Refereed Paper

The use of a heuristic process to evaluate an online information retrieval interface

S. M. Zabed Ahmed,
Cliff McKnight,
Charles Oppenheim

Authors

S. M. Zabed Ahmed, Department of Information Science and Library Management, University of Dhaka, Bangladesh

Cliff McKnight and **Charles Oppenheim**, Department of Information Science, Loughborough University, Loughborough

Email: c.mcknight@lboro.ac.uk

Abstract

This paper presents the results of a heuristic evaluation with the Web of Science interface. Three human factors experts carried out their independent evaluation. The findings were then analysed and combined to discuss them with expert members to reach a consensus on usability issues identified. The heuristic evaluation helped to identify a number of both positive and negative aspects in the Web of Science interface. The key strength of the then current interface was its consistency in terms of conventions used, screen layouts, minimum use of colours, and use of graphics and icons. The main weakness lay in its functionality, i.e., searching, navigation, online help, etc. The results show the effectiveness of a heuristic approach to evaluating user interfaces to online information retrieval systems.

Acknowledgement

We would like to thank the Commonwealth Scholarship Commission in the UK for its financial support to one of us (SMZA). The contributions of expert panel members in the heuristic evaluation are gratefully acknowledged.

Introduction

Heuristic evaluation (Nielsen, 1992; Nielsen, 1994) is a usability inspection method. It uses a short list of heuristic guidelines and a small number of evaluators. Each evaluator is given the heuristic guidelines to go through the interface independently to identify usability problems. All evaluations must be completed before the evaluators are allowed to communicate with each other. The theory behind this is that a single evaluator will miss out many of the problems in a user interface, but different evaluators will find different problems. Thus, much better results are obtained by combining the results from several independent evaluations.

Nielsen's (1994) list of heuristic principles has been frequently used in heuristic evaluation. This list contains the following ten usability heuristics that represent what any system with good usability is expected to have:

(1) visibility of system status;
(2) match between the 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 user recognise, diagnose, and recover from errors; and
(10) help and documentation.

This paper describes the use of heuristic processes to evaluate the usability of the Web of Science interface using Nielsen's heuristic principles. Three human factors experts carried out their independent evaluation first, and then we analysed and combined the findings according to one of the heuristic principles. The main objectives were to: (a) assess the strengths and weaknesses of the Web of Science interface, and (b) assess the power of heuristic evaluations.

Related research work

There are two main approaches to considering usability evaluation of any interface: empirical and analytical. Empirical techniques involve testing with real users, whereas analytical techniques require experts assessing usability using established theories and methods. The heuristic evaluation method has been applied to evaluate the usability of several traditional computer interfaces. In an early heuristic evaluation, Nielsen and Molich (1990) found that individual evaluators performed quite badly in finding usability problems in the interfaces they evaluated. The proportion of problems each evaluator found varied between 20% and 51%. The aggregated results from several individual evaluators, however, showed that a group of three to five evaluators working separately would find between 55% to 90% of the problems. Based on this, Nielsen and Molich recommended using heuristic evaluation with three to five evaluators. In a later study, Nielsen (1992) found that "double" usability specialists (expert in both usability evaluation and the interface to be evaluated) were better than naive or regular usability specialists for conducting a heuristic evaluation. Some studies also compared the effectiveness of different evaluation methods in finding usability problems in traditional interfaces (e.g., Jeffries et al., 1991; Desurvire et al., 1992; Cuomo and Bowen, 1994; Consolvo and Towle, 2005). These studies showed that heuristic evaluation found more problems than other techniques. It also proved to be the most cost-effective method.

The heuristic evaluation method has been applied to assess the usability of several library web sites, digital libraries, OPACs, and IR interfaces. Mangiaracina and Marchetti (1998) described the heuristic evaluation of GUI-based EINS interface. The findings of the evaluation were then used to design the EINS-Web interface. Warren (2001) discussed the heuristic evaluation of URICA OPAC system using published user interface guidelines. The evaluation showed a number of usability problems in the interface. Peng et al. (2004) applied heuristic principles in order to develop a questionnaire to assess the usability of GEMS interface at Nanyang Technological University, Singapore. This questionnaire was surveyed with 88 students, although heuristic evaluation should involve few usability experts evaluating elements of an interface against a checklist of heuristics or design principles. Some studies focused on creating prototypes for library web sites, OPACs, and IR systems in order to obtain feedback on designs by means of heuristic evaluation (e.g., Van House et al., 1996; McMullen, 2001; Allen, 2002). The final version of the interface design evolved after few iterations and user testing. Several studies also used multiple evaluation methods in predicting usability problems in library environments (Doubleday et al., 1997; Cogdill 1999; Dickstein and Mills, 2000; Blandford et al., 2004). These studies showed that the results of heuristic evaluation are both valid and useful.

Methods used

Three human factors experts conducted their individual heuristic evaluations of the Web of Science interface at Loughborough University, UK. They were all "double" specialists, i.e., experts in both usability engineering and the Web of Science interface to be evaluated. They were all provided with Nielsen's ten heuristic principles to guide their individual evaluations. Experts were told not to discuss among themselves the problems they found during their evaluations. They were told to look for usability issues in the Web of Science interface where they are confused or feel the users would be confused. They were also told to suggest a solution to the problem identified, if possible. After their independent evaluation, the findings were analysed and combined to discuss them with the expert members to reach a consensus on the usability issues that emerged from their independent evaluations.

Results of the evaluations

The goal of the heuristic evaluation is to cite violations of usability heuristics in a user interface. The result of the evaluation is thus a list of usability problems in the Web of Science interface. There were also numerous aspects of a good design in the interface identified by the

experts. We first discuss the positive features, followed by the findings of usability problems in the Web of Science interface.

Web of Science: positive features

The experts' comments on the Web of Science interface were generally favourable. They all noted that the interface showed a general appreciation of HCI issues, usability requirements, and the tasks that its users wish to accomplish. The homepage described what services are available. The main texts as well as the texts associated with each link were written in a style that was readily comprehensible to most users.

The Web of Science featured a simple layout, maintained a high degree of internal consistency, and was carefully organised. It followed standard conventions and most information appeared in a natural and logical order. The use of graphics was conservative thus minimising the time needed to download pages. Overall, it maintained a consistent look and feel throughout the interface. The fonts, font sizes and colours were consistent across pages. There was consistency in the use of key terms. The Web of Science banner appeared in a consistent position all over the interface. Likewise, the text justification was consistent across pages. Navigation throughout the Web of Science was also relatively consistent. All of the pages had "Home", "Help" and "Logoff" buttons at the top. The interactive principle of web browsing was readily suited to the information search process in the Web of Science.

The interface used a conventional form fill-in style. This type of interface is considered to be convenient for naive users as all options are presented in their context, and the users only need to fill in the relevant boxes. The Full Search option allowed sophisticated search queries for improved and better-targeted results. A search on topic, person or place displayed papers that satisfy the search criteria. The title was hyper-linked to the full record, which included links to "Cited references", "Find Related Records", and "Times Cited" options. The first revealed the list of references that the author had cited in the original article. Many of these references have links associated with them, which, if activated, revealed the full record. This process could be continued repeatedly to explore deeper into the knowledge base. The "Find Related Records" link retrieved articles whose reference list included at least one of the sources cited by the original article. The "Times Cited" option indicated how many times the current article had been cited by other papers. This link could be activated to reveal exactly who had cited the original work.

The Web of Science was well documented. The "Help Contents" button in each section of the help facility was very useful. Moreover, searchers had the opportunity to email feedback about their experiences with the system.

Web of Science interface: violations of usability heuristics

The experts identified a number of usability problems in the Web of Science interface. All usability problems found were analysed against one of Nielsen's ten heuristic principles. Suitable solutions to the problems were also identified where possible. Sometimes solutions could be drawn from the nature of the problem itself.

Visibility of system status. The interface used selectable coloured buttons for users to issue basic commands, i.e., "Home", "Help", "Log off", etc. These buttons, however, did not give strong feedback that they were selected. The searchers may not realise that they have not accurately selected a button for some time if the network is slow. Also, there was little feedback when the system responded to a search request, except the Windows hourglass. A representation of activity would be helpful, although it is difficult to indicate the actual rate of progress or task completion time since network traffic volumes may affect all these. The search screens were divided into horizontal bands which either provided information, allowed users to enter the query terms, or input commands. These bands could have been differentiated by colours to emphasise the actions required by the users.

Match between system and the real world. The interface was extremely modular, and users were forced to choose one route or another without necessarily understanding its relationship to their task. The different purposes and extents of the various "Search" options should have been clearly explained. Besides, the Easy Search did not offer any significant advantages over the Full Search and in many respects would frustrate even naive searchers. This was especially true in cited reference searching since all the citing papers would be mixed together. Moreover, the date limit in the Full Search was based on the date the article was added to the database rather than its date of publication. Thus, limiting search to 2002, for example, would ignore many of the articles published in 2002.

The Easy Search and Full Search options asked users to specify the database(s) to be searched before entering the query terms. This was counter intuitive. The user would presumably choose to expand the scope of the initial search if it was not successful. The interface did not allow users to change databases without going out from the search session and back in from either Easy Search or Full Search screen. The Place/Address searching required searchers to use abbreviations, for example, "dept" for department, "univ" for university, etc. The help screens provided users with a long list of common abbreviations. Similarly, "Cited Work" included the name of the original source document which was abbreviated. This list was also very long, and users had to copy and paste (or type) the name from the list into the search box. The Web of Science should have made these features easier to handle for naive users.

The interface used information science jargon, e.g. "citation index", "cited references". Experienced searchers may well be familiar with these terms, but more background information could have been given for naive users. The Easy Search button should have appeared at the top of the list in the homepage rather than the Full Search. Also, the icon displaying the Web of Science banner on the homepage provided no real information.

The search results did not show the number of hits per term in a query. The users had to search for individual keywords if they wanted to undertake sub-set searching. In addition, the word "Summary" on the top of the results screen was rather inappropriate.

This implies a summary of the articles retrieved whereas it actually displayed a list of retrieved records. Likewise, the use of the word "Lookup" in cited reference search was unusual. It has, incidentally, been dropped in the latest version of the interface, launched in Summer 2005. The results pages allowed users to mark records for later printing or downloading. However, it was not obvious whether the users had to "submit" for each page, or whether they could "submit" at the end after marking on several pages. A short overview of the "search/mark/submit" process would have been helpful.

User control and freedom. There was no clear path for refining/modifying a search query. The user had to navigate back to the search screens using the browser's "Back" button. A general navigation bar within the pages would have been useful. Moreover, a text-based version of the interface should have been built in parallel with the existing one for users to choose if they wish.

Consistency and standards. The command buttons were given a consistent location which was helpful, but their absolute position changed depending on the number of buttons displayed on a given screen. This could be confusing for novice searchers. The number of results for a search appeared at the bottom of the results screen. It would have been useful to see the number of results at the top of the search results as well. Moreover, the search result pages displayed title information sometimes in upper and sometimes in lowercase letters. It was not clear why some titles were displayed in uppercase and some records were not highlighted in the cited reference list.

Error prevention. The interface should have offered an emergency exit button in order to get out from the system anytime. This exit button would also encourage naive users to venture

onto more advanced options since they could always try out unknown options, knowing they have the ability to get out of trouble without repercussions. Moreover, every action should be reversible so users can go back to a previous state in a session.

Recognition rather than recall. The users were constantly forced to take initiatives due to the lack of prompts or guidance in the interface. The availability of various search operators (Boolean, proximity, truncation, etc.) could have been presented visually to the users by including the operators as options to select in a template.

Flexibility and efficiency of use. The search results were shown in blocks of 10 records at a time. The interface supplied no way to display all these records at one time in full format. The Web of Science interface should have offered more flexibility in designating the number of records users want to be displayed and in which format. The Full Search allowed users to save searches that could be used in a later session. It would have been useful if the saved query could have been run automatically and results sent via email as an alerting service, much as the old ISI ASCA Service used to do.

Aesthetic and minimalist design. The Web of Science navigation was made difficult by the inability to move easily between screens, and having too many screens. The navigation skipped certain levels in the hierarchy. For example, the homepage gave access to Easy Search and Full Search. Topic, Person and Place searches were accessible from Easy Search. It is important, however, that users could access Easy Search from Topic search. Navigation throughout the Web of Science required all users to return back to the homepage. It was also impossible to move from Easy Search to Full Search; rather the user needed to return to the homepage and then navigate down to Full Search. The Full Search screen was cluttered and contained too much information.

Help users recognise, diagnose, and recover from errors. The treatment of error messages in the Web of Science was generally inadequate. All error messages should be specific and should offer no more technical detail than necessary. The error messages should have listed search tips for effective query formulations and provided a link to online help for further guidance. The error message should also have provided a link back to the search page to modify the query.

Help and documentation. The online help option, when available, was difficult to scan and extract information from. Task-oriented online help should have been included, so that users could have found the best way to look for what they need. The "examples" link took users to a section in the full help. It would be useful to see a short page of examples with a link to the full help.

The Web of Science treated multiple words in a topic search as a phrase. It was not obvious from the online help that the users need to type in brackets as part of a search query, e.g., "(A or B) and C" or whether phrases, e.g., "Information Technology" would count as a single search term without needing enclosing brackets or quotes. Examples on this would have been useful.

Discussion and conclusion

It is clear that our heuristic evaluation provided an important critique of the then current Web of Science interface. Based on our experience with developing a new methodology for user-centred design and evaluation of IR interfaces (Ahmed et al., 2006), we recommend using both qualitative and quantitative methods for evaluating interfaces. We suggest heuristic evaluation for an interface in its early stages to find major problems. Usability testing is recommended for use later in the development process with the final version of the interface design.

The heuristic evaluation helped us to identify both positive and negative aspects of the Web of Science interface. The key strength in the interface is its consistency in terms of conventions used, screen layouts, minimum use of colours, and use of graphics and icons. The

main weakness lay in its functionality, i.e., searching, navigation, online help, etc. These would certainly affect users' search performance and their satisfaction with the interface negatively (Ahmed et al., 2004; Ahmed et al., 2005).

We believe that this paper demonstrates the value of adopting a heuristic approach to the evaluation of online information retrieval interfaces and commend the use of such methods to all those developing such interfaces in the future.

References

Ahmed, S. M. Z., McKnight, C. and Oppenheim, C. (2004) A study of users' performance and satisfaction with the Web of Science IR interface. *Journal of Information Science*, 30(5), 459-468.

Ahmed, S. M. Z., McKnight, C. and Oppenheim, C. (2005) A study of learning and retention with a Web-based IR interface. *Journal of Librarianship and Information Science*, 37(1), 7-16.

Ahmed, S. M. Z., McKnight, C. and Oppenheim, C. (2006) A user-centred design and evaluation of IR interfaces. *Journal of Librarianship and Information Science*, 38(2) [in press]

Allen, M. (2002) A case study of the usability testing of the University of Florida's virtual library interface design. *Online Information Review*, 26(1), 40-53.

Blandford, A., Keith, S., Connell, I., and Edwards, H. (2004) Analytical usability evaluation for digital libraries: a case study. In: *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries*, 7-11 June, Tucson, AZ. New York, ACM, 27-36.

Cogdill, K. (1999) *MEDLINEplus Interface Evaluation: Final Report*. Unpublished Report, Human-Computer Interaction Lab (HCIL), University of Maryland.

Consolvo, S., and Towle, J. (2005) Evaluating an ambient display for the home. In: *CHI '05: Extended Abstracts on Human Factors in Computing Systems*, 2-7 April, Portland, OR. New York, ACM, 1304-1307.

Cuomo, D. L., and Bowen, C. D. (1994) Understanding usability issues addressed by three user-system interface evaluation techniques. *Interacting with Computers*, 6(1), 86-108.

Desurvire, H., Kondziela, J., and Atwood, M. (1992) What is gained and lost when using evaluation methods other than empirical testing. In: Monk, A., Diaper, D., and Harrison, M. D., eds. *Proceeding of the HCI '92: People and Computers VII*, 15-18 September, York, UK. Cambridge, Cambridge University Press, 89-102.

Dickstein, R., and Mills, V. A. (2000) Usability testing at the University of Arizona Library: how to let the users in on the design. *Information Technology and Libraries*, 19(3), 144-151.

Doubleday, A., Ryan, M., Springett, M., and Sutcliffe, A. (1997) A comparison of usability techniques for evaluating design. In: DIS '97: *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, 18-20 August, Amsterdam, The Netherlands. New York, ACM, 101-110.

Jeffries, R., Miller, J. R., Wharton, C., and Uyeda, K. M. (1991) User interface evaluation in the real world: a comparison of four techniques. In: CHI '91: *Proceedings of the Conference on Human Factors in Computing Systems - Reaching through Technology*, 27 April-2 May, New Orleans, LA. New York, ACM, 119-124.

Mangiaracina, S. and Marchetti, P.G. (1998) WWW interface design, driven by heuristic evaluation: the EINS-Web Project. In: *Proceedings of the 8th DELOS Workshop on User Interfaces in Digital Libraries*, Stockolm, 18-21 October. Available at: www.ercim.org/publication/ ws-proceedings/ DELOS8/marchetti.pdf [28 September 2005]

McMullen, S. (2001). Usability testing in a library Web site redesign project. *Reference Services Review*, 29(1), 7-22.

Nielsen, J. (1992) Finding usability problems through heuristic evaluation. In: *CHI '92: Proceedings of the Conference on Human Factors in Computing Systems*, 3-7 May, Monterey, CA. New York, ACM, 373-380.

Nielsen, J. (1994) Heuristic evaluation. In: Nielsen, J., and Mack, R. L., eds. *Usability Inspection Methods*. New York, John Wiley, ISBN 0471018775

Nielsen, J., and Molich, R. (1990) Heuristic evaluation of user interfaces. In: *CHI '90: Proceedings of the Conference on Empowering People: Human Factors in Computing Systems - Special Issue of the SIGCHI Bulletin*, 1-5 April, Seattle, WA. New York, ACM, 249-256.

Peng, L. K., Ramaiah, C. K., and Foo, S. (2004) Heuristic-based user interface evaluation at Nanyang Technological University in Singapore. *Program: Electronic Library and Information Systems*, 38(1), 42-59.

Van House, N., A., Butler, M. H., Ogle, V., and Schiff, L. (1996) A user-centered iterative design for digital libraries: the Cypress experience. *D-Lib Magazine*. Available at: http://www.dlib.org/dlib/february96/02vanhouse.html [5 October 2005]

Warren, P. (2001) Why they still cannot use their library catalogues. *Proceedings of the Information Science Conference*, Krakow, Poland, 19-22 June. Available at: proceedings.informingscience.org/IS2001Proceedings/pdf/WarrenEBKWhy.pdf [28 September 2005]