Metaphors of Human Thinking: A New Tool in User Interface Design and Evaluation

Kasper Hornbæk & **Erik Frøkjær**

University of Copenhagen, Denmark

KHornbaek@nik.ku.dk, erikf@diku.dk

Abstract: Understanding human thinking is crucial in the design and evaluation of human-computer interaction. Inspired by introspective psychology, we present five metaphors of human thinking. The aim of the metaphors is to help designers consider important traits of human thinking. We illustrate how to use the metaphors for usability evaluation and how good and poor user interfaces can be appreciated in terms of the metaphors. An experiment with 87 subjects show that usability evaluation by metaphors of human thinking compared to heuristic evaluation uncovers usability problems that are assessed as more severe on users and more complex to repair.

Keywords: Human thinking, inspection technique, usability evaluation method, heuristic evaluation, quantitative study

1 Introduction

We present five metaphors of human thinking that have been highly useful in our research and teaching of human-computer interaction (HCI). The metaphors build upon the classical book 'Principles of Psychology' by William James (1890), which by Peter Naur (1995) has been made more easily accessible to HCI professionals. Each metaphor has been selected to capture aspects of thinking important to successful human-computer interaction, aspects often forgotten or not respected properly.

Metaphors in the HCI literature have been used in describing certain styles of interfaces, e.g. the desktop metaphor, and as a vehicle for representing and developing designs of interfaces. Our use of the term metaphor is different, in that the metaphors are not in any way intended as interface metaphors, nor do we imagine the metaphors to form part of designs. Rather, the aim of the metaphors is to support the evaluator/systems designer in a focused study of how well certain important aspects of human thinking are taken into account in the user interface (UI) under evaluation/design. The metaphors are intended to stimulate, generate insight, and break fixed conceptions. These uses of metaphors have been thoroughly studied in the literature on creative thinking (Gardner 1982) and illustratively applied by Sfard (1998) in the educational domain.

2 Metaphors of Human Thinking

To convey a sense of the utility of the metaphors, we outline for each metaphor (1) the underlying understanding of human thinking, (2) an example of the use of the metaphor in UI design, and (3) some key questions to consider in evaluation. Further illustrations of the usefulness of the metaphors in HCI can be found in Frøkjær & Hornbæk (2002) and Hornbæk & Frøkjær (2002).

Metaphor of Habit Formation. Habits are shaping most of our thought activity and behaviour—e.g. as physical habits, automaticity, all linguistic activity, and habits of reasoning. The metaphor is: Habit formation is like a landscape eroded by water. We propose this metaphor to indicate how a person's formation of habits leads to more efficient actions and less conscious effort, like a landscape through erosion adapts for a more efficient and smooth flow of water. Creeks and rivers will, depending on changes in water flow, find new ways or become arid and sand up, in the same way as a person's habits will adjust to new circumstances and, if unpracticed, vanish.

In design, there is an abundance of examples of user interfaces that violate human habits. One example is adaptive menus, used for example in Microsoft Office 2000. Adaptive menus change the layout of the menu according to how often menu items are used, for example by removing or changing the position of items seldom used. However, adaptive menus make it impossible to form habits in the selection of menu items, since their position may be different from when they were previously selected. A study by Somberg (1987) showed the efficiency of constant position placement of menu items compared to menus that change based on use frequency. Somberg, however, did not explicitly link habit formation to the usefulness of constant placement of menu items.

In usability evaluation, this metaphor calls for considering: Are existing habits supported? Can effective new habits, when necessary or appropriate, be developed? Can the user use common key combinations? Is it possible for the user to predict, a requisite for forming habits, the layout and functioning of the interface?

Metaphor of the Stream of Thought. Human thinking is experienced as a stream of thought—in the continuity of our thinking, the richness and wholeness of a person's mental objects, of consciousness, and subjective life. The metaphor is: Thinking as a stream of thought. This metaphor was proposed by James to emphasize how consciousness does not appear to itself chopped up in bits: 'Such words as "chain" or "train" do not describe it fitly. It is nothing jointed; it flows'. Particular issues can be distinguished and retained in a person's stream of thought with a sense of sameness, as anchor points, which function as 'the keel and backbone of human thinking' (James 1890, vol. I, p. 459).

In design, a simple, yet effective, attempt to recreate part of the richness of the stream of thought when users return to resume interrupted work, is Raskin's design of the Canon Cat (Raskin 2000). When the Canon Cat is started, the display immediately shows up as it was before work was suspended. Not only does this allow the user to start thinking about the task at hand while the system is booting. It also provides help in remembering and recreating the stream of thought as it was when work was interrupted.

In usability evaluation, this metaphor calls for considering: Is the flow in users' thought supported in the interface by recognizability, stability and continuity? Does the application make visible and easy accessible such interface elements that relate to the anchor points of users' thinking about their tasks? Does the application help users to resume interrupted tasks?

Metaphor of the Dynamics of Thinking. Here is considered the dynamics of human thinking, the awareness shaped through a focus of attention, the fringes of mental objects, association, and reasoning. The metaphor is: Awareness as a jumping octopus in a pile of rags. This metaphor was proposed by Naur (1995) to indicate how the state of thought at any moment has a field of central awareness, that part of the rag pile in which the body of the octopus is located; but at the same time has a fringe of connections and emotions, illustrated by the arms of the octopus stretching out into other parts of the rag pile. The jumping about of the octopus indicates how the state of human thinking changes from one moment to the next.

In design, modal dialog boxes prevent the user from switching to potentially relevant information—in Microsoft Word, for example, it is not possible to switch back to the document to look for a good file name once the 'save as ...' dialog has began.

In usability evaluation, this metaphor calls for considering: Are users' associations supported through flexible means of focusing within a stable context? Do users associate interface elements with the actions and objects they represent? Can words in the interface be expected to create useful associations for the user? Can the user switch flexibly between different parts of the interface?

Metaphor of the Incompleteness of Utterances. Here is focused on the ephemeral character of utterances and their incompleteness in relation to the underlying thinking. The metaphor is: Utterances as splashes over the waves to the rolling sea. This metaphor was proposed by Naur (1995) to emphasize how utterances are incomplete expressions of the complexity of a person's current mental object, in the same way as the splashes tell little about the sea below.

For design, one implication of the metaphor of utterances as splashes over the waves is that we must expect users to describe the same objects and functions incompletely and in a variety of ways. Furnas et al. (1987) investigated the diversity in words used for describing commands and everyday objects. On the average, two participants described the same command or object by the same term with less than 20% probability. The most popular name was chosen only in 15-35% of the cases. Furnas et al.'s suggestion for relieving this problem is called the unlimited alias approach, where terms unknown to the system may be interactively related to existing commands or object names. This approach

is coherent with the metaphor and uses interactivity to clarify the intentions of the user. However, it would partly go against the metaphor of habit formation.

In usability evaluation, this metaphor calls for considering: Does the application support changing and incomplete utterances? Are alternative ways of expressing the same information available? Are interpretations of users' input made clear? Does the application make a wider interpretation of input than users intend or are aware of?

Metaphor of Knowing. Human knowing is always under construction and incomplete. The metaphor is: Knowing as a site of building in progress. Also this metaphor was proposed by Naur (1995) and meant to indicate the mixture of order and inconsistency characterizing any person's insight. These insights group themselves in many ways, the groups being mutually dependent by many degrees, some closely, some slightly. As an incomplete building may be employed as shelter, so the insights had by a person in any particular field may be useful even if restricted in scope.

In design, mental models have been extensively discussed. Consider as an example Norman's (1983) description of the use of calculators. He argues that the use of calculators is characterized by users' incomplete understanding of the calculators, by the in-stability of the understanding, by superstitions about how calculators work, and by the lack of boundaries in the users' understanding of one calculator and another. These observations by Norman are coherent with the ideas expressed by the metaphor of knowing.

In usability evaluation, this metaphor calls for considering: Are users forced by the application to depend on complete or accurate knowledge? Is it required that users pay special attention to technical or configuration details before beginning to work? Do more complex tasks build on the knowledge users have acquired from simpler tasks? Are users supported in remembering and understanding information in the application?

3 Experiment

To understand the effectiveness of metaphors of human thinking (MOT) as an usability evaluation technique, we conducted an experiment comparing MOT to heuristic evaluation (HE; Nielsen & Molich 1990). Eighty-seven computer science students used either HE or MOT to evaluate a web application (http://punkt.ku.dk). Each subject individually performed the evaluation supported by scenarios made available by the developers of the web application. Forty-four subjects received as description of MOT a psudonymized version of Hornbæk & Frøkjær (2002); 43 subjects received a description of HE from Nielsen (1993), 19-20 & 115-163. In all, subjects identified 911 problems.

In order to find problems that are similar to each other, we undertook a consolidation of the problems. In this consolidation, the two authors grouped together problems perceived alike. This resulted in a list of 341 consolidated problems. Next the client (i.e. the person who manages the development of the web application and is responsible for developing the design) assessed each consolidated problem. We asked the client to assess for each consolidated problem: (a) severity (on a scale from 1 to 3), (b) if design ideas were gotten from the problems (yes or no), (c) if the problem was novel (yes or no), and (d) the perceived complexity of solving the problem (on a scale from 1 to 4).

Table 1 shows a few important results from the experiment. By analysis of variance, we find no difference between the number of problems subjects identified with the two techniques, $F(1,85)=1.76$ p>.1.

Analyzing the client's assessment of the severity of problems, we find a significant difference between techniques, $F(1,85)=15.51$, p<.001. The client assesses problems identified with MOT as more severe (mean=2.21; standard deviation=0.73) than problems found by HE (M=2.42; SD=0.87).

	HE (N=43)	MOT (N=44)
Number of problems	11.3 (6.2)	9.6 (5.7)
Severity ***	2.4 (0.9)	2.2 (0.5)
Complexity ***	3.2 (1.0)	3.00 (0.8)
Novel problems***	3.8 (2.8)	2.0 (1.5)
Design ideas	2.5 (1.9)	2.2 (2.2)

Table 1: Usability problems identified with heuristic evaluation (HE) metaphors of human thinking (MOT). Severity was graded 1, 2 or 3, where 1 was given to a very critical problem and 3 was given to cosmetic problem. Complexity was graded from 1 to 4, where 1 was given to a very complex problem and 4 to a simple problem. All other rows refer to the average number of problems found by a subject. ***=significant difference between techniques.

The complexity of the problems identified is significantly different between techniques, $F(1,85)=12.94$, $p<.001$. The client assesses problems found with MOT as more complex to solve (M=3.00, SD=0.80) compared to those found by HE (M=3.21, SD=0.96).

Concerning the number of novel problems, HE identifies significantly more than MOT does, $F(1,85)=14.59$, $p<.001$. For both techniques, novel problems on the average are less severe (M=2.31; SD=0.75), are less complex (M=3.48; SD=0.71), and 41% are only found by one subject, suggesting that novel problems are mostly cosmetic and somewhat esoteric problems.

For reading and performing the evaluations, the subjects reported spending for MOT on average 4.0 hours (SD=2.3) and for HE 5.8 hours (SD=3.8). This difference is significant and large (Mann-Whitney U=546.5, z=-2.88, $p<.01$).

4 Conclusion

General properties of thinking activity known to all of us by introspection are emphasized through five metaphors, which build upon the work of William James and of Peter Naur. The metaphors catch psychological aspects of habit formation, stream of thought, awareness, utterances, and knowing. With the possible exception of awareness, these aspects of human thinking are rare in recent years of HCI literature (cf. Frøkjær & Hornbæk 2002). From commonly available user interfaces and from a selection of empirical studies, the utility of the metaphors was illustrated by their ability to clarify designs and notions in HCI. We suggest that the metaphors, by virtue of their psychological recognizability and focus on basic aspects of thinking, can help designers consider important traits of human thinking.

When using MOT as a usability evaluation technique, evaluation through the five metaphors is focused on the users' mental activity. It is remarkable how MOT in this first experiment has given good results compared to HE, the usability inspection technique most widely used in industry. HE usually performs very well in comparison with other inspection techniques, e.g. cognitive walkthrough and GOMS-based techniques (Cockton et al. 2003). It must be emphasized that our results are preliminary and have to be challenged by further studies. What happens when MOT is used for evaluating interfaces in non-traditional use contexts, when the evaluators are more proficient, or when MOT is used in design work? In this experiment, however, usability evaluation by metaphors of human thinking showed to be viable.

References

Cockton, G., Laveryn, D. & Woolrych, A. (2003), Inspection-Based Evaluations, in Jacko & Sears *The Human-Computer Interaction Handbook*, Lawrence Erlbaum, 1118-1138.

Frøkjær, E. & Hornbæk, K. (2002), Metaphors of Human Thinking in HCI: Habit, Stream of Thought, Awareness, Utterance, and Knowing, *in Proceedings of HF/OzCHI 2002*.

Furnas, G. W., Landauer, T. K., Gomez, L. M., and Dumais, S. T. (1987), The Vocabulary Problem in Human-System Communication, *Communications of the ACM*, vol. 30, no. 11, 964-971.

Gardner, H. (1982), *Art, Mind and Brain: A cognitive approach to creativity.* Basic Books.

Hornbæk, K. & Frøkjær, E. (2002), Evaluating User Interfaces with Metaphors of Human Thinking, *in* N. Carbonell & C. Stephanidis, eds., Proceedings of 7th ERCIM Workshop "User Interfaces for All, *Lecture Notes in Computer Science* 2615, Springer-Verlag, pp. 486-507.

James, W. (1890), *Principles of Psychology.* Henry Holt & Co.

Naur, P. (1995), *Knowing and the Mystique of Logic and Rules.* Kluwer Academic Publishers.

Nielsen, J. (1993), *Usability Engineering.* Academic Press.

Nielsen, J. & Molich, R. (1990), Heuristic evaluation of user interfaces, *in Proceedings of CHI'90* pp. 249-256.

Norman, D. (1983), Some Observations on Mental Models, in *Mental Models*, D. Gentner & A. L. Stevens, eds., Erlbaum, Hillsdale, NJ, pp. 7-14.

Raskin, J. (2000), *The Humane Interface: New Directions for Designing Interactive Systems.* Addison-Wesley.

Sfard, A. (1998), On two metaphors for learning and on the dangers of choosing just one, *Educational Researcher*, vol. 27, no. 2, pp. 4-13.

Somberg, B. L. (1987), A Comparison of Rule-Based and Positionally Constant Arrangements of Computer Menu Items, *in Proceedings of CHI+GI'87* ACM Press, pp. 255-260.