Evaluation of Peripheral Displays

Andreas Meiser, Georg von Peuerbach Gymnasium, Linz, Austria,
www.georgvonpeuerbach.at, meiser@georgvonpeuerbach.at

***Abstract*— There are many treatises on the evaluation of devices that make information of any kind available. However the evaluation of peripheral displays turns out to be difficult as their users do not interact with them directly. Most of the evaluations done on this topic concentrate on the results of the evaluation for producers and designers of peripheral displays.**

This paper proves that the fundamental attitude of the user towards this kind of display must be a part of the evaluation.

As neither possible users nor possible evaluators have a lot of experience with these displays the preconceptions of the people involved are more important for this kind of evaluation than for traditional displays.

These preconceptions have a direct influence on the results of the evaluation.

***Index Terms*— evaluation, peripheral displays, poll, users preconceptions**

***ACM Classification Keywords*— D.2.2: User Interfaces; H.5.2: Evaluation/methodology. General Terms: Human Factors**

I. INTRODUCTION

The task of the peripheral displays is to provide awareness with minimal attention and to present information without distracting or burdening the user.

In this sense this paper doesn't differ between peripheral and ambient displays. Mankoff et al. describe ambient displays as aesthetically pleasing displays of information which sit on the periphery of the user's attention[13].

Although a great number of displays of this kind can be found – its spectrum reaching from a simple wind sock indicating wind force and direction on highways to aesthetically designed objects like frosted orbs providing information through different colours and brightness[14][24] and to what looks like a modern painting on an office wall that might in fact show when the next bus leaves from a nearby station[21] – there are just a few attempts of evaluating these displays.

The main reason is that the user doesn't interact directly with these devices. The problem partly lies in the fact that it is often not clear how well the user realizes the peripheral display and therefore is in the position to use the presented information[9]. The conclusion must be drawn that common evaluation methods have to be adapted for the peculiarities of peripheral displays. The basis of this analysis are the methods for Usability Evaluation by Nielson[19].

They comprise usability inspection (heuristic, cognitiv walkthrough) and usability tests (thinking aloud, constructive interaction, videoconfrontation, coaching method, focus groups, observations, questionnaires). These methods are standards for Usability Evaluation and are well documented[10][11][16][17].

This paper tries to demonstrate that for the evaluation of peripheral displays not only the adaptation of these methods is absolutely necessary but that also completely new categories must be considered which arise from the peculiarity of the display as well as from the user's and evaluator's attitude towards this kind of display.

This paper was written for a seminar in the winter term 2005/06 on the topic "peripheral displays" at the Institute of Pervasive Computing, Joh. Kepler University, Linz Austria and was handed in for an expert's valuation on December 12th, 2005.

II. RELATED WORK

Evaluation of peripheral displays is not a straightforward task. An obvious approach is to measure usability. Referring to the special characteristic of peripheral displays (they are designed to operate in the periphery, while other tasks occur in the foreground) an evaluation must measure not only usability, but also other important factors like awareness and distraction[7][15].

This paper gives examples of some techniques used to evaluate these factors. Almost no work can be found considering the evaluator's preconceptions towards peripheral displays.

A. Heuristics

In his work Nielson mentions the following set of heuristics: (1) Visibility of system status; (2) Match between system and the real world; (3) User control and freedom; (4) Consistency and standards; (5) Error prevention; (6) Recognition rather than recall; (7) Flexibility and efficiency of use; (8) Aesthetic and minimalist design; (9) Help users recognize, diagnose and recover from errors; (10) Help and documentation.

Nielsen found that 3-5 novice evaluators find 40-60% of known issues when applying his canonical list of heuristics[18].

These results refer to the evaluation of traditional displays.

Considering that peripheral displays are not used in the same way as traditional displays, as users don't interact with them directly, but perceive them on the periphery of their attention, Mankoff et al. developed an adapted set of heuristics[13]. They eliminated some heuristics (numbers 3,4,5,7,9,10) and added 5 heuristics more specific to ambient displays.

This resulted in the following adapted set: (1) Useful and relevant information; (2) "Peripherality" of display; (3) Match between design of ambient display and environments; (4) Sufficient information design; (5) Consistent and intuitive mapping; (6) Easy transition to more in-depth information; (7) Visibility of state; (8) Aesthetic and Pleasing Design. They used this adapted set to evaluate 2 ambient displays.

The purpose of the one called 'busmobile' is to provide information about popular bus lines, the second one called „Daylight Display", has its purpose in providing information about whether it is dusky, light, or dark outside.

The analysis showed that the adapted set of heuristics provided better results than Nielson's original set, especially if you consider that heuristic evaluation is more effective if a small number of evaluators is able to find a high percentage of problems.

The analysis showed that a single evaluator using the new heuristics finds 22% of known major problems on average, and eight evaluators are sufficient to find about 70% of known problems. In contrast, using Nielsen's heuristics, a single evaluator will only find about 13% of major issues and eight evaluators find about 50% of known issues.

However it turned out that the adapted set did not adequately address error conditions. The conclusion drawn by Mankoff et al. is that a combination of Nielson's set of heuristics and the adapted set might show the best results. The final result of their work is the following set of heuristics: (1) Sufficient information design; (2) Consistent and intuitive mapping; (3) Match between system and real world; (4) Visibility of state; (5) Aesthetic and pleasing design; (6) Useful and relevant information; (7) Visibility of system status; (8) User control and freedom; (9) Easy transition to more in-depth information; (10) "Peripherality" of display; (11) Error prevention; (12) Flexibility and efficiency of use.

B. Questionnaires

Based on works in which awareness and distraction in connection with peripheral displays were measured directly and quantitatively[2][3][4][12], T. Matthews, A. Dey and J. Mankoff analysed 2 peripheral displays which should support the management of emails[14].

The group of people interviewed get a large amount of emails every day. They stated that they felt obliged to open incoming emails immediately in order not to miss an important message. In most cases the emails turned out to be either spam or emails of subordinate importance.

The peripheral displays tested by Matthews, Day and Mankoff should improve the situation. On the one hand they used a ticker, a common type of onscreen display that shows scrolling text, for their study, on the other hand they used a commercial display, a physical, frosted Orb (Ambient Orb) that sits on the user's desk and changes color in response to user input. In contrast to the ticker, the Orb displays information off the desktop and more abstractly.

They used the adapted set of heuristics developed by Mankoff et al.[13] to optimize the design of the tested displays. Then they carried out Lab Studies and Field Studies.

The lab study methods included selfreporting of awareness and distraction, objective records of performance on questionnaires about the peripheral display contents. The field study included selfreporting of awareness, distraction, and usability; objective records of performance on questionnaires about the peripheral display contents and qualitative information from interviews.

To measure awareness, they used self-reporting questions in which users were asked to tell how much attention they paid to a display, and knowledge questions, in which was tested how much information they had retained from the display objectively.

To measure distraction, also self-reporting questions were used and also primary task completion speed and accuracy were measured.

To measure usability, participants were asked to rate the displays against the heuristics used in the former heuristic evaluation, and also interviews were made about the participants use of the displays.

Apart from results for the use of displays (for example most of the users constantly checked their emails all the same, so that they didn't pay attention to the information provided by the peripheral display. Nevertheless they could avoid opening unimportant emails more often. Furthermore the ticker was considered as less disturbing than expected, …) new findings were made considering test methods on awareness and distraction. A combination of selfreports of awareness and distraction in combination with interviews about the use of the displays seems absolutely necessary.

In addition an objective measuring of awareness and distraction in comparison to the measurement of the work results is useful.

C. Comprehension

Holmquist[9] directs his attention to the consideration of comprehension, i.e. how well a user understands (and, consequently, is able to use) an ambient display.

He describes three levels of comprehension, where each is a pre-requisite of the next: 1. That information is visualized; 2. What kind of information is visualized; 3. How the information is visualized.

As an example he uses a display that looks like an abstract painting but in reality provides information about the next bus leaving from a nearby station[21].

Holmquist shows that comprehension over time is an important factor when evaluating ambient displays.

D. In situ evaluation

S. Consolvo and J. Towle[6] evaluated an ambient display called the CareNet Display[5].

It is an interactive digital picture frame that augments a photograph of an elder with information needed by the friends and family who provide the regular care she needs to remain at home. The status of the elder's meals, medications, outings, activities, mood, falls were shown and calendar were updated throughout the day as events occurred.

Consolvo and Towle carried out both, a heuristic evaluation as well as an in situ evaluation. They used the set of heuristics developed by Mankoff et al.[13] and proved its effectiveness. On the other hand the heuristic evaluation could not reveal certain problems, which became evident in the in situ evaluation. For example the display couldn't be well seen from different angles, its radiation was very bright during the night, which was considered as disturbing.

This leads to the conclusion that a combined evaluation shows better results, which has already been stated by Heather Desurvire [8].

E. Simulation

Young and Mann tested peripheral displays in connection with GPS-based guidance aids[23]. Most systems use a horizontal array of LEDs – called a lightbar – to indicate lateral error to machinery operators. LEDs to the left or right of centre light up when the machine is moving away from the correct path. Young and Mann built a 1420 mm wide lightbar composed of 23 lighted elements, where each element was a 46 mm square cluster of 16 LEDs and compared this system with a purchasable one, consisting of 23 single LEDs. For this purpose they simulated the control cabin of a high clearance filed sprayer because this machines were the most popular machines for GPS guidance aids. Twenty-two undergraduate and graduate students were used in the simulator study. All subjects had experience driving vehicles, and 11 of 22 subjects had experience driving agricultural equipment

During the simulation not only the effectiveness of the display was measured, but also further parameter, like for example the average reaction time and the heart rate measurement of the participants. Additionally the participants were handed out questionnaires to get subjective feedback. The study showed that the peripheral display resulted in an improvement of the steering performance. Some subjects noted that they could see the large lightbar better so they could continue to interpret lightbar information while looking at other things through the windshield. The larger lightbar was easier to notice when a person looked away and looked forward again.

III. PRECONCEPTS

Most of the treatises on the evaluation of peripheral displays that have been published so far regard the evaluation mainly as a means to get better criteria for the design of peripheral displays[1][20][22].

The decisive question in daily use is, in addition to the objective advantage that can be reached by using the peripheral display, how far peripheral displays are accepted by people.

As the evaluation is often carried out by the users themselves the mental attitude of the potential user is an important factor for the interpretation of evaluation results that must not be neglected.

A. Research Method

To examine this problem more closely, a poll on the fundamental acceptance of peripheral displays was carried out on a high school. The following scenario was the basis of this poll:

Schools offer several open days for parents a year. Although all teachers have a consulting hour every week, most of the parents make use of the open days. For this reason parents have to queue for quite a while until they can talk to the teacher. The parents are often annoyed that they are kept waiting for so long. For teachers it's sometimes difficult to keep talks short. As the teachers do not know anything about the number of parents waiting outside the room, it's difficult for them to judge the correct period of time that is at their disposal for each parent.

The teacher never knows how many parents there are outside and for how long they have already been waiting. Every now and then the door is opened by waiting parents, which is considered as most disturbing. On the other hand it could be that nobody is waiting outside and the teacher would have time for giving detailed information.

The test to offer each parent just a certain period of time has proved to be unsatisfactory in certain situations, as it may sometimes happen that a detailed conversation is absolutely necessary and cannot be delayed. Basic solutions concerning the organisation like putting one's name on a list have been tried out in different versions, but all proved to be inconvenient.

To handle this problems a purchasable peripheral display (an ambient orb[24]) was suggested as a possible solution.

These displays should be put in all the classrooms. By changing their colour and brightness they should inform the teachers about how many parents are waiting outside.

B. Test

38 teachers aged between 30 and 55 having at least 3 years of teaching experience and therefore the same experience with open days were asked.

The teachers were distinguished by their sex and if they teach scientific subjects or not.

This poll was carried out at two different levels. On the one hand the aim was to find out about the teachers' general attitude towards the peripheral display described above. On the other hand specific questions on the teachers' personal

situation on open days were asked, as well as questions on the effectiveness of the display, expected distractions, acceptance by the parents and their general importance for the school.

The participants of this evaluation got questionnaires giving the following information: a summarizing presentation of the situation on open days, a description of the function of an ambient orb including its photo.

There have been several changes in the system and organisation of this school during the last few years and because of this the teachers are highly sensitive towards innovations. This may lead to spontaneous defensive reactions towards innovations. It was of high importance to inform the participants that every positive response would in no case lead to the installation of an ambient orb automatically without additionally discussions.

The influence of possible problems in connection with the installation of an ambient orb on the answers given was tried to be avoided.

It was explicitly mentioned that the questionnaire deals with the general attitude towards an ambient orb and that questions on how to finance or how to install it are of absolutely no relevance.

The questionnaire offered some space for additional remarks.

C. Results

The answers show no decisive difference concerning males and females and no significant difference concerning the subjects taught. No proof could be found that teachers of scientific subjects basically show a more positive attitude towards peripheral displays than teachers of other subjects.

38 teachers were asked. The result of the first level (teachers' general attitude towards the peripheral display) showed that 10 (26.3%) of them completely refused the display, 4 (10.5%) considered it as brilliant and 24 (63.2%) expressed a slightly optimistic view ("We could try").

In response to specific questions the following insights were gained: Anyone (100%) out of the first group, nobody (0%) of the second group and 9 (37.5%) out of the third group considered the display as no improvement, additionally 16 (66.7%) out of these 24 called it "interesting gadgets".

The results concerning the supposed distraction were of great interest. Out of these 10 teachers refusing the display, 7 (70%) considered it as distracting, none of the 4 supporting teachers saw a distraction and 3 (12.5%) of the remaining group saw a distraction.

Concerning their own distraction and the one of the parents all 7 teachers of the first group thought parents would be distracted as well, the 3 teachers of the last group saw just their own distraction.

The question about a possible innovation for the school was not answered positively by all the teachers of the first group. 6 (25%) teachers of the second group answered the question positively and 2 (50%) of the last group did so.

Important additional remarks were that the expenses for the introduction should not be too high. Some were of the opinion that just the teacher should see the ambient orb. Some also thought about improvements like the use of peripheral displays not only on open days at school but also to present further information in order to make its use more efficient.

D. Discussion

The poll resulted in a balanced correlation between negative attitude and supposed distraction (Pearson $\rho = 0.59$). The correlation between negative attitude and supposed distraction of the parents is very high (Pearson $\rho = 0.9$). This fact is of great interest because there are not any objective hints that parents could be more distracted by an ambient orb than by waiting outside the room for an unknown period. So you may come to the conclusion that evaluating the distraction of peripheral displays without evaluating the evaluator's attitude towards this kind of displays appears to be problematical.

The possible argument that heuristic evaluations are mainly carried out by professional evaluators can be proved wrong by a simple financial argument. Evaluation in statal or non-profit oriented organisations (where schools belong to) do not have the required financial means to employ professional evaluators. In most organisations of this kind the people concerned evaluate the innovation.

A difference in the objective criteria like sex or subject taught was expected. The poll shows that it was impossible to make assumptions about the answers in advance. The poll was originally carried out anonymously. In the following discussion many teachers revealed in which way they had answered. On this account it was possible to compare the teacher's judgement to other known positive or negative attidudes associated with innovations at this school.

No significant relation between the teachers' judgement of peripheral displays and other projects done at school became evident. No proof could be found that teachers refusing other projects at school also showed a negative attitude towards the ambient orb.

This facts indicate that the basic attitude of the people involved must be found out before the evaluation so that it can be taken into consideration for the interpretation of the evaluation results.

It could be expected that not only the evaluation of distraction but also evaluation in general will be affected by user's and evaluator's preconceptions towards peripheral displays.

IV. FUTURE WORK

Future works should find out reasons for the refusal of peripheral displays. If the reason for the refusal can be reduced to the fact that peripheral displays are new and therefore strange or partly they are even seen as gadgets, then the importance of this paper must be considered with priority.

Carrying on the work of Holmquist[9] seems interesting because the fact that it is a question if a display actually

presents information or rather in which way some information is presented should be analysed in connection with the user's general attitude.

More work is required to gather further information about the consequences of this work concerning not only the evaluation of distraction but also the evaluation of peripheral displays in general.

ACKNOWLEDGMENT

Many teachers of my school took part in the poll and in further discussions. I would like to thank them for their time and effort. Many thanks also to Gerhild Tulzer for translating the manuscript.

REFERENCES

[1] Ames, M., Dey, A., "Description of design dimensions and evaluation for Ambient Displays", UC Berkeley Computer Science
[2] Arroyo, E., Selker, T., "Arbitrating multimodal outputs: Using ambient displays as interruptions", Proceedings HCII'03, 2003.
[3] Bartram, L., et al., "Moving icons: Detection and distraction", Proceedings Interact'01, 2001.
[4] Cadiz, J. J., et al., "Designing and deploying an information awareness interface", Proceedings CSCW'02, pp. 314–323.
[5] Consolvo, S., Roessler, P., Shelton, B.E., "The CareNet Display: Lessons Learned from an In Home Evaluation of an Ambient Display", Proceedings of the 6th Int'l Conference on Ubiquitous Computing: UbiComp '04, (Sep 2004), pp.1-17.
[6] Consolvo, S., Towle, J., "Evaluating an Ambient Display for the Home", Proceedings CHI 2005, April 2–7, 2005, Portland, Oregon, USA.
[7] Day, J. A., "Peripheral Displays for Information Awareness", available: www.cc.gatech.edu/grads/d/Jason.Day/documents/ miniProject2WriteUp.pdf
[8] Desurvire, H., "Faster, Cheaper! Are Usability Inspection Methods as Effective as Empirical Testing?", In: Usability Inspection Mehods. Ed. by Jakob Nielsen and Robert L. Mack. New York u.a.: Wiley, 1994. S. 195.
[9] Holmquist,L.E., "Evaluating the Comprehension of Ambient Displays", Proceedings CHI 2004, April 24–29, 2004, Vienna, Austria.
[10] Lewis, C., Wharton, C., "Cognitive Walkthroughs", In: Handbook of Human-Computer Interaction: M. Helander et al. 2nd, completely rev. ed. Amsterdam u.a.: Elsevier. p. 718-725.
[11] Leßmann, C., Wolff, M., "ErgoNet - ein netzbasiertes, multimediales System zur Unterstützung von Ad-hoc Usability tests", Diplomarbeit, Universität Bremen, FB 3 Mathematik/Informatik. Bremen, 1999. p. 42-46
[12] Mamykina, L., et al., "Time aura: Interfaces for pacing", Proceedings CHI'01, pp. 144–151.
[13] Mankoff, J., Dey, A.K., Hsieh, G., Kientz, J., Lederer, S., Ames, M., "Heuristic Evaluation of Ambient Displays", Proceedings of the Conference on Human Factors in Computing Systems: CHI '03, (Apr 2003), pp.169-76.
[14] Matthews, T., Dey, A., Mankoff, J., "A Comparison of Two Peripheral Displays for Monitoring Email", Proceedings of CHI 2004, pp. 463-470. [MDMCR04].
[15] McCrickard, D. S., Catrambone, R., Stasko, J. T., "Evaluating Animation in the Periphery as a Mechanism for Maintaining Awareness", Proceedings of INTERACT 2001, Tokyo, Japan, July 2001, pp. 148-156.
[16] Nielsen, J., Molich, R., "Heuristic evaluation of user interfaces", Proceedings of Conference on Human Factors in Computing Systems (CHI'90), pp. 249–256. ACM Press, 1990.
[17] Nielsen, J, "Usability Engineering", Boston a.o.: AP Professional, 1993. p. 224
[18] Nielsen, J., "Ten usability heuristics", http://www.useit.com/papers/heuristic/heuristic_list.html, accessed on August 20, 2002.
[19] Nielsen, J., "Usability inspection methods", In: Tutorials of the ACM CHI '95 (07. - 11. May 1995), pp. 377-378.
[20] Osuntogun A., "Design, Implementation and Evaluation of a Peripheral Display", University of South Florida, Summer Undergraduate Program in Engineering Research at Berkeley, (SUPERB) 2003
[21] Skog, T., Ljungblad, S.,Holmquist, L.E., "Between Aesthetics and Utility: Designing Ambient Information Visualizations", Proceedings of InfoVis 2003, IEEE, 2003.
[22] Young, S., Mann, D. D., "Development and Evaluation of Peripheral Display for Agricultural Guidance Aids", AIC 2002 Meeting, CSAE/SCGR Program, Saskatoon, Saskatchewan, July 14 - 17, 2002
[23] Wai-ling Ho-Ching, F., Mankoff, J., Landay, J. A., "From Data to Display: the Design and Evaluation of a Peripheral Sound Display for the Deaf", EECS Department, University of Cal., Berkeley, Technical Report No. UCB/CSD-02-1204, 2002
[24] available: http://www.ambientdevices.com

Andreas Meiser has been teaching Mathematics, Physics and Computer Science at the Georg von Peuerbach High Scool, in Linz, Austria since 1995. He is responsible for further vocational training in the field of informatics on this school and furthermore administrates the schools computernetwork. His main interest can be found in the introduction and evaluation of new media at this school. Presently he researches into the development and evaluation of web-based learning methods and strategies.