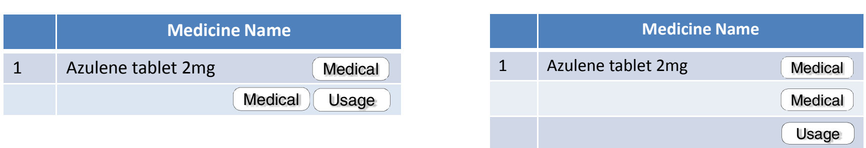


Fig. 2. Problem of "Text box operations"

Problem (1): Text Box Operations

[Problems] Users enter a medicine name in a text box and a new text box appears under it. The new text box branches the process in the input state. In the case where a user presses the Enter key without an input string, the usage search screen appears. In the case where a user enters a string, the medicine name search screen appears. For this reason, users can't predict the operational procedure visually.

[Modified user interface designs]
(1) Modified user interface design 1: The "medicine" and "usage" fields were put on the text box. This made the operation procedure clear.
(2) Modified user interface design 2: In addition to improvement (1), a text box for usage was always displayed.



(a) Design 1 (b) Design 2

Fig. 3. Modified user interface designs for the "Text box operation" problem

Problem (2): Item Selection

[Problems] Operation becomes onerous because a new window appears each time an item is narrowed down. Because the main screen hides behind many sub screen windows, a user can't see his operational behavior on the main screen. In general, a user clicks an item and presses the fix button and the selected result is reflected. But the selected result is reflected on the main screen by a single click in this interface. If a user makes a mistake in selecting an item, recovery requires many steps (he must look back at the main screen and delete the item).

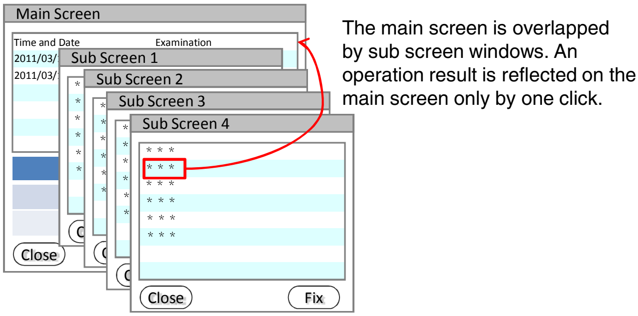


Fig. 4. Problem of "Item selection". The main screen is overlapped by sub screen windows. An operation result is reflected on the main screen only by one click.

[Modified user interface designs]
(1) Modified user interface design 1 (Single selection, depth of data is fixed):
Items are narrowed down in one window. In the last step, the "Fix" button is clicked to finalize the operation.

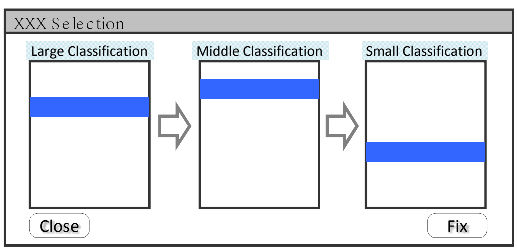


Fig. 5. Modified user interface design 1

(2) Modified user interface design 2 (Single selection, depth of data is not fixed):

・Items are narrowed down on the same screen.
・Breadcrumb navigation shows the current location.

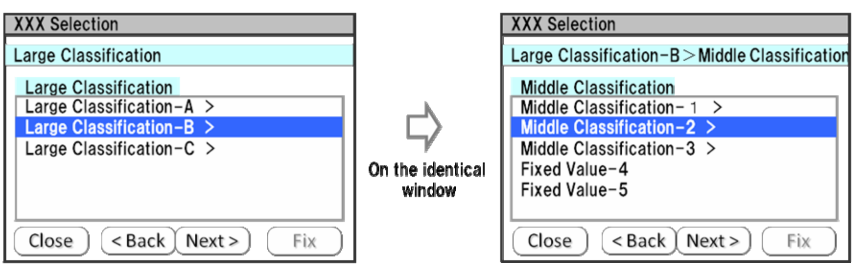


Fig. 6. Modified user interface design 2

(3) Modified user interface design 3 (Multiple selection, depth of data is not fixed):
- In addition to improvement (2), the screen has a list of selected items. The users can delete an item with the "Delete" button.
- The users push the fix button in the last step and selected items only are reflected on the main screen

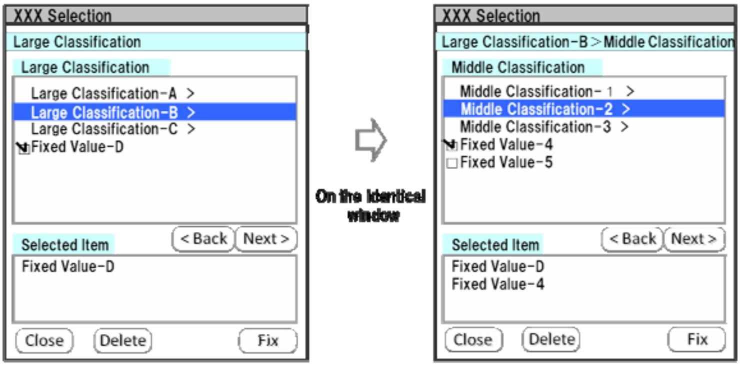


Fig. 7. Modified user interface design 3

Task Selection

The problem of "Text box operations" was that users couldn't understand the operational procedure visually. The "medicine" and "usage" buttons were put in the text box in the modified user interface design. This made the operation procedure clear. In this case, because operational objects were added, it was expected that prediction time of the improved version would be longer than the current version.

The problem with "Item selection" was that users couldn't look at their operational behavior on the main screen, and the selected result was reflected on the main screen by a single click. In the modified user interface design, items are narrowed down in one window. Clicking the "Fix" button finalized the result in the last step.

In this case, it was expected that prediction time of the improved version would be as long as the current version for the correct step. However, it was expected that prediction time of the improved version would be shorter than the current version in the recovery step. Thus, it was decided that the improvement effect would be assessed with not only correct operation steps but also with recovery operation steps for operation mistakes. We selected tasks that follow these two scenarios and applied the tool to predict task execution time. "Text box operations" has one task and "Item selection" has three tasks (correct step and two recovery steps).

Result of Applying the Tool

[Prediction for task execution time]

Prediction for execution time of the current version and the improved version were as follows (See Figure 8).

(1) Text box operations
The prediction time of the current version was the shortest. Both of the improved versions were 1.5 seconds longer than the current version.
(2) Item selection (Correct step)
The prediction times of improved versions 1-3 were shorter than the current version. The time of improved version 2 was the shortest and 1.7 seconds shorter than the current version.
(3) Item selection (Error-recovery step 1: A user executes the task with recovery operation because he made a mistake at the middle step.)
The prediction time of the current version was the shortest. The times of improved versions 1, 2 and 3 were shorter than the current version. The shorter of them was improved version 1 and 4.4 seconds shorter.
(4) Item selection (Error-recovery step 2: A user executes the task with recovery operation because he made mistake at the last step.)

The times of improved versions 1 and 3 were much shorter than the current version. The time of improved version 2 was 2.1 seconds shorter than the current version. The shorter of them was improved version 3 and 16 seconds shorter.

[Analysis of prediction time]
To analyze the cause of the prediction time differences, the timeline that shows the computational process of the predictions for human performance was checked. Described below is only the result of analysis for “Text box operations.”

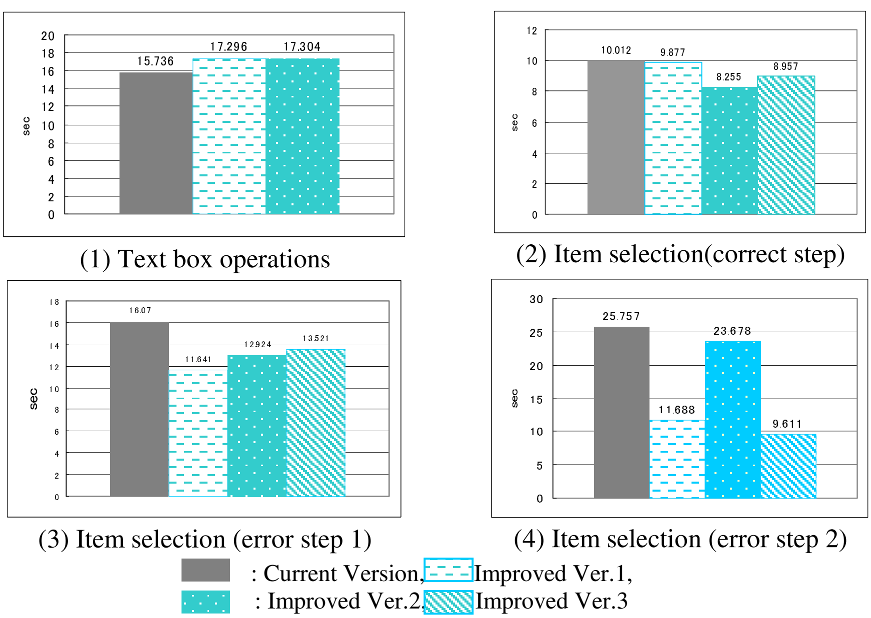


Fig. 8. Prediction time for each task

Figure 9 shows that procedural step ④ causes the temporal difference between the current version and improved version, that is, the difference between the procedure to only press the enter key on a text box in the current version and the procedure to move the cursor to the "Usage" button with a mouse and click it in the improved version. By contrast, location change of the "Together" check box in the improved version didn't cause temporal difference between the current version and improved version.

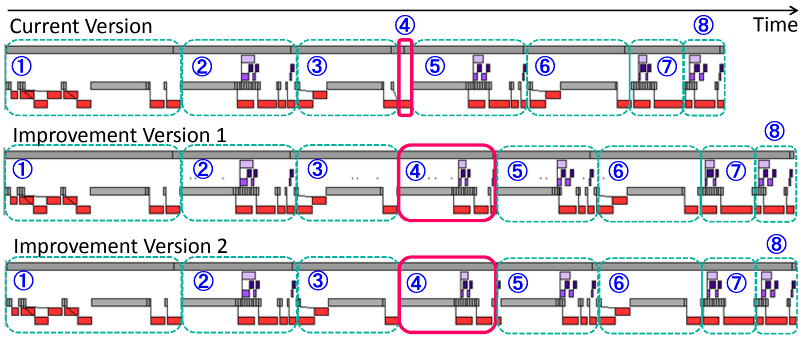


Fig. 9. Timeline of predictions in "Text box operations" task: Box ① - ⑧ represent procedural step. The step ④ causes the temporal difference of prediction time.

Evaluation Result of Modified User Interface Design

[Enter in text box]
The modified user interface designs made the procedure more easily understandable and its execution time was 1.5 seconds longer than the current version. Thus, we explained "the modified user interface designs were more understandable" and "the degradation of efficiency was 1.5 seconds" to clients and gave them quantitative metrics for UI selection.

[Item selection]
In the correct step, there is not much difference of prediction time between the current version and the improved version. In the recovery step, however, the times of the improved versions were much shorter than the current version. This is because users can look at selected items before finalizing them or go back to the last state if they make a mistake in the improved versions. It failed to consider the recovery step as it attached more importance to operational efficiency in the current version.

Because the time of the improved version was 0.1 – 0.7 seconds shorter than the current version in the correct step and was also 6 – 16 seconds shorter in the recovery steps, it was found that efficiency of the improved version was higher than the current version in not only correct operation steps but also recovery operation steps for operation mistakes.

5 Lessons Learned and Future Work

[IP phone]
Prediction time difference between the current version and the improved version showed the improvement effect of the modified user interface designs because they were made to improve efficiency.

[Electronic health record]

- We made sure that task execution time was effective as a metric that shows the improvement effect of the modified user interface design because it is easily understandable for clients.
- When the improvement is made for other than efficiency attributes (e.g.. understandability), it is difficult to "directly" verify the effectiveness of the modified user interface designs from an efficiency standpoint. For this case, however, the tool also gave clients data that is necessary for assessment of the improvement effect.

We obtained the following knowledge from clients or usability engineers.

- In this paper, the tool only shows that efficiency does not go down, and does not show the improvement effect on understandability directly. It would appear that to evaluate trade-offs between decreased efficiency and increased understandability, a tool that could predict understandability would be a useful addition to usability evaluations.
- Besides execution time, it would appear that the number of clicks with a mouse, necessity of shifting between mouse and keyboard, and distance of moving the cursor would be useful for usability evaluation. These could be computed using outputs of the tool. We plan to use these attributes in usability evaluation.

In this paper, we explained that “the modified user interface designs were more understandable” and that “the degradation of efficiency was 1.5 seconds” to clients and gave them quantitative metrics for UI selection. However, we could not evaluate the improvement effect on understandability directly. Since our results showed predictions of decreased efficiency with designs that we hoped improved understandability, a tool that can predict understandability would be useful. CogTool-Explorer [5] is a research version of CogTool that uses Information Foraging Theory [6] to predict understandability. In addition, we hope develop a technique that can estimate cognitive load or its impact on user performance and to expand the applicable scope of the tool using it. We are also planning to accelerate putting this tool in practical use by implementing it into a user interface development system.

References

1. Chi, E.H., et al.: The bloodhound project: automating discovery of web usability issues using the InfoScent simulator. In: Conference on Human Factors in Computing Systems, pp. 505–512 (2003)

2. Blackmon, M.H., et al.: Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 31–40 (2005)
3. Carroll, J.M.: HCI Models, Theories, and Frameworks: Toward a multidisciplinary science. Morgan Kaufmann Pub., San Francisco (2003)
4. John, B.E., et al.: Predictive human performance modeling made easy. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 455–462 (2004)
5. Teo, L., John, B.E.: Towards Predicting User Interaction with CogTool-Explorer. In: Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting, New York City, New York, September 22-26, pp. 950–954 (2008)
6. Pirolli, P.: Information Foraging Theory: Adaptive Interaction with Information. Oxford University Press, USA (2007)