Using Think Aloud Protocol with Blind Users: A Case for Inclusive Usability Evaluation Methods

Sambhavi Chandrashekar
Faculty of Information Studies
University of Toronto
Toronto, CANADA
+1 416 709 6270
sambhavi.chandrashekar@utoronto.ca

Tony Stockman
Department of Computer Science
Queen Mary University of London
London, U.K.
+44 (0) 207 882 5202
tonys@dcs.qmul.ac.uk

Deborah Fels
School of Information Technology Management
Ryerson University
Toronto, CANADA
+1 416 979 5000
dfels@ryerson.ca

Rachel Benedyk
UCL Interaction Centre
University College London
London, U.K.
+44 20 7679 5200
r.benedyk@ucl.ac.uk

ABSTRACT
There is a need to assess the applicability of conventional Usability Evaluation Methods to users with disabilities, given the growing importance of involving these users in the usability evaluation process. We found that conventional Think Aloud Protocol cannot be used as is, and will require modification to be useful, when evaluating websites with blind users.

Categories and Subject Descriptors
H.5.2 [User interfaces]: Evaluation/methodology, User-centred design, Theory and methods.

General Terms
Performance, Design, Human Factors.

Keywords

1. INTRODUCTION
Usability evaluation of systems forms an integral part of a user-centered design process. However, designers are generally not aware that individuals with physical, sensory or cognitive disabilities often fall into a system’s intended target user group. The majority of usability studies of computer systems have been carried out with able-bodied users [1]. Consequently, Usability Evaluation Methods (UEMs) that have evolved over the years tend to be designed for such users.

With increasing emphasis being placed on inclusion and universal usability, there is a growing trend to include users with disabilities in usability studies during the formative and summative design phases, and to assess conformance with disability-related regulations. Typical studies where users with disabilities are involved are (i) evaluation of website accessibility, and (ii) support for website usability guidelines addressing the needs of people with disabilities [3,4,8].

The fact that these users tend to have unique and different computer interactions compared with their able-bodied counterparts prompted us to examine whether their interactions with UEMs would also be different, specifically with Think Aloud Protocol (TAP). TAP is a verbal protocol method popularly used to gather usability data during system evaluation by asking the users to vocalize their thoughts, feelings and opinions concurrently while interacting with the system [5]. In this paper, we present our experiences and observations in using TAP during the evaluation of a website by blind users.

2. STUDY
As part of a broader study undertaken to analyze the distinction between accessibility guidelines and usability guidelines for inclusive website design [2], user testing was conducted on an experimental website. Six vision-impaired students were asked to evaluate a website specifically designed to assess the level of performance resulting from non-conformance vs. conformance to accessibility guidelines. Four of these users were blind and used the JAWS screen reader while the other two had low vision and used the Windows Magnifier. Users were asked to complete six tasks that were read out to them one at a time. The TAP method requires users to first read out a passage aloud to prepare themselves for talking out loud while performing their tasks later. This was found to be not possible for the four blind users as their method of reading text was to listen to it using a screen reader. Appropriate alternative training materials and methods to help them read out loud were not obvious to us and there seemed to be no literature outlining such materials or methods for TAP with blind users. TAP was still attempted for our study despite this limitation in training.

During the study, all users were prompted when they fell silent for more than about 15 seconds. However, most of the time, the blind users did not respond when prompted whereas the two low-vision users always responded when prompted. The sessions were audio recorded and the data was analyzed for the prime purpose of the study, i.e., assessing the effect of the non-conformance vs. conformance features of the website on user performance. These
data were reported previously [2]. The current focus is the efficacy of the TAP method.

3. DISCUSSION

First, we identified an important issue regarding TAP training for blind users. Since they used the screen reader to read text, getting them to read out loud from short passages at the outset, as required by the method, was not feasible. Merely explaining the TAP process appeared to be insufficient to get them geared up to talk aloud during the test session. There were no obvious alternative training strategies and materials that matched these users’ capabilities and skills.

The second important issue was that the blind users did not offer as many comments as the users with low vision, even when prompted regularly. Further, they expressed reluctance to repeatedly pause and restart the screen reader for articulating their comments as it interrupted their task performance. Coyne & Nielsen also reported a similar observation, that providing instructions before the session and offering prompts during the session did not seem to encourage blind users to talk often during computer-based tasks requiring screen reader use [4]. Possible reasons for this reluctance could be that there was competition for cognitive or perceptual resources between speech production for talk aloud and audio signal processing to understand the screen reader output, or that this particular set of users was not accustomed to talking aloud and were shy. However, during the interviews conducted after the tasks, they were able to comment on their browsing experience retrospectively.

Some researchers [1,4,8] have reported using TAP with blind users; Coyne and Nielsen [4] mentioned that people with disabilities require special considerations. However, none has studied the efficacy of TAP with blind users, nor offered any alternatives. There has been limited research on using TAP with users having other disabilities. For example, Roberts & Fels explored the use of the Think Aloud Protocol with deaf sign language users. They found that TAP could not be used in its conventional form, and devised a modified method called Gestural Think Aloud Protocol in order to include deaf sign language users [7].

Screen reader use seems to require a high cognitive effort from users as it involves hearing the page content in a sequential order, hearing the repeating information on each page (e.g., banners and users as it involves hearing the page content in a sequential order, as required by the method, was not feasible. Merely explaining the TAP process appeared to be insufficient to get them geared up to talk aloud during the test session. There were no obvious alternative training strategies and materials that matched these users’ capabilities and skills.

The second important issue was that the blind users did not offer as many comments as the users with low vision, even when prompted regularly. Further, they expressed reluctance to repeatedly pause and restart the screen reader for articulating their comments as it interrupted their task performance. Coyne & Nielsen also reported a similar observation, that providing instructions before the session and offering prompts during the session did not seem to encourage blind users to talk often during computer-based tasks requiring screen reader use [4]. Possible reasons for this reluctance could be that there was competition for cognitive or perceptual resources between speech production for talk aloud and audio signal processing to understand the screen reader output, or that this particular set of users was not accustomed to talking aloud and were shy. However, during the interviews conducted after the tasks, they were able to comment on their browsing experience retrospectively.

Some researchers [1,4,8] have reported using TAP with blind users; Coyne and Nielsen [4] mentioned that people with disabilities require special considerations. However, none has studied the efficacy of TAP with blind users, nor offered any alternatives. There has been limited research on using TAP with users having other disabilities. For example, Roberts & Fels explored the use of the Think Aloud Protocol with deaf sign language users. They found that TAP could not be used in its conventional form, and devised a modified method called Gestural Think Aloud Protocol in order to include deaf sign language users [7].

Screen reader use seems to require a high cognitive effort from users as it involves hearing the page content in a sequential order, hearing the repeating information on each page (e.g., banners and copyright information) and processing information in small portions as delivered by the screen reader. Users must remember the whole page by forming connections in their mind between these small portions. The synthetic speech of the screen reader also lacks prosodic cues (emphasis and rhythm), which are important for parsing of speech. All this puts additional pressure on the cognitive resources of the users affecting their ability to simultaneously accomplish any other tasks that require thinking.

The reluctance of these users to comment during the study could also be viewed in the light of Wickens’ multiple resource model [9], which purports that humans can divide attention between the visual and auditory channels better than between two visual or two auditory channels. To use TAP, blind users must use two auditory channels - one for listening to screen reader output and one for talking aloud. Rapid shifts of attention within the same modality may cause problems due to sensory overload resulting in confusion and potential discomfort. This may impede their ability to formulate and vocalize comments while continuously hearing the screen reader, causing a reluctance to comment.

Retrospective data may not be as effective as concurrent data in revealing the human decision making process [6]. However, to supplement, or even substitute, the procedure of collecting concurrent verbal protocols, which does not appear to work optimally in the case of blind users, retrospective verbal protocols could be collected immediately after completion of each task. The relative effectiveness of these two procedures as well as their combined effectiveness as a multimethod strategy needs further study.

4. CONCLUSION

It appears that TAP, in its popular form as a concurrent verbal protocol method, may not be effective for use with blind persons using a screen reader to access websites. Further research is required to determine how best to modify this protocol for use with these users.

5. REFERENCES


