IASDR07
INTERNATIONAL ASSOCIATION OF SOCIETIES OF DESIGN RESEARCH THE HONG KONG POLYTECHNIC UNIVERSITY 12TH TO 15TH NOVEMBER 2007

METHODOLOGICAL ISSUES IN RESEARCH INTO WEB DESIGN FEATURES

Virgínia Tiradentes Souto

Universidade de Brasília, Departamento de Desenho Industrial, Campus Universitário Darcy Ribeiro - SG1, Brasília - DF, 70910-900, Brazil, vsouto@unb.br

ABSTRACT:

Different methods have been used in order to investigate the layout of websites. There is a question about which methods produce valid findings. This paper gives a brief overview of different methods commonly used to investigate the layout of websites. It also uses two case studies in order to compare the findings from different methods: one that used experimental research followed by a questionnaire and another that used think-aloud test followed by an interview. The findings suggest that a combination of different empirical methods seems to be an effective way to investigate different aspects of the layout of websites. The use of a global approach as an alternative method to investigate the layout of websites is proposed. Finally, this paper draws attention to the relevance of increasing the number of investigations done by design researchers into the web design field.

Keywords: **design methods, interface design, World wide web**

1. INTRODUCTION

The Internet's growth in the last decade has made researchers become more interested in investigating the ease of use of websites. Different methods have been used in order to investigate the layout of websites (e.g. experimental research, think-aloud, observation). However, there is a question about which of the different methods used are most appropriate for evaluating the effect of website layout on users´ performance, perception and preference. The validity of empirical methods used to investigate layout aspects has been discussed by some researchers (Lund 1999; Dyson 2004).

Studies on the effect of users´ performance and preference for websites are usually called usability studies. Many of these studies are used to investigate a particular case, whereas there are other studies that investigate the existence of a natural phenomenon which can be applied to many cases.

Some of these methods have been shown to be more appropriate for achieving some types of findings (e.g. user performance differences), whereas there are other methods that seem more appropriate for gathering other types of findings, such as preference differences.

In order to illustrate the differences in two of the most common types of methods used to investigate the layout of websites, two case studies are reviewed. These studies are briefly described and the advantages and disadvantages of the methods used in these studies is discussed.

An alternative approach to these most common methods used to investigate the layout of websites is also proposed. This approach may be called the global approach (Dyson 2004) or kitchen sink approach (Muter 1996) and consists of comparing layouts which vary in a number of features, instead of isolating a few variables. This approach is described and commented on.

The paper starts with a brief overview of the most common methods used to investigate the layout of websites and ends with conclusions about the relevance of design research in this subject.

2. METHODS COMMONLY USED TO INVESTIGATE WEB DESIGN FEATURES

There are many different methods that can be used to investigate web interface design. These methods may or may not involve users in the investigations. Methods that involve expert evaluation or reviewers in a web interface investigation instead of users are usually called usability inspection methods. On the other hand, methods in which the involvement of users is critical to the investigation are called empirical methods. Empirical methods used in order to investigate the usability of websites can be divided into two groups: empirical usability methods and empirical research methods.

Both usability inspection methods and empirical usability methods may be included in the 'clinical research' category described by Buchanan (2001). This is because both groups of methods are commonly used to solve a particular, individual case. On the other hand, empirical research methods are included in the applied research category. Empirical research methods are used to investigate natural phenomena. In contrast with usability methods, research methods aim to find answers to questions related to many individual cases. The most common usability inspection methods, empirical usability methods, and empirical research methods are briefly described below.

2. 1. USABILITY INSPECTION METHODS

Usability inspection methods aim to predict the problems that users might find while interacting with an interface (Virzi 1997; Rogers and Sharp 2002). These methods involve experts and reviewers instead of users and therefore are called non-empirical methods. Examples of inspection methods are: action analysis, cognitive walkthrough, and heuristic evaluation.

The most common usability inspection method is heuristic evaluation (Holzinger 2005). This method involves experts evaluating whether the elements of the interface are conforming to a set of usability principles called heuristics (Nielsen and Molich 1990). Another common usability inspection method is cognitive walkthrough. It involves the simulation of each step of a user's problem-solving process when doing a task (Mack and Nielsen 1994).

Some advantages of these methods are the cost (Rogers and Sharp 2002), the rapidity, and identification of major and minor problems (Holzinger 2005). The main disadvantages are: non-involvement of the end user and inability to identify unknown users' needs (Holzinger 2005).

2. 2. EMPIRICAL USABILITY METHODS

As mentioned above, empirical usability methods are used to improve the usability of a particular case. There are many different kinds of empirical usability methods. The most common empirical usability methods (also called usability test methods) are: performance measurement, observation, questionnaires and thinking aloud.

Performance measurement is one of the most common empirical usability methods. Performance measurement aims to evaluate user performance in order to improve usability design. This method can be used to compare different interfaces and also to assess whether usability aims have been met (Nielsen 1993). User performance usually measures the time taken and the number of errors made by a group of test users performing a set of tasks (Nielsen 1993). Some limitations of this method are that it 'emphasizes first-time usage' and covers a limited number of interface features (Shneiderman 1998).

Questionnaires are another well-established method in which the opinions of the users about the user interface are collected (Holzinger 2005; Rogers and Sharp 2002). Questionnaires are especially useful to measure subjective satisfaction and preferences of the users (Holzinger 2005). Interviews, a similar method to questionnaires, can be defined as 'a conversation that has structure and a purpose' (Kvale 1996).

Another type of user testing is observation. It involves watching and listening to users (Preece, Rogers and Sharp 2002). Although an observation test does not always require technological apparatus, the design of a test and the data analysis of an observation test can be complex. Observation can be useful to find out what users do, their work context, and how well technology supports them. This method can be done in either a controlled environment or in a natural environment (Rogers and Sharp 2002).

Finally, another common type of user testing is called think-aloud method (also called concurrent protocols). This method refers to a test where users are asked to verbalise their thoughts while performing tasks. It provides information on how users view the system and which problems users

have with the interface (Holzinger 2005). Think-aloud method is considered one of the main empirical methods for improving user interfaces (Nielsen 1993; 1994). Section head: 14pt single spacing, all caps, aligned left.

2. 3. EMPIRICAL RESEARCH METHODS

Empirical research methods can be divided into three main approaches: quantitative, qualitative, and mixed approaches (Creswell 2003). Quantitative research is concerned with measurements and statistical analysis (e.g. experiments). Qualitative research concerns descriptions and open-ended data analysis. In the mixed approaches both quantitative and qualitative data are collected.

Quantitative research, more specifically experimental research, is the most common empirical research method used to investigate interfaces. In experimental research the investigator manipulates and controls independent variables (i.e. properties of objects or events) and observes dependent variables (i.e. the data) (Howell 1982). Examples of dependent measures commonly used are: performance times, preference, accuracy, and efficiency.

Experimental research is considered both a complex method (Shneiderman 1998) and the most accurate of the research methods (Atwater and Babaria 2001). According to Kerlinger (1986) the controlled experiment is the ‘desired model of science’ for research ‘with the purpose of discovering, classifying, and measuring natural phenomena and the factors behind such phenomena’. The main disadvantages of experimental research are: usually the results of experimental research are more specific and less general than in non-experimental research; and few research questions can be explored because of the control requirement (Atwater and Babaria 2001).

Experimental research has some similarities to performance measurement, such as: they take place in a similar 'physical setting' and deal with the collection of similar kinds of data (Dumas and Redish 1999). However, they also have some strong differences, such as their aims. As previously stated, experimental research aims to test whether or not some specific phenomenon exists, whereas user testing aims to evaluate user performance in order to improve usability of a particular design. These methods also differentiate the number and choice of participants, number of variables, replication and documentation, and data analysis.

Other common types of empirical research methods are: observation, interview, and questionnaire methods. These methods are similarly used in both usability methods and research methods. However, as mentioned above, the aims and the requirements of usability and research methods are different.

3. CASE STUDIES

This paper compares two studies on the layout of government websites. The first study used experimental research method followed by a questionnaire response whereas the second study used think-aloud method followed by an interview. These studies were chosen because they were planned to investigate similar features of the layout of websites and used different methods to do so. It is relevant to compare these methods (experimental research and think-aloud method) because they are both very frequently used to investigate the design of websites and their outcomes demonstrate some marked differences. Experimental research is used to find out whether or not a phenomenon exists (considered applied research), whereas think-aloud is commonly used to investigate the usability of a particularly interface (considered clinical research).

As mentioned above, in experimental research one or a few variables should be isolated and the influence of other variables should be controlled. These characteristics mean that this method has some limitations. It may mean that some research questions related to interface design cannot be answered through experimental research. For example, when the layout of an interface is investigated as a whole, and a confounding variable appears. Furthermore, the interface designed for experimental research may appear unfamiliar in some cases due to the need to isolate the controlled variables.

In order to get the users' thoughts and/or perception on typical websites, types of methodologies other than experimental research can be used. Other methods can provide the perception of the user in relation to a more naturalistic environment, such as think-aloud protocols. Think-aloud method (described in section 2.2.) is commonly used in Human-Computer Interaction studies. Many authors claim that the think-aloud test is an effective method for getting people's thoughts on mediating the completion of a task (e.g. Ericsson and Simon 1993; Birns, Joffre, Leclerc and Paulsen 2002).

Think-aloud method is frequently used in cognitive psychology research (Branch 2000) and also in graphical user interface (GUI) studies (e.g. Hughes Packard and Pearson, 1997; Branch 2000). Most of the investigations on websites which use the think-aloud methodology are done in order to improve the usability of a specific website (e.g. Benbunan-Fich 2001; Bunz 2001; Zimmerman, Akerelrea and Buller 2003). Therefore, these studies are considered empirical usability methods, not empirical research methods.

Both studies reviewed here combine a user test with a method of getting the users' perception. While the first study followed the user test with a questionnaire response, the second study followed the user test with an interview. They were both applied to get the users' preference for and perception of the interfaces. However, questionnaire response is usually considered a more limited method than interview. A possible advantage of interview over questionnaire method is that, with the former, participants may be more spontaneous in giving their thoughts. This is because in an interview they do not have to write their thoughts down and therefore may give a more detailed contribution. In addition, with unstructured interviews the interviewer may add new questions in order to get a deeper understanding of participants' thoughts, which is not possible with questionnaires.

These two studies are just briefly described. For a complete description of these studies and their findings see Souto (2006). The main point of including these studies in this paper is to show how different methods can produce different findings while the same features are investigated. They also show that different methods may be used to produce complementary findings and/or to confirm their findings, or not.

3. 1. CASE STUDY ONE: EMPIRICAL RESEARCH METHOD FOLLOWED BY A QUESTIONNAIRE

The first study aimed to explore the effect of the number of links and the number of groups of links on visual search of government websites. Two numbers of links on a single screen (24 and 48), three numbers of groups of links (2, 4 and 12) and two types of tasks (known-search and exploratory-search) were investigated. Figures 1 and 2 illustrate two out of the six combinations of layouts investigated.

Figure 1: Example of 48 links divided into 2 groups.

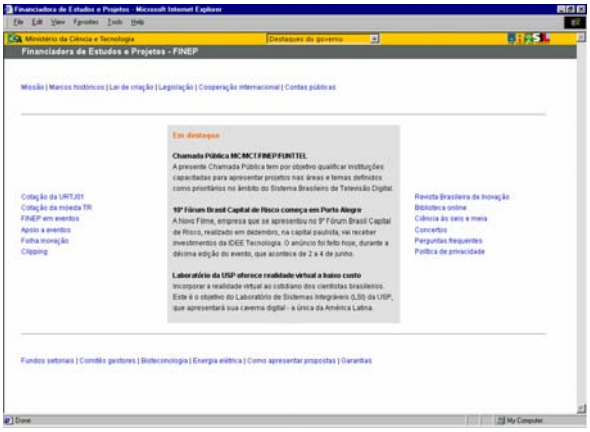

Figure 2: Example of 24 links divided into 4 groups.

The design used in this study was the same for all versions, changing only the investigated variables and the aspects related to them (e.g. number of links per group, and position of links on screen). The number of links and the number of groups were chosen based on a previous survey. Participants were tested individually. Participants were required to answer three sets of questions using three versions of websites. Each of these versions had a different number of groups of links. Half of the participants used the three versions with 24 links and the other half with 48 links. After completing the three sets of questions, participants were asked to fill in one questionnaire with their personal information and another with their subjective judgments.

This study investigated four main research questions. The kind of question investigated was, for example: ´Does the number of groups of links (2, 4 or 12) affect the ease with which the user finds information?´ As an example of the kind of findings of this method were: participants were faster using 24 links than 48 links; no significant differences in accuracy between 24 and 48 links; no significant differences in the number of wrong clicks between 24 and 48 links.

The type of question and the findings that experimental research shows are usually very focused and precise. On the one hand, the findings of this type of method may provide the researcher with the possibility of creating principles or establishing connections among many individual cases; on the other hand, it limits the way researchers draw conclusions and how far they can generalise. For example, in this study one of the possible conclusions is that there is evidence that users prefer fewer links to many links on screen. However, the tests were done with a fixed number of links (24, 48 links). Therefore, the conclusions of this study need to bear in mind the particularities of the variables investigated.

Another difficulty with experimental research is the test of some unfamiliar layouts. These unfamiliar layouts may be necessarily used in order to balance the layout. In this study, for example, it may be argued that some aspects of the design used are not commonly seen in current government websites (e.g. placing a large number of links horizontally across the top of the screen, as shown in figure 1). Therefore, in order to test a more realistic interface, another type of method rather than experimental research should be used.

3. 2. CASE STUDY TWO: THINK-ALOUD METHOD FOLLOWED BY AN INTERVIEW

The aim of second study was to investigate the users' perception of the arrangement of links in government websites. The study investigates real online government websites in the users' natural environment of Internet use. This study also investigates where users look first when using a government website, the search strategy and the link characteristics which help or hinder users in finding information on screen. After the users had experienced the online websites, an interview was conducted to get the users' perception of the arrangement of links in these websites.

Four online Brazilian government websites are the object of the present study. Figure 3 and Figure 4 show examples of two homepages of these websites. Each participant used all the websites and was tested individually. Participants performed five tasks in each of these websites. The test was done in their own environment of Internet use, either at work or home. The participants were required to think aloud and explain (a) where they look on the screen and (b) the links they choose to click on while they try to find answers for questions using four online government websites. After participants had finished the tasks, they were interviewed about their perception of the websites they had used.

Figure 3: Illustration of the homepage of the Ministry for Foreign Relations (MRE) website.

Figure 4: Illustration of the homepage of the Ministry of Development, Industry and Foreign Trade (MDIC) website.

Examples of research questions are: What are the users' strategies when searching for links on websites? What do the users think about the number of links and the number of groups of links on a web page? Examples of the findings of this study are: where the participants look first and their search strategy are both dependent on the layout of the website; clean website and not too many links on one web page are characteristics that users found helped them find information on screen.

Think-aloud method appeared to be an effective method to find users´ perception of an interface. In the study described, many different aspects that users perceived in relation to the interface were pointed out. This method makes it possible to gain a deeper understanding of users´ views

in relation to the layout of a website. The advantage of this method is especially clear when it is combined with an interview, as in the case study described. This is because the interview helps to elucidate users´ thoughts that were not spontaneously described during the think-aloud test.

3. 3. DISCUSSION OF THE CASE STUDIES

The findings of the first study show that the use of experimental research followed by questionnaire responses is an effective way of investigating the research questions on the layout of websites. This is because the data collected in the first study showed the users´ performance and preference for each of the variables investigated. However, as mentioned above, there are some restrictions on when experimental research followed by a questionnaire response can be applied. In cases where a more realistic environment is needed, this method may not be appropriate.

The findings of the second study show that the use of think-aloud test followed by an interview is a valid method for identifying user perception in relation to the layout of websites. This is because the data collected in the second study captured the thoughts of many users and elucidated their perceptions of the topics investigated. In addition, through these methods researchers can investigate how users perform tasks with real websites in their natural work environment. Therefore, the author of this paper agrees with the claim of some researchers that the think-aloud test is an efficient method for investigating the usability issues of websites.

However, think-aloud followed by an interview does not seem to be the most appropriate method for indicating the effects of some specific layout features, such as the types of links on screen. This is because asking participants to think aloud while doing the test may be disruptive and may therefore influence the performance of the participants. Therefore, the use of a think-aloud test together with a performance test needs to be treated with some caution. In general terms, a think-aloud test with interview seems more appropriate for investigating users' thoughts, while experimental research seems more appropriate in the investigation of users' performance.

As mentioned above, one of the problems with experimental research is the fact that one or a few variables should be controlled. On the other hand, non-controlled experiments, like the second case study, may not be the most appropriate method for indicating the effect of specific layout features of the interface. Therefore, it seems that there is a need for a different approach to

researching web design features. A global approach described below may be an alternative method to investigate the layout of websites.

4. GLOBAL APPROACH ON RESEARCH INTO WEB DESIGN FEATURES

Dyson (2004) discusses the validity of some experimental research in which some typography variables are isolated. She reviews empirical studies on how physical text layout affects reading from screen. The text layout variables investigated were: columns, line length, interlinear spacing and window size. Dyson (2004) states that designers may not agree with the validity of studies that isolate some typographic variables in order to investigate the layout. The problem of isolating some typography variables is that there are variables that cannot be isolated. Lund (1999) also agrees with Dyson (2004) in that typographical variables should not be treated individually. He arguments that confounds may be produced when variables, apart from the one investigated, are kept invariant. For example, the line length is related to different variables: typeface, size of the typeface, and interline space. So it may not be possible to isolate some variables without producing a confound. As Dyson (2004) points out, ´as line length increases without a corresponding increase in interlinear spacing, the ratio between the two variables changes´. Therefore, in this situation it is not possible to know whether the effect is caused by line length or interline space. This may, therefore, produce studies with a lack of internal validity.

Dyson (2004) suggests two alternative approaches to solve this problem. One approach is to use systematic replications to test the relationship of different variables. This approach demands a huge effort as the effect of individual features should be tested in relation to other features, through systematic replications. In spite of the amount of work needed to test many different variables, Paterson and Tinker, summarised in the well-known book Legibility of Print (Tinker, 1965), used this approach for many years to investigate type size, line length, and interlinear space. They did these studies to test the legibility of printed media. This type of approach has not been yet investigated in digital media. In contrast, the number of possibilities has normally been constrained (Dyson, 2004).

According to Dyson (2004), another alternative approach to solving the problem of internal validity of experiments may be the global approach. It consists of comparing layouts which vary in a number of features, instead of isolating a few variables. This approach may be used to research

different aspects of an interface. This global approach was also described by Muter (1996). He called it the ‘kitchen sink’ approach.

A recent study has used this approach to investigate print material layouts (Lonsdale, 2005). This investigation used a combination of typographic features intended to improve legibility. It followed the legibility guidelines in order to created optimised layout and tested this with other layouts that did not follow legibility guidelines. The study found that the use of legibility guidelines can help users to search and answer questions from examination materials. Therefore, typographic variables should be considered in order to produce reliable examination materials.

A relevant point about the global approach is that it requires the use of more expertise than other types of methods. This is because more inconstant variables mean that there is a need for knowledge about designing the test material, as the inter-relationships should be considered. Therefore, the global approach contrasts with the experimental approach in the expertise required by the researcher. In experimental research, less expertise is necessary as only one or a few variables are tested. This leads to the conclusion that the global approach should be used to investigate the layout of websites by researchers with a background in design.

5. CONCLUSIONS

The case studies described in this paper showed that a combination of different methods may be needed to evaluate the effect of some layout features in websites. The findings showed that the different methods used in these studies (experimental research followed by a questionnaire, and think-aloud followed by an interview) are effective in finding the users' performance in and preference for finding information on web pages. In addition, think-aloud was useful as a complementary approach to the experimental research. Finally, these case studies showed the limitations of each of these methods. Therefore, different approaches may be needed in order to build up more effective knowledge about the effect of layout features of websites.

A possible different approach that may be used to investigate the layout of websites may be the global approach. This approach may be useful to get the users´ performance in using many variables and not treating them individually. This approach demands a deeper knowledge of typographical interfaces than experimental research. This is because the selection of level, the number of variables and their inter-relationships needs to be made very carefully, and therefore demands someone with expertise in the subject area (design). However, there are very few

studies using this approach. Consequently, more studies using this method should be done in order to check the validity of the findings and the usefulness of this type of research.

Finally, it is relevant to consider that most of the research into the layout of websites is done by computer scientists and not by design scientists. Zimmerman et al. (2007) in their paper ´Research through design as a method for interaction design research in HCI´ compare HCI research with design research. They show that there is a barrier between these two areas, and also that HCI scientists know little about what design researchers do. As mentioned above, expertise in the typographic area seems essential in order to formulate some types of study. Therefore, it seems that there is a need for more design researchers to investigate the effects of the layout of websites on users´ performance, preference and perception. The good news is that design research is in a period of expansion, with the establishment of many research centres, associations, journals, and conferences around the world. As Cross (2006) states: ´Forty years on, design research is alive and well, and living in an increasing number of places´.

ACKNOWLEDGMENTS:

Virgínia Tiradentes Souto was supported by a doctoral grant from CNPq – Brazil (National Council for Scientific and Technological Development).

REFERENCES:

Atwater, L. and Babaria, K. (2001). Controlled experiments. Choosing Human-Computer Interaction Appropriate Research Methods, University of Maryland. Retrieved 23 March 2004, from the World Wide Web: http://www.otal.umd.edu/hci-rm/cntlexp.html.

Benbunan-Fich, R. (2001). Using protocol analysis to evaluate the usability of a commercial web site. Information and Management, 39(2): 151-163.

Birns, J. H., Joffre, K. A., Leclerc, J. F. and Paulsen, C. A. (2002). Getting the whole picture - the importance of collecting usability data using both concurrent think aloud and retrospective probing procedures. Proceedings of 11th Annual UPA Conference - Humanizing Design, 8-12 July, Orlando, Florida, USA [CD-ROM].

Branch, J. L. (2000). The trouble with think alouds: generating data using concurrent verbal protocols. In A. Kublik (Ed.), Proceedings of the 28th Annual Conference, CAIS 2000: Dimensions of a Global Information Science, 28-30 May, Edmonton, Canada, Canadian Association for Information Science. Retrieved 10 May 2005, from the World Wide Web: http://www.slis.ualberta.ca/cais2000/branch.htm.

Buchanan, R. (2001) Design Research and the New Learning. Design Issues 17(4) 3-23.

Bunz, U. K. (2001). Usability and gratifications: towards a website analysis model. Peer-reviewed, Rutgers University. Retrieved 20 November 2004 from the World Wide Web: http://www.scils.rutgers.edu/~bunz/NCA2001usability.pdf.

Creswell, J. W. (2003). Research Design: Qualitative and Quantitative Approaches. London, Sage.

Cross, N. (2007). Forty years of design research. Design Studies, 28(1), 1-4.

Dumas, J. S. and Redish, J. (1999). A Practical Guide to Usability Testing. Exeter, Intellect.

Dyson, M. C. (2004). How physical text layout affects reading from screen. Behaviour & Information Technology, 23(6): 377-393.

Ericsson, K. A. and Simon, H. A. (1993). Protocol Analysis: Verbal Reports as Data. Cambridge, The MIT Press.

Holzinger, A. (2005). Usability engineering for software developers. Communications of the ACM, 48(1): 71-74.

Howell, D. C. (1982). Statistical Methods for Psychology. London, Duxbury.

Hughes, J. E., Packard, B. W. and Pearson, P. D. (1997). Reading classroom explorer: visiting classrooms via hypermedia. 46th Annual National Reading Conference Yearbook, Chicago, USA, National Reading Conference. Retrieved 18 May 2005 from the World Wide Web: http://www.readingonline.org/research/explorer.

Kerlinger, F. N. (1986). Foundations of Behavioral Research. London; Holt, Rinehart and Winston.

Kvale, S. (1996). Interviews: an Introduction to Qualitative Research Interviewing. London, Sage.

Lonsdale, M. S. (2005). Effects of the typographic layout of reading examinations materials on performance. Unpublished PhD Thesis, Department of Typography and Graphic Communication, The University of Reading, Reading, UK.

Mack, R. L. and Nielsen, J. (1994). Executive summary. In J. Nielsen and R. L. Mack (Eds.). Usability Inspection Methods. New York, John Wiley & Sons: 1-23.

Muter, Paul (1996) Interface design and optimization of reading of continuous text, in van Oostendorp, H. and de Mul, Ajaak, Eds. Cognitive Aspects of Electronic Text Processing, chapter 8, Norwood, N.J., Ablex: 161-180.

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

Nielsen, J. (1994). Estimating the number of subjects needed for a thinking aloud test. International Journal of Human-Computer Studies, 41(3): 385-397.

Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces. In J. Carrasco, J. Whiteside (Ed.), Proceedings of the ACM CHI 90 Human Factors in Computing Systems Conference, 01-05 April, Seattle, USA, ACM Press: 249-256.

Lund, O. (1999) Knowledge construction in typography: the case of legibility research and the legibility of sans serif typefaces. Unpublished PhD Thesis, Department of Typography and Graphic Communication, The University of Reading, Reading, UK.

Preece, J., Rogers, Y. and Sharp, H. (2002). Interaction Design Beyond Human-Computer Interaction. New York, John Wiley & Sons.

Shneiderman, B. (1998). Designing the User Interface: Strategies for Effective Human-Computer Interaction (3th edition). Reading, Addison-Wesley.

Souto, V. T. (2006). The arrangement of links in government websites. Unpublished PhD Thesis, Department of Typography and Graphic Communication, The University of Reading, Reading, UK.

Virzi, R. A. (1997). Usability inspection methods. In M. Helander, T. K. Landauer and P. V. Prabhu (Eds.), Handbook of Human-Computer Interaction. Oxford, Elsevier: 705-715.

Zimmerman, D. E., Akerelrea, C. and Buller, D. (2003). Lessons learned from building a health web site: implications for technical communicators. STC's 50th Annual Conference Proceedings, 18-21 May, Wyndham Anatole, USA. Retrieved at 10 March 2005, from the World Wide Web: http://www.stc.org/ConfProceed/2003/PDFs/STC50-046.pdf

Zimmerman, J., Forlizzi, J., Evenson, S. (2007) Research through design as a method for interaction design research in HCI. Proceedings of the SIGCHI conference on Human factors in computing systems, 28 April-3 May, San Jose, USA: 493-502.