Cogn Process (2010) 11:285–288
DOI 10.1007/s10339-010-0365-9

LETTER TO THE EDITOR

Usability evaluation with screen reader users: a video presentation of the PCTA's experimental setting and rules

Stefano Federici · Simone Borsci · Maria Laura Mele

Received: 3 June 2010 / Accepted: 9 June 2010 / Published online: 14 July 2010

Abstract In the study entitled "Web usability evaluation with screen reader users: Implementation of the Partial Concurrent Thinking Aloud technique" (Federici et al. 2010), we have proposed a modified protocol of usability evaluation technique for blind users, which integrates the features of the concurrent and the retrospective techniques. This new technique, called partial concurrent thinking aloud (PCTA), while respecting the properties of classic verbal protocols, overcomes the structural interference and the limits of concurrent and retrospective protocols when used with screen reader users. In order to facilitate understanding and acquisition of the PCTA's technique for practitioners and researchers, we have video recorded three different verbal protocols by visualizing five experimental sections. In the first two videos, we have compared a concurrent with a retrospective's verbal protocol of a sighted user, showing the difference of the verbalizations provided by the user in these two conditions. The third video shows the structural interference of the screen reader, during a blind user concurrent thinking aloud. In the last two videos, we show the difference of a blind user behaviour when PCTA or retrospective protocol is adopted. The videos clearly visualize the advantage of the PCTA use in respect of the two other protocols. In conclusion, the visualization of the PCTA technique confirms that this new verbal protocol promotes and guarantees a more user-driven usability assessment with disabled people, by better involving screen reader users, overcoming the structural interference and the limits of the concurrent and retrospective protocols.

Keywords Concurrent think aloud · Integrated model of usability evaluation · Partial concurrent think aloud · PCTA · Retrospective think aloud · Usability evaluation · Verbal Protocol

Introduction

In the usability evaluation, both retrospective and concurrent thinking aloud technique (TAP) could be used according to the study aims and goals. Nevertheless, when a usability evaluation is carried out with blind people, several studies propose to use the retrospective TAP: indeed, using a screen reader and talking about the way of interacting with the computer imply a structural interference between action and verbalization (Guan et al. 2006; Takagi et al. 2007). Indeed, as Strain et al. (2007) have noticed, the use of a screen reader "interferes with any dialogue between moderator and participant. Perceptual studies have shown that it is possible for humans to deal with two voices at once (the so-called cocktail party effect); however, due to cognitive limitations, people often have a difficult time talking and listening at the same time". These authors are referring to the Kemper et al. (2003) study about the costs of doing two things at once for adults. Undoubtedly, basic cognitive studies provided a lot

Electronic supplementary material The online version of this article (doi:10.1007/s10339-010-0365-9) contains supplementary material, which is available to authorized users.

S. Federici
Department of Human and Education Sciences,
University of Perugia, Perugia, Italy

S. Federici · S. Borsci (✉) · M. L. Mele
ECoNA, Interuniversity Centre for Research on Cognitive Processing in Natural and Artificial Systems,
Sapienza University of Rome, Rome, Italy
e-mail: simone.borsci@gmail.com

of evidence supporting the idea that individuals can listen, verbalize, or manipulate, and rescue information in multiple-task condition. As Cherry (1953) showed, subjects, when listening to two different messages from a single loudspeaker, can separate sounds from a background noise, recognize the gender of the speaker, the direction, and the pitch (cocktail party effect). At the same time, subjects that must verbalize the content of a message (attended message) listening to two different messages simultaneously (attended and unattended message) have a reduced ability to report the content of the attended massage, while they are unable to report the content of the unattended message. Moreover, Ericsson and Kintsch (1995) showed that, in a multiple-task condition, subjects' ability of rescuing information is not compromised by an interruption of the action flow (as it happens in the concurrent thinking aloud technique) thanks to the "Long Term Working Memory mechanism" of information retrieval. Even if users can listen, recognize, and verbalize multiple messages in a multiple-task condition and they can stop and restart actions without losing any information, other cognitive studies (Cherry 1953) underlined that the overlap of activities in a multiple-task condition have an effect on the goal achievement. Kemper et al. (2003), analysing the users' abilities to verbalize actions in a multiple-task condition, showed that the fluency of a user's conversation is influenced by the overlap of actions. We do not know how and how much the content of verbalizations could be influenced by the strategy of verbalization (i.e. the modification of fluency and the complexity in a multiple-task condition). Anyway, we well know that users in the concurrent thinking aloud verbalize the problems in a more accurate and pertinent way (i.e. more focused on the problems directly perceived during the interaction) than in the retrospective one (Hoc and Leplat 1983; Bowers and Snyder 1990; Hannu and Pallab 2000). The pertinence is granted to the user by the proximity of action-verbalization-next action; this multiple-task proximity compels the subject to apply a strategy of verbalization that reduces the overload of the working memory. However, for blind users, this time proximity between action and verbalization is lost: the use of the screen reader, indeed, increases the time for verbalization (i.e. in order to verbalize, blind users must first stop the screen reader and then restart it). The use of retrospective TAP with disabled users remains only a functional solution, for two main reasons: (i) First, it permits to overcome the user's cognitive limitations, but it fails to analyse the user's performance during an interaction, as the concurrent TAP does; (ii) Second, since the efficiency of concurrent technique greatly decreases when used with blind people in comparison with sighted users, practitioners prefer to use the retrospective model over the concurrent, even though, in this way, the number of verbalizations remarkably decreases. Therefore, as a lot of studies show, the retrospective technique provides a less effectiveness of data and a less pertinent users' verbalization (Bowers and Snyder 1990; Ericsson and Simon 1993; Borsci and Federici 2009). In order to reduce the screen reader influence (structural interference) without losing the advantages of the proximity within action, thinking, and identification of the problems (pertinence of users' verbalization), we have improved a new TAP technique, called Partial Concurrent Thinking Aloud (PCTA) (Borsci and Federici 2009; Federici et al. 2010), by unifying the advantages of both concurrent and retrospective models.

Properties of PCTA

Our aim is to build up a usability assessment technique eligible to maintain the advantages of concurrent and retrospective protocols while overcoming their limits. Therefore, we have analysed the PCTA technique's efficiency with both blind and sighted users. In order to do so, we split the PCTA method into two sections, one concurrent and one retrospective. (1) The first section is a modified concurrent protocol built up according to the three concurrent verbal protocols criteria described by Ericsson and Simon (1993). (i) The first criterion is *Subjects should be talking about the task at hand, not about an unrelated issue.* In order to respect this rule, the time between problem retrieval, thinking, and verbalization must be minimized to avoid the influence of a long perceptual reworking and the consequent verbalization of unrelated issues. Blind participants, using a screen reader, increase the time latency between identification and verbalization of a problem. To minimize this latency, users are trained to ring a desk-bell that stops both time and navigation by creating a memory sign (i.e. ring the bell). Then, they restart immediately the navigation; (ii) the second criterion is *To be pertinent, verbalizations should be logically consistent with the verbalizations that just preceded them.* For any kind of user, it is hard to be pertinent and consistent in a concurrent verbal protocol. Therefore, the practitioners could generally interrupt the navigation and ask for a clarification or stimulate the users to verbalize in a pertinent way. In order to do so and stop navigation to screen reader users, we propose to negotiate a specific physical sign with them: The practitioner, sitting behind the user, will put his hand on the user's shoulder; (iii) the third criterion is *A subset of the information needed during the task performance should be remembered*. The concurrent model is based on the link between working memory and time latency. The proximity between the occurrence of a thought and its verbal report allows users to verbalize on the basis of their working memory. (2) The second PCTA

section is a retrospective one in which users analyse those problems previously verbalized in a concurrent way. The memory signs, created by users ringing the desk-bell, overcome the limits of classic retrospective analysis; indeed, these signs allow the users to be pertinent and consistent with their concurrent verbalization, thus avoiding the influence of long-term memory and perceptual reworking. According to our new verbal protocol technique, the PTCA evaluation is based on a threefold structure (Borsci and Federici 2009; Federici et al. 2010):

1. In order to minimize proximity between action, thoughts, and verbalization, visual impaired users interrupt the navigation ringing a desk-bell next to the mouse (i.e. memory sign);
2. Practitioners can touch users' shoulders with a hand as a physical sign (negotiated during training) in order to interrupt the navigation and ask about the action performed;
3. The retrospective session analysis is focused on those memory signs created during the concurrent session analysis.

The experimental videos

Video 1 and 2: A sighted user verbal protocol: Concurrent vs. Retrospective:

In the concurrent thinking aloud (video 1), users express their problems, strategies, stress, and impressions without the influence of a "rethinking" perception. The influence of the long-term memory (LTM) is minimized because the user spends just 5 s from the thinking to the verbalization. Therefore, the user's verbalizations are more pertinent (focused on navigation problems). The user's focus attention is more on the action and the influence of user's past experience on the verbalization is minimized: the working memory (WM) is mainly involved in the process of verbalization.

In the retrospective thinking aloud (video 2), by using the LTM and making a cognitive reconstruction of their previous navigation experience with system, user tells a story of their actions, strategies, and navigation problems. In the retrospective thinking aloud, the interaction experience is stored by users in the LTM. When the user watches the recorded video, their judgments, experiences, and expertise are mixed with the interaction experience. The video clearly shows that the facilitator has to help user in his reconstruction more than in the concurrent verbalization. The verbalizations of the problems are more influenced by user's opinions and judgments of the entire navigation of the product. For example: If a user finds the system unsatisfactory, the verbalizations might suffer of this user's judgment. The user's verbalizations are longer than those provided in the concurrent analysis since user has already made a "judgment" about the interface (he tells a story well knowing its end).

Video 3: The structural interference in a screen reader user thinking aloud

In the third video, a screen reader user tries to engage a concurrent thinking aloud; he is able to analyse the interface speaking during the interaction and stopping the screen reader. However, the video shows that this multiple-tasks condition has a very bad consequence for the user navigation, in fact: (i) he has to go over a link or a text repeatedly in order to listen again the screen reader when he verbalizes during the interaction; (ii) in the multiple-tasks condition, user suffers of an heavy cognitive overload, which decreases his verbalizations.

Video 4 and 5: a screen reader user verbal protocol: The retrospective and the PCTA

The video of the retrospective analysis (video 4) shows that the verbalizations of the screen reader user are comparable to those provided by the sighted user. The screen reader user tells a story, and the facilitator has to help him in order to obtain a pertinent verbalization.

In the PCTA (video 5), by using the memory signs, user can avoid the LTM influence in the information rescue. Those signs allow user to have a pertinent analysis of screen readers' action (i.e. the interaction) and to verbalize in a pertinent way. When user, focusing his attention on the recorded screen reader, listens to the memory signs, he is alerted that in certain point there is a problem. Using the memory sign, user can rescue in the LTM, by minimizing the influence of user experience, which kind of problems he has found during the navigation. Indeed, the user can avoid re-thinking of interaction by using the sign as a memory rescue tool. These memory signs are created by user during the concurrent step for himself in the retrospective step. There is a difference for the user between watch a video and recognize where they have a problem and which kind of problem they have found (as in the retrospective protocol) and in watch a video with signs that indicate where are the problems. As in the concurrent protocol in the PCTA, users can indicate immediately where they find a problem, by speaking about the specific problem indicated by the sign. The video clearly shows that the verbalizations provided by the user in the PCTA are pertinent and short as it happen in the concurrent verbalization of a sighted user.

Acknowledgments We thank Dr Giammario Mascolo—psychotechnologist and assistant chair of Psychotechnologies for Adaptation and Integration at "Sapienza" University of Rome—and Domenico Polimeno—web and software developer, expert in the usability evaluation and design research—for their participation to the video realization.

References

Borsci S, Federici S (2009) The partial concurrent thinking aloud: a new usability evaluation technique for blind users. In: Emiliani PL, Burzagli L, Como A, Gabbanini F, Salminen A-L (eds) Assistive technology from adapted equipment to inclusive environments—AAATE 2009. IOS Press, Amsterdam, pp 421–425

Bowers VA, Snyder HL (1990) Concurrent versus retrospective verbal protocols for comparing window usability. Human Factors Society 34th Meeting, 8–12 October 1990 HFES, Santa Monica, pp 1270–1274

Cherry EC (1953) Some experiments on the recognition of speech, with one and with two ears. J Acoust Soc Am 25:975–979

Ericsson KA, Kintsch W (1995) Long-term working memory. Psychol Rev 102:211–245

Ericsson KA, Simon HA (1993) Protocol analysis: verbal reports as data, Revised edn. MIT Press, Cambridge

Federici S, Borsci S, Stamerra G (2010) Web usability evaluation with screen reader users: Implementation of the Partial Concurrent Thinking Aloud technique, Cognitive processing 11 (3). doi: 10.1007/s10339-009-0347-y

Guan Z, Lee S, Cuddihy E, Ramey J (2006) The validity of the stimulated retrospective think-aloud method as measured by eye tracking. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 1253–1262

Hannu K, Pallab P (2000) A comparison of concurrent and retrospective verbal protocol analysis. Am J Psychol 113:387–404

Hoc JM, Leplat J (1983) Evaluation of different modalities of verbalization in a sorting task. Int J Man-Mach Stud 18:283–306

Kemper S, Herman RE, Lian CHT (2003) The costs of doing two things at once for young and older adults: talking while walking, finger tapping, and ignoring speech or noise. Psychol Aging 18:181–192

Strain P, Shaikh AD, Boardman R (2007) Thinking but not seeing: think-aloud for non-sighted users. CHI '07 extended abstracts on Human factors in computing systems, ACM, New York

Takagi H, Saito S, Fukuda K, Asakawa C (2007) Analysis of navigability of web applications for improving blind usability. Comput-Hum Interact 14:13–37