Working Paper

The disruptive effect of Think Aloud

By

Janni Nielsen
Carsten Yssing

No. 20-2004

Institut for Informatik

Handelshøjskolen
i København

Howitzvej 60
2000 Frederiksberg

Tlf.: 3815 2400
Fax: 3815 2401
http://www.inf.cbs.dk

Department of Informatics

Copenhagen
Business School

Howitzvej 60
DK-2000 Frederiksberg
Denmark

Tel.: +45 3815 2400
Fax: +45 3815 2401
http://www.inf.cbs.dk

The disruptive effect of Think Aloud

Janni Nielsen
Department of Informatics
Copenhagen Business School
janni.nielsen@cbs.dk

Carsten Yssing
Department of Informatics
Copenhagen Business School
cy.inf @cbs.dk

ABSTRACT
Thinking Aloud is the most commonly used technique used to test users´ interaction with computers. The assumption is that Think Aloud gives access to what goes on in the users´ minds. However, interfaces are multi modal and play heavily on user´s visual perception. Reflecting upon Think Aloud (TA), we ask the question: what happens when users are required to verbalise their visual perceptions and interactions? We argue that TA may have a disruptive effect, suggesting that other techniques be considered. With a theoretical distinction between focal and subsidiary awareness and a focus on the sense making process, we develop a frame for test of user´s visual interaction which rely on the coordination between hand/mouse and eye/cursor.

Author Keywords: Think Aloud, visual perception, interaction, test

Introduction
In HCI practice as well as in research Thinking Aloud (TA) stands out as unique. It is popular [1] and often referred to as the usability method. In a survey of methods and techniques used by Danish HCI practitioners (75% of respondents) and researchers (25% of respondents) TA came out as the single most frequently applied technique [4]. This should not come as a surprise as the technique is included in the HCI curriculum taught at universities [5]. TA is tempting because only few users are needed, it may be used by non-specialists and it promises access to people´s minds. Besides, Jakob Nielsen [10] has tirelessly promoted TA and argued for its cost effective benefits.

The understanding beneath the use of TA in usability testing is that the technique gives access to mental behaviour [11], and insight into cognitive activities ".. that may not be visible at all" [8]. Branch [2] has argued that TA provides "the most complete and detailed description of the information-seeking processes …". However, she argues that concurrent verbalisation is problematic "… when the information is difficult to verbalise because of its form ..". Karsenty [9] suggests that TA puts a cognitive load on the user, requiring a cognitive involvement that may interfere or compete with the cognitive requirements of the task, and Preece [13] cautions us about the added strain on users. But what do we get when we ask people to think aloud?

Thinking Aloud
Ericsson and Simon [6] discuss the use of introspective data in the study of task directed cognitive behaviour. Their understanding is that most performance measure rely on some kind of verbal data and they argue that a sentence is the verbal realisation of thought. They assume we can verbalise what we are learning while in the process of learning and we can verbalise what we know if questioned shortly after the process has taken place. Because it is still retained in our Short Term Memory (STM). However, if there is a time span before recall, we have to rely on our Long Term Memory, and we will produce descriptions and explanations – not report our immediate thoughts.

In their classic text on protocol analysis from 1984, they distinguish between three kinds of cognitive processes: 1) *Talk aloud* is a vocalisation of thoughts which are already coded in verbal form (internal speech), and accessible in STM, 2) *Think aloud* is verbalisations of thoughts also held in STM, but coded in other forms, e.g. visually and 3) *Retrospective reports* are verbalisations of thoughts not held in STM, e.g. descriptions and explanations. Retrospection is similar to thinking aloud, but more error prone in comparison to what the user actually did and saw because it is descriptions, or recall from long term memory.

Percept and concept
But how do users´ experience TA? In the research literature there seems to be little interest in the way the users experience the technique. Teaching graduate students TA techniques, they were asked to focus on how they experienced the technique. The students reports revealed a number of problems. They experienced that

- they thought faster than they could speak
- thought processes were much more complex than they could verbalise
- TA interfered with the interaction and the task and
- TA did not come naturally

These findings suggest there is "a problem in assuming that performance measures have to rely on some kind of verbal data, and that a sentence is the verbal realisation of thought"[12]. Some performances are beyond words, though they may be observed, and even registered. Besides, thoughts are not mainly verbal and directly accessible in oral speech, and there is not a 1:1 relationship between the thoughts, the actions and the words spoken. Furthermore the user is interacting with

multi modal interfaces: colours, layout, sound, graphics, animations – and visualisation is the essential feature. This requires a mental interaction which is based in visual perception and it comes into being through an act of sense giving [3). However, TA requires that attention shifts focus from giving sense to that which is perceived and does not exist as concept to constructing sentences or words and expressing them aloud,. Perceptions and actions must be transformed to talk, and even if the speech is immediate and run concurrently with the thoughts – users attention has to shift focus from understanding to verbalisation. The process of sense making is interrupted because attention keeps changing object, and verbal overshadowing may be the result. [15]

Visual perception is a sense making process and verbalisation of thoughts may have a disruptive effect on the interaction. When the students report that they think more than they can verbalise, that they think faster than they can vocalise – it may be because they sense these extremely complex mental process as almost instant mental processes. The students do not describe the cognitive process taking place – but the way in which they experience it.

Mental interaction and visual perception

Interfaces are visual and dynamic and visual perception is the basis for the user´s interaction. It seemed essential to us that any test of the interface should be able to capture the visual interaction, without being interrupted by verbal and with this understanding we embarked upon experimenting with test techniques. Point of departure for our experimental work was the observation of how users´ eyes follow the cursor movements on the screen, never turning attention to focus on the hand controlling the cursor. We decided to use this hand-eye/mouse-cursor coordination and explored it in different steps.We shall return to the steps after having introduced the theoretical underpinning for this understanding and our approach.

Polanyi [14] makes a distinction between focal and subsidiary awareness. When a blind person uses a stick to feel her way through a space her focus is not on the stick – nor on the end of the stick, but on the meeting of the stick with objects or surfaces. This extension of the senses outwards away from oneself and into the world we assume also to be the case in the interaction of hand and eye/mouse and cursor. The senses are projected outwards and onto the meeting of the eye with the cursor on the screen. This projection is tied to Polanyi understanding of the sense-making process - as when a person is reading a text - the reader does not focus on each letter, or on each word, nor on the sentence as a whole. These all serve as subsidiaries for her focal awareness which is projected outward, towards meaning. This understanding of the sense making process, we assume also to be the case of the interaction with visual dynamic interfaces, hence the visual perception is the sense-making.

Experimental work with visual tests

The aim was to develop a test for visual interaction with web-sites, and this meant including a graphical dimension, a use dimension and the dimension of embedded communication in site. Because evaluation of websites should not be based in functionality alone nor in aesthetics and visual design. Hence the test should include functionality, aesthetics and the imbedded communication in the site as to target group and the object of the site as well as structure. The following keywords served as guidelines for the design of tests:

- *User-oriented*, with functionality tested by specifying different information retrieval task
- *Aesthetics-oriented*, with focus on the visual image and that which captures user´s eye: the visual perception
- *Imbedded design communication*, with focus on the intended target group for the web site, the structure and the goal with the design

The users were graduate students from Copenhagen Business School participating in an full HCI course (450 student hours). Their background is in economics, system development and organisational development. The test were developed over a number of courses and all together approximately 150 students have been tested over a period of 3 years.

We started out with a very simple pen and paper test where the students were asked to let the pen follow their eyes roaming around in a picture. This experiment revealed a number of problems. It was impossible to see where the roaming of the picture started and ended and impossible to say anything about the reading directions. Neither was it possible to see if students eyes rested in different places in the pictures. Besides, the pen tended to get in the way of the eyes looking.

A part solution to the problem was to design the test to that the image was projected on the wall and students hand and pencil would follow the eyes by drawing on a transparency. This technique also allowed us to make comparisons between the way their eyes roamed the images, by simply putting the transparencies on top of each other. The solution to the problem with reading direction became instructing the students to mark starting point with a square and end point with an arrow. The last problem was the question of whether the eyes would come to rest in the picture at different points was to ask students to draw a circle whenever their eyes rested in the image, and to number these circles after the test.

The result of this test is in the process of being analysed, and it looks interesting. However, our attempt to get beyond the disruptive effect of TA – only resulted in similar problem as with TA. The user had to shift attention from letting the eyes roam the image to paying attention to when the eyes rested in image and to focus on drawing circles and retrospectively number them Hence the solution did not solve our problem with TA, but it did point us in directions which seem promising, and we eventually designed the test to be run on computers and with automatic logging of cursor movements hence getting beyond the problems of verbal interruption.

The following is an example of the instruction to the student. The instructions is introduced in steps, one after the other, when students stop action:

Step 1: "You work as a web-architect in a design company, and have a day where it is a little quiet at work. You decide to use the time and satisfy your curiosity about what is new within graphical design on the web. You enter http://www.art-in-pixel.com. What captures you? What do you see?

Step 2: art-in-pixel has designed the CD cover for EMI and Class 95, the Christmas collection, and you would like to take a closer look.

Step 3: Who is the target group for the site? What is the goal of the designer with the site? What is the structure of the site?

Testing Visual Interaction

Just as TA has to be introduced to the users, and the users have to go through a learning process before valid empirical data can be collected, so do test of visual interaction We are in the middle of the analytical work but so far we have identified the following steps for testing visual interaction. The sequence in which they are introduced to the user is constructed to allow for a learning process:

- **direct coordination eye-hand:** visual reading with pen on transparency on top of image
- **semi-indirect coordination eye-hand:** visual reading of image projected on wall: hand and pen on transparency
- **indirect coordination eye-hand:** visual reading/interaction with computer interface and recording of cursor movement and screen.

In the final step of the test a *Retrospective recall* is collected. As specified above in the indirect coordination eye-hand, the visual interaction is being recorded. This recording is then replayed with a simultaneous dialogue, between user and researcher is being carried out with voice over on a new recording. The replay recording can be stopped and started as the dialogue unfolds, placing a minimum stress on the user.

From TA to Visual Interaction and retrospective recall

With the indirect coordination eye-hand techniques we suggest that we succeed in visual capture of visual interaction, that is capturing concurrent visual perception which takes us beyond the disruptive effect of verbalisation.

With the indirect coordination eye-hand followed by retrospective recall we also move beyond some of the problems of retrospective explanations and description. Because it is not a free recall, but is controlled by the actual sequence recorded, we do not rely directly on user´s nor researcher´s memory. Hence the technique may take us beyond the problems of descriptions and explanations.

This seems in accordance with the understanding of Ericsson (2002), "The least reactive method to access participants ´memory of an experience is to instruct them to give a retrospective report (Ericsson and Simon 1984/1993, p. 378). The participants are instructed to recall the sequence of their thoughts and experiences during the target event and to report those thoughts/experiences that they can definitely recall, thus avoiding guessing or reporting thoughts that they think they might have had ….. In fact, the process of recalling one´s thoughts appears to strengthen the overall memory … ". [7]. However, in using retrospective recalls to retrieve data from memory one must be cautions, because the requirement to describe and explain actions may change the recall of what actually happened. A related problem, as also pointed out by Ericssooon, is when participants are asked to verbalise more on their performance than what they spontaneously recall. Ericsson´s solution to the problems are to develop standardised non-reactive procedures on the basis of instructions to retrospective reports (p. 987).

This may be development strategy within the information processing framework. However, in an HCI perspective and with focus on web-interfaces, the interest is not only to eliminate errors and noise, or to secure as accurate recall as possible. It is also interesting how users think about the website, after the interaction. User recall should not be investigated only as concurrent accurate memory, but also in retrospective, as a learning opportunity and a aesthetic experience. However, it is important to distinguish between the two. Besides, the retrospective recall is a possibility for dialogue where it is, in fact, possible to capture how users reason about their perception of an interface.

References

1. Boren, M. T. & Ramey, J. (2000), "Thinking aloud: Reconciling theory and practice", *IEEE Transactions* 43, no. 3, pp. 261-278, 2000

2. Branch, J. L. (2000), Investigating the information-seeking processes of adolescents, *Lib. & Inf. Science Research,* vol. 22,4, pp. 371-392.

3. Bruner, Jerome (1990) Acts of Meaning, Harvard University Press, Cambridge, 1990

4. Clemmensen, T. & Leisner, P., (2002), Community knowledge in an emerging online professional community, IRIS25 Denmark

5. Dix, A., Finlay, J., Abowd, G. & Beale, R., *Human-Computer Interaction*, Prentice Hall, 1997

6 Ericsson, K. A. & Simon, H. A.(1984), *Protocol Analysis. Verbal reports as data*, Cambridge, Massachusetts, 1984.

7. Ericsson, K. Anders, Applied Cognitive Psychology, 16: 981–987 (2002), Towards a Procedure for Eliciting Verbal Expression of Non-verbal Experience without Reactivity: Interpreting the Verbal Overshadowing Effect within the Theoretical Framework for Protocol Analysis

8. Hackos, J. T. & Redish, J. C.(1998), *User and Task Analysis for interface Design*, Wiley, USA, 1998

9. Karsenty, L. (2001), Adapting verbal protocol methods to investigate speech systems use, *Applied Ergonomics,* vol. 32, no. 1, pp. 15-22.

10. Nielsen, J.(1994), "Estimating the Number of Subjects Needed for a Thinking Aloud Test", *Int. J. of Human-Computer Studies,* **41**, 3, pp. 385-397

11. Nielsen Janni, Clemmensen T., Yssing C. (2002), Getting access to what goes on in people's heads? NordiCHI, proceedings. Denmark, p. 101-110

12. Nielsen Janni and Carsten Yssing (2003) Getting beyond the disruptive effect of Think Aloud, Proceedings of the Third Danish Human-Computer Interaction Research >Symposium, 67-70

13. Preece, J., Rogers, Y. & Sharp, H.(2002), *Interaction Design: Beyond Human-Computer Interaction*, John Wiley & Son, 2002.

14. Polanyi, Michael (1967) Logic and Psychology*, American Psychologist 23*, p. 27-43

15. Schooler Jonathan W., Stellan Uhlsson and Kevin Brooks (1993), Thoughts Beyond Words. *J, of Exp. Psychology*, 1993, Vol 122, 2, 166-183