Considerations of Efficiency and Mental Stress of Search Tasks on Websites by Blind Persons

Junichi Iizuka[1], Akira Okamoto[1], Yasuo Horiuchi[2], and Akira Ichikawa[3]

[1] Tsukuba University of Technology 4-12-7, Kasuga, Tsukuba City, Ibaraki, 305-8520 Japan
jiizuka@k.tsukuba-tech.ac.jp, okamoto@k.tsukuba-tech.ac.jp
[2] Chiba University 1-33, Yayoi-cho, Inage-ku, Chiba City, Chiba, 263-8522 Japan
hory@faculty.chiba-u.jp
[3] Waseda University 2-579-15, Mikasima, Tokorozawa City, Saitama 359-1192 Japan
a.ichikawa@aoni.waseda.jp

Abstract. We examined what kind of rating index was usable for verification of usability of websites for blind persons. The search time had a strong correlation with the NASA-TLX WWL scores. This would suggest a possibility to evaluate the usability by the search time. On the other hand, in respect of the accessibility check tool, its verification result had no correlation with NASA-TLX WWL scores, so it could not be used as a tool for verification of usability. We must develop a new usability check tool for blind persons. If we place functions with high usage frequency and high level of importance at a top of the website where a user can easily recognize them, it not only gives us a high level of visibility but also is effective for a blind person using voice output web browser as well.

1 Introduction

Many institutions including public institutions provide diverse information through Web sites. Therefore, it is important to make the information on Web sites accessible (assurance of accessibility) for the disabled and the aged persons [1] as well as for the non-disabled and young persons.

We also think that easy access to Web sites for disabled and aged persons is important. ISO9241-11 [2] defines "Usability" as "Level of efficacy, efficiency and user satisfaction of a product when the product is applied to achieve the specified target." However, guidelines and rating indices have not been clearly determined on consideration of the usability for the level of satisfaction required by a disabled person. Whether or not compliance to the accessibility guideline improves the usability has been little clarified.

In this study, we will discuss what kind of rating index is available for verification of the usability on Web sites for blind persons.

To begin with, when we define the usability as "Level of user satisfaction", we think that the level of satisfaction can be shown by the amount of mental workload to some extent. First, we measured the time to obtain information from Web sites. Second, we measured mental workload on Web sites by using NASA Task Load Index (NASA-TLX) [3] which is a typical measuring method of mental workload.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2009, LNCS 5616, pp. 693–700, 2009.

The reason why we measure the time is that we think of correlations between the required time and mental workload based on the following assumption; if the user can obtain information within a short time, the user's mental workload is assumed to be small, but if the user is slow to obtain the information, the user's mental workload is assumed to be large.Also, we examined if the accessibility check tool was effective for verification of usability or not.

As an experimental task, we selected "book search" on Web site of the library. The reasons for selecting this task are as follows. (1) It is a Web site which is required to be universally used by anyone including disabled person, (2) It contains many technical elements, which will affect the usability of the disabled persons. These technical elements include (a) browsing: various information is listed on the Web, (b) input operation: users must key-in the information such as keywords for search. and (c) tabular form output: the search result is displayed in tabular form.

In this study, we formulated the following hypotheses.

Hypothesis A: If the search time is correlated with mental workload, the length of search time indicates the degree of usability.

Hypothesis B: If the Hypothesis A is proved to be true and the search time is correlated with the verification result by the check tool, the check tool can be used for evaluation of usability.

2 Experimental Method

Subjects consist of 6 blind students of Tsukuba University of Technology and 6 sighted staffs. The blind students belong to the Department of Computer Science who are familiar with voice output web browser for personal computers. The sighted staffs are university officials using personal computers every day. In the experiment, the blind subjects used their own computers to which their familiar voice output web browser is installed. The sighted subjects used the common personal computers.

2.1 Verification with Check Tool

We surveyed compliance condition of 10 libraries web sites in Chiba City under Japanese Industrial Standard, JIS X8341-3 [4] by using the accessibility check tool.

JIS X8341-3 is the Japanese Standard that prescribe Web accessibility with the same purpose as WCAG (Web Content Accessibility Guidelines) 1.0 [5] We used "WebInspector ver.5.0" which is most common in Japan as an accessibility check tool [6]. From the front page to pages showing the details of books searched were verified. When the Web page was applying "frames", we verified HTML file that the "frame" was indicating. If the Web page was applying style sheets, we verified both pages of HTML files and CSS(Cascading style sheets) files.

2.2 Experiments of Book Search

In the beginning, the subjects conducted the book search of the other city's library and understood the experimental procedure. Then, they searched the same book at the library site of Chiba City. We defined the search time as time ranged from time point

when the front page of the library was shown to time point when the subjects could confirm display of the detailed information page of the book they searched. We divided the subjects into 2 groups and reversed the order of libraries to be searched for elimination of the order effect of the experimental result.

2.3 Evaluation of Mental Workload

- When the book search was finished, we asked the subjects to report the evaluated values of 6 rating subscales of NASA-TLX on a 10-point scale.
- We conducted the paired comparison among 6 rating subscales of NASA-TLX, and asked the subjects to weight the items. We multiplied the evaluation value in (1) by the weight obtained through the paired comparison, and then summed them up to calculate "NASA-TLX weighted workload(WWL) scores".
- We conducted interviews with them on impressions in the course of search operation.
- We repeated "2.2 Experiments of Book Search", on the next library site.

3 Experimental Result

3.1 Mental Workload and Search Time

NASA-TLX WWL scores show a condition that the more the number is, the heavier mental workload is. Fig 1 shows the average of NASA-TLX WWL scores of 6 blind subjects and the average search time of 6 blind subjects in each Web site. Figure 2 shows the average of NASA-TLX WWL scores of 6 sighted subjects and the average search time of 6 sighted subjects in each Web site.

Correlation coefficient between the average NASA-TLX WWL scores and the average search time of the blind subjects was r = 0.671. Correlation coefficient between the average NASA-TLX WWL scores and the search time was r = 0.751.

Based on the results, we found out a strong correlation between NASA-TLX WWL scores and the average search time in respect of blind subjects and sighted subjects.

Next, Figure 3 shows the relation between the search time and NASA-TLX WWL scores of all the subjects for 10 Web sites.

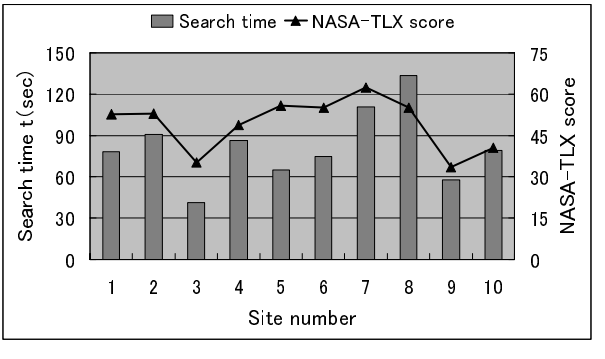

Fig. 1. The average NASA-TLX WWL scores and the average search time of blind subjects

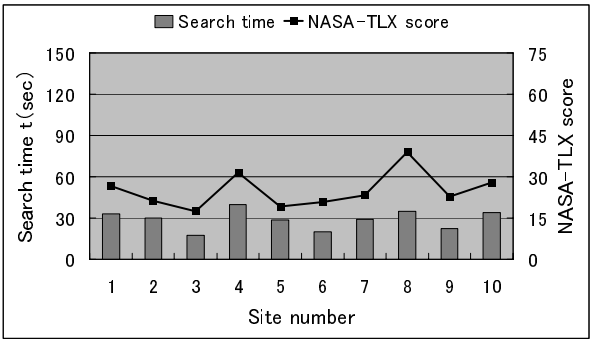

Fig. 2. The average NASA-TLX WWL scores and the average search time of the sighted subjects

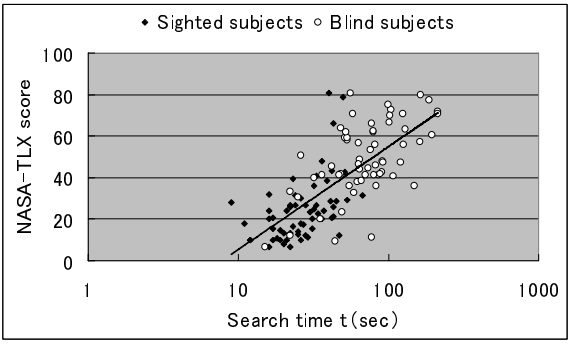

Fig. 3. The average search time of blind subjects and sighted subjects

The sighted subjects showed shorter search time and lower NASA-TLX WWL scores. On the other hand, blind subjects had longer search time and higher NASA-TLX WWL scores compared to sighted subjects. When we applied the collinear approximation to all the data, we obtained the following relational expression.

$$\text{NASA-TLX WWL score} = 21.5 \times \log(t) - 44.1 \quad (1)$$

3.2 Mental Workload and the Number of Errors on Web Sites

Figure 4 shows the average value of NASA-TLX WWL scores and the number of errors i.e., number of incompliant items, detected by the check tool of 6 blind subjects. Figure 5 shows the average value of NASA-TLX WWL scores and the number of errors detected by the check tool of 6 sighted subjects on each Web site.

The correlation coefficient between NASA-TLX WWL scores and the number of errors of blind subjects was $r = -0.077$. The correlation coefficient between NASA-TLX WWL scores and the number of errors of sighted subjects was $r = 0.083$.

This result clarified that there was little correlation between NASA-TLX WWL scores and the number of errors of blind subjects and sighted subjects.

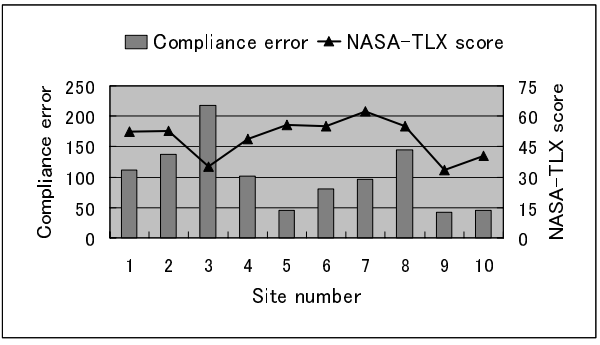

Fig. 4. The average NASA-TLX WWL scores and the number of errors of the library sites of blind subjects

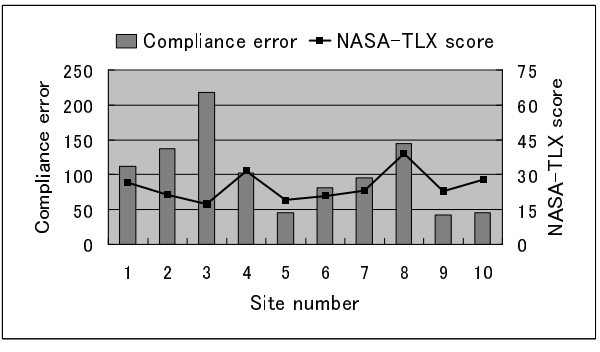

Fig. 5. The average NASA-TLX WWL scores and the number of errors of the library sites of sighted subjects

4 Discussion

We will verify the Hypothesis A formulated in "1. Introduction". As shown in Figure 1 and Figure 2, blind subjects and sighted subjects had strong correlations between NASA-TLX WWL scores and the search time, which supported the Hypothesis A In respect of a task of book search in this study, we obtained a result that the longer the search time was, the lower the usability was and that the shorter the search time was, the higher the usability was. This suggests that the search time can be an index to judge the quality of usability.

NASA-TLX defines 6 mental workload items and evaluates them in details. It requires time and efforts for execution. On the contrary, the measurement of the search time requires fewer loads on the side of the subjects, so it can be assumed to be a simple evaluation method.

Formula (1) shows the relationship between NASA-TLX WWL score and the search time. If we set a certain NASA-TLX WWL score for acceptable easiness of some book search task, we can convert it to the target value of the search time. In addition, if we assume the time required for one operation, we can estimate the number of operations

permitted for the search. However, NASA-TLX and the search time require measurement by the subjects.

As an easier method, we thought up the hypothesis b by focusing attention on if the accessibility check tool was available for evaluation of usability or not.

As shown in Figure 4 and Figure 5, there is little correlation between NASA-TLX WWL scores and the number of errors by the check tool in blind subjects and sighted subjects. In other worlds, we cannot evaluate the usability with the check tool. Therefore, we need to develop a new check tool to improve the usability for the disabled people.

Table 1 shows the summary of impressions on Web sites by blind subjects and sighted subjects.

Table 1. Easiness and difficulty in understanding in blind subjects and sighted subjects

		Blind subjects	
		Easy in understanding	Difficult in understanding
Sighted subjects	Difficult in understanding	(2) Measures for blind subjects had opposite effects on sighted subjects.	(1) Both the accessibility and usability are low.
	Easy in understanding	(3) Actuation of universal design.	(4) Accessibility for the blind subjects is low.

The category of easiness in understanding for blind subjects and sighted subjects (3) is applied with universal design. On the other hand, the category of difficulty in usage for both subjects (1) provides low accessibility and usability. The category of which is easily understandable for blind subject but hardly understandable for sighted subjects (2) includes pages mainly based on text without pictures or tables. The category which is understandable at a glance for sighted subjects but hardly usable for blind subjects using voice output web browser (4) includes pages mainly using pictures and tables.

Figure 6 shows the average search time of 6 blind subjects on the horizontal axis and the average search time of 6 sighted subjects on the longitudinal axis for each Web site. The dashed line shows the average search time of sighted subjects and blind subjects. Symbols and numbers in the Figure 6 show the number of library sites.

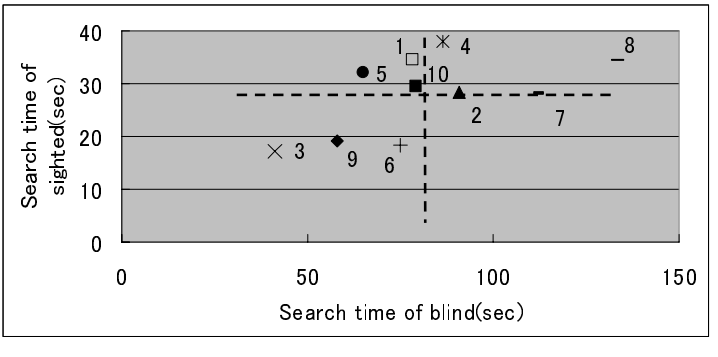

Fig. 6. The average search time of blind subjects and sighted subjects

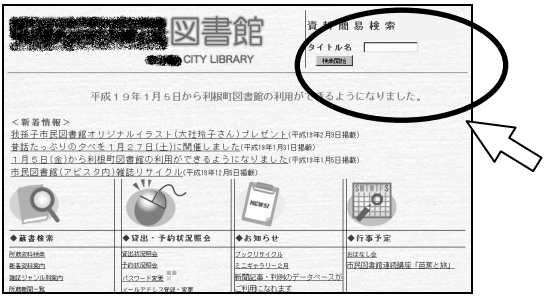

Fig. 7. A case example in which the book search is placed on the upper right of the page

A positional relation in this chart is the same as the one in Figure 1. In other words, the first quadrant is hardly understandable for blind subjects and sighted subjects, and the third quadrant is easily understandable for both subjects.

For example, the site No.3 (x mark in Figure 6) shows that the search time by both blind subjects and sighted subjects are below average. We confirmed the content of interview as well as reviewed the applicable Web pages to examine the reason. Figure 7 shows the front page of No.3.

As the entry field of book search is placed on the top line of the screen, sighted subjects responded that independent search function allowed them to view easily. Generally, voice output web browser will sequentially read from the upper left to the lower lines in the Web page. As voice output web browser read the field of book search right after display of this page, blind subjects could easily understand it.

Sighted subjects took the search time at average level for Sight No.7 (- mark in Figure 6), but blind subjects took the longer search time. Figure 8 shows the book title entry page. Three fields are shown in line to input key words. Sighted subjects can instantly understand the placement, but voice output web browser does not read three fields understandably. At the interview, we had responses from blind subjects that they had hard time understanding what to input in each field.

Based on the result above, we will think of the requirements for sites with high usability. It is effective for sighted users to decrease the amount of information within the page to an appropriate level and increase the level of legibility of the information which

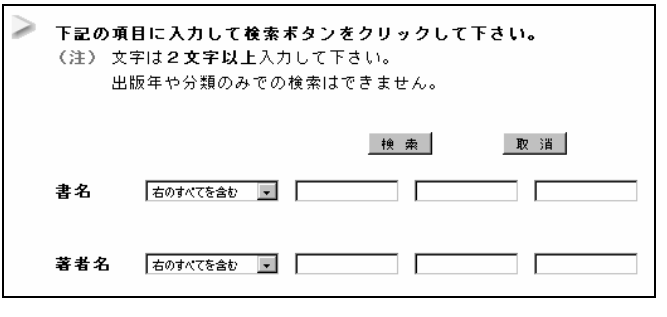

Fig. 8. A case example in which three entry forms are placed in line

is frequently used by the users. For blind subjects, it is effective to concern the accessibility for voice output web browser and appropriately arrange the sequence of reading. If these requirements are satisfied at the same time, we have Web site compliant to universal design.

5 Conclusion

The book search time has a strong correlation with NASA-TLX WWL scores, and the length of search time shows quality in usability. More specifically, the result indicates that we may evaluate the usability by the search time. On the other hand, the verification result of the accessibility check tool showed a weak correlation with NASA-TLX WWL scores, so it is not usable as a tool for verification of usability. Therefore, we need to newly develop a usability check tool for the disabled people.

If we place the book search function at a visible place in the top of Web page, it provides high level of legibility for all the users and accessibility for blind subjects using voice output web browser. In general, it is better to place functions with high frequency use and importance in this style.

In this study, we discussed a task of book search of the library sites, and we will further continue experiments with other tasks to clarify methods for improvement of usability for use of Web sites by blind persons.

Acknowledgement. I deeply appreciate for students and officials of Tsukuba University of Technology for your cooperation on this experiment. I am also grateful for Associate Professor Kazumitsu Shinohara at Osaka University for giving us valuable advice relating to NASA-TLX.

References

1. Harper, S., Yesilada, Y.: Web Accessibility: A Foundation for Research Human-Computer Interaction Series. Springer, New York (2008)
2. International Organization for Standardization, http://www.iso.ch/iso/home.htm
3. Hart, S.G., Staveland, L.E.: Development of a NASA-TLX (Task load index): Results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 139–183. North-Holland, Amsterdam (1988)
4. Japan Standards Association, http://www.jsa.or.jp/default_english.asp
5. Web Content Accessibility Guidelines 1.0. Web Accessibility Initiative, World Wide Web Consortium (May 1999), http://www.w3.org/TR/WAI-WEBCONTENT/
6. FUJITSU, Fujitsu Accessibility Assistance, Web Accessibility Inspector, http://www.fujitsu.com/global/accessibility/assistance/