Journal of Universal Computer Science, vol. 4, no. 4 (1998), 449-460
submitted: 10/12/97, accepted: 28/12/97, appeared: 28/4/98 © Springer Pub. Co.

Evaluation of On-line Help

Ise Henin
(University of Victoria, Victoria, BC Canada
ise@uvic.ca)

Abstract: This paper looks at a variety of on-line help systems and at guidelines for their design; and indentifies general problem-solving strategies which are important for the effectiveness and usability of on-line help. The lack of a suitable evaluation instrument is identified and a questionnaire to address this need is developed: the *On-line Help Evaluation Checklist.* The new instrument is to assist instructional designers (who develop courses that require computer problem-solving skills of the target audience) to assess the adequacy of a tool's on-line help. The instrument is subsequently applied to the evaluation of software tools to be used in a first-year, university-level course on instructional instrumentation.

Category: Methodological

Keywords: on-line help; evaluation; instructional design; navigation; hypertext;

1 Introduction

This study deals with computer documentation, specifically, with on-line help. Rarely does someone who regularly works with computers admit to reading a computer manual or using a software program's on-line help, yet everybody has problems or questions while working on a computer. Software used to be supplied with printed documentation covering several yards of wall space on dusty shelves, today the application and its documentation may fit on a CD-ROM and is often supplied without a paper-based manual and with minimal printed documentation. Tutorials, reference materials and guides for trouble-shooting are more and more frequently computer-based, either supplied with the program or available in electronic form—such as from the developers' web site for downloading. All such materials, and a variety of others, can be referred to as on-line documentation, with on-line help belonging to the subset of documentation that is to assist with computer problem solving.

1.1 The Context

The setting for the study is (a) to identify types of on-line help systems; (b) to review current pertinent literature on problem-solving with (and the design of) on-line help; (c) to develop, from a theoretical framework, an approach to evaluate on-line help; (d) to apply the approach to evaluate the on-line help supplied with software tools identified for use in a first-year, university level, instrumentation technology course; and (e) to recommend strategies for supporting the students with suplementary on-line materials where the tools offer ineffective on-line help.

1.2 The Focus

The focus of this report is to summarise (a) and (b) in the context of the problem, and to describe the steps taken in (c) that have led to the development of an on-line help evaluation instrument used in (d) and (e) as part of a needs analysis for course development.

2 Method and Procedure

This project combines a theoretical framework and a practical application, in which literature is reviewed and design tools are evaluated, in order to establish a framework for developing an instrument for solving a specific educational problem.

2.1 Theoretical Framework

The theoretical framework, based on a literature review, presents research into problem-solving behaviour of computer users, a detailed review of design guidelines developed specifically for on-line help, and identifies an evaluation instrument.

2.2 Preliminary Investigation

The preliminary investigation phase, during which a variety of on-line help systems are examined and their characteristics described, concentrates on the problems encountered while trying to use the, somewhat dated, on-line help evaluation instrument identified in the previous section. This investigation has uncovered, inter alia, the need to improve the instrument, and develop a new tool for use within this context.

2.3 Design and Construction

The design and construction phase is approached in three stages: during the first phase, a new evaluation instrument is developed in draft form; this phase is followed by testing and evaluation, during which various evaluators are observed applying the instrument and their comments are recorded; and a revision phase, during which the observations made during testing are incorporated.

2.4 Implementation

During the implementation phase, the new instrument, the *On-line Help Evaluation Checklist,* which was to form the basis for a needs analysis, is applied to evaluate the on-line help for the software tools to be used in Instrumentatietechnologie 1 (ISM-1) a course that was under development for the 1995-96 academic year at the Educational Technology faculty of the University of Twente.

3 Results and discussion

In this section the results of the literature review, the preliminary investigation and the design and construction phase are presented, and a short summary of how the new tool is applied in the implementation phase is discussed.

3.1 Definition of the problem

The students in the new version of the ISM-1 course will need to become proficient in a short period of time with several, often quite complex, computer software programs, and acquire the skills relatively independently even though they are not expected to have prior experience in the use of computers. But to acquire skills, students have to have access to learning resources. Can the resources for learning to use software effectively be found in the on-line help already associated with the applications, or do instructional materials have to be specifically provided for ISM-1 students?

3.2 Literature review

Literature shows that on-line help is a tool for assisting computer users in resolving problems quickly [Duffy, Mehlenbacher, and Palmer (1992)]. On-line training materials—which support the goal of learning—can guide a learner through a series of exercises to illustrate a concept and promote understanding. On-line reference materials—which provide an exhaustive treatment of a given subject—are useful only to those wishing to invest time to understand a subject at a certain level of depth. On-line help, on the other hand, is designed to answer the question "How do I?" and its goal is to support performance, not broad-based learning. On-line help assists in error correction rather than detection, and access to the system is always user initiated. Therefore, according to the authors mentioned above, on-line help has to be targeted to relate directly to the task in question; be accessible in an efficient manner; and facilitate transfer from the help system to the problem task.

Research that looks into the behaviour of people consulting printed manuals can be relevant in this context, even though it is not conducted with on-line help in mind, because it examines problem solving from a user's perspective. Certain design methods for incorporating problem-solving information in user guides are applicable to on-line help, as well. Designing manuals, using a "minimalist" approach, is discussed by Lazonder (1994), and Lazonder and van der Meij (1995) who stress that presenting problem-solving information in a way that facilitates detection, diagnosis, and correction of errors improves performance and corrective error-handling skills. Minimalist documentation provides extensive problem-solving information and strategies for its display, positioning, and indexing.

The issue of problem-solving with computer documentation has been studied in detail by van der Meij (1996[1]). The author suggests that the reason a person who encounters a problem prefers to ask someone about solving it, rather than consulting a manual or the on-line help, is the ability to negotiate meaning: defining the problem is seen as the most difficult part of finding a solution. Van der Meij presents a general model for problem solving which outlines three stages: experiencing the problem (which includes seeing it and deciding to address it);
expressing the problem (which requires defining it, deciding on a source and selecting a search method); and processing the problem-solving information (which includes extracting and evaluating the information followed by solving the problem).

A model for the design and evaluation of on-line help, which lists 22 designgoals in eight tasks, is presented by Duffy, et al. (1992) with the focus on:

- the importance of target-audience analysis since the designer must assume a level of prior knowledge and must state which information is

[1] The article cited in this report was published in September 1996, but when the study took place in the summer of 1995, the version accepted for publication on July 12, 1995 was used.

deemed a prerequisite to understanding

- the importance of supporting a varied vocabulary, defining terms implicit in the vocabulary, and providing non-technical terms and a wide array of synonyms

- the provision of constant, fixed entry points to the system; and multiple access methods, such as through keyboard shortcuts and contents maps

- the need to facilitate the scanning process by providing either scrolling fields or a paging mechanism

- the provision of concise, goal-oriented, and task-based contents with elaboration and procedural information provided only on request

- the need to bring the information to the user rather than requiring the user to search for it

- the necessity of user testing of the navigation method

- the need for concurrent availability of the application and its help system to facilitate the transfer of the problem-solving information to the task

While design guidelines for on-line help systems exist, information on their usefulness is scarce. There appear to be few, if any, instruments for evaluating on-line help. Duffy, et al. acknowledge the need to evaluate on-line help, and, lacking a suitable tool, have developed an instrument of 42 questions, titled the *Help Design Evaluation Questionnaire* (HDEQ) for the evaluation of help by software developers. Shneiderman (1987) has developed a generic (short) and a detailed (long) version of an instrument for the evaluation of interactive computer systems, but since these questionnaires are concerned mainly with the human computer interface of the evaluated software tool itself, there is not enough emphasis placed on the on-line help aspect to be applicable to this study.

Both tools are designed for formative evaluation during the design process, a time when revisions are still possible. Once the end-user has access to the tool, the evaluation becomes a summative one, and many items become irrelevant since it is too late to make changes to the on-line help tool.

The previous section outlines general problem-solving strategies that are incorporated in effective on-line help systems, and identifies instruments that might be useful for a system's evaluation. The next section looks at types of help systems and describes how the HDEQ instrument was used for an evaluation of a complex help system.

3.3 Types of on-line help systems

On-line help can be grouped into three categories, systems that explain screen functionality, based on the “bubble” dialogue technique borrowed from cartoon strips; systems that mirror printed documentation that might include some additional, rudimentary navigation tools; and systems that use hyperlinks and search mechanism to exploit the advantage of the computer for accessing large amounts of problem-solving data.

A system using the dialogue method, defines “hot spots” in such a way that, when the pointing device is moved over an item, the item is explained, either in a “bubble” at the screen-location of the item or in another, predefined, area. Disadvantages of the method are that the display may disrupt the task, since the mouse must remain at exactly the right location for the text to remain visible; that the information may be hard to find repeatedly since it is not always obvious which “hot spot” generates the display; and that there is no way of keeping the information on the screen since the mouse is activating the “bubble” instead of the object that the help is sought for. Although systematic information searches are not available, this type of help can, nevertheless, be sufficient for thelimited scope of very small applications.

The print-based systems consist of manuals, previously or simultaneously provided in print form. The information is usually well formatted and, once printed, can be organized with the user’s own notes; it may also be more detailed and complete than had it been designed for the screen. Readability, however, may leave much to be desired, since a font, to be readable on the screen, has to be much larger than a printer-destined font; usability is further limited, when the information is presented in scrolling windows, that cannot keep enough information visible at any one time to enable the user to apply the instructions to the problem.

A hypertext-based, on-line help system takes full advantage of interactivity and offers highlighted words or objects that, upon selection, cause additional information to be presented. Such a system can present information in any form (such as text, graphics, and animations) and its key feature is organisation from general to detail: information is stored in small enough units to be displayed in a minimal chunk of information at the lowest hierarchical level. This type of system requires good navigation tools; it is important that one can find one’s way back to the original starting point; and sophisticated systems supply “site maps” that let the user know which area, in relation to the overall system, is currently being accessed.

After identifying the common types of help systems, a complex, hypertext-based help system is selected for performing a trial evaluation using the the HDEQ instrument identified in the literature review.

3.4 Applying the *Help Design Evaluation Questionnaire*

To test the HDEQ questionnaire, it was used to evaluate the Macintosh Guide supplied with Apple's MacOS 7.5.2 operating system. The guide lists over 300 topics, has an underlying educational design, and is well integrated into the computer platform's standard way of providing help. It has been given good reviews, and several ISM students have indicated its usefulness, therefore one would expect a high evaluation rating as result.

The evaluation of the Macintosh Guide with HDEQ took over an hour and required several hours prior learning to understand the system sufficiently to be able to answer the questions. The guide as tested, achieved a score of .79; the maximum obtainable score is 1. The guide ranked high in contents, comprehension and link to application, but rather low on navigation, format, and menu selection, these results were rather suprising.

The HDEQ does not appear to be well suited to help systems based on the graphical user interface (GUI) prevalent today. A system will rank low (thus inferior) if it does not contain lots of text and many navigational buttons, even if it incorporates a good search mechanism and is well designed from a human computer interface viewpoint. Since HDEQ appears to have been developed to evaluate character-based screens, it rates a product that displays a profusion of information higher than one that displays only the minimum amount necessary. The highest rating goes to a screen that lists 15 to 50 initial choices, and nests 7 to 15 levels of hierarchy. While a menu with 15 to 50 lines of text might (just) be readable, one would not consider a screen with 50 icons, hyperlinks or buttons to be well designed. It is hard to imagine someone moving 15 levels down to find help information, especially since there are no assurances that the information actually exists.

Shortcomings such as the above, excessive length of the instrument, and the requirement that the evaluator be an expert to test every aspect of the selected tools' on-line help made it clear that this tool was not suitable for the practical part of this project.

3.5 Design of the *On-line Help Evaluation Checklist*

While the designers of HDEQ have assumed that end users would neither have the expertise nor the time to evaluate a help system, the premise at the outset of this new design process for the *On-line Help Evaluation Checklist* (see Figure 1) is that course developers may be interested in an evaluation tool if they perceive the tool as useful, easy to understand, and not requiring a large time investment.

The process of defining the questions went through several iterations, for simplicity; a "yes/no" format was to be used. The 22 goals identified by Duffy et al. (1992) were reworded into 30 questions; 40 guidelines added that were derived from Stevens & Stevens (1995); 10 ideas were incorporated from

Collis & Verwijs (1994), van der Meij (1996), and Lazonder (1994); and 20 questions pertaining to search criteria (an area neglected in the literature) added by the author. These 100 questions were placed in the categories contents, screen design and navigation, and search capabilities; ranked by importance; and the top 15 questions in each category selected for inclusion in the draft instrument.

Testing of the instrument on the Macintosh Guide (where a high rating would be expected), an overall rating of 89% was attained, slightly higher than that attained with HDEQ, even though the new instrument is much shorter, and requires no subject matter expertise. While encouraging, the results cannot say much about the draft instrument's reliability or validity, but the observations could be used for fine-tuning, since the procedure uncovered that word order and choice of words were problematic, that there were no instructions, and that there were no fields to tally the results.

The time-frame for this project was too small to perform extensive statistical analyses on tests of the draft instrument, or to test it with a wide range of products. Therefore, for the formative evaluation, seven members of the faculty and staff were approached, including a subject matter expert in the area of on-line help, to use the instrument for an evaluation while their observations would be recorded and used to improve the instrument. The focus was to be on contents, language, and ease of use. Participants were asked to read the "what to do" column before proceeding, and told that the "notes" and "examples" columns were intended for reference, if and when needed. The participants were then asked to start the software tool, locate the on-line help, and spend a few minutes familiarising themselves with the system before proceeding to attempt to answer all applicable questions. The importance of deciding on yes/no answers was emphasised.

The most important design flaw that became evident during the testing phase was that the "yes" or "no" format of the evaluation instrument's scale did not adequately represent the choices the evaluator might wish to make. Leaving a question blank could mean that an item is not applicable, that the evaluator does not know, or that the evaluator does not consider the question important within the context of the evaluation. Adding a column "not applicable or not important" (abbreviated to "n/a") would remove this ambiguity: If "n/a" is selected, the question is not counted in the tally of the results. This would prevent a non-response from skewing the results. When evaluators were asked to calculate the ratings by following the listed instructions, an error in the instructions was uncovered. Other changes were made as result of ambiguity in wording, and some questions were reordered.

After incorporating all corrections and suggestions for improvement, and rewriting the instructions on the back of the form, the instrument was finalised (see Figure 1) and contains 45 questions relating to contents, navigation and visual design, and search capability. Checkboxes for the questions were added for "yes", "no," and "n/a" (not applicable). The results

On-line Help Evaluation Checklist

Software Package	Version Number	Platform & Operating Operating System	Evaluation Date & Time	Evaluator	Rating in Percent	Evaluation Purpose
Where installed						Instructions or review.

Quality of Contents & Navigational Access to Contents	N/A	Yes	No
Is starting the on-line help (OLH) easy?	☐	☐	☐
Is closing help or stopping OLH easy?	☐	☐	☐
Is using the OLH easy to figure out?	☐	☐	☐
Is there "help" on help?	☐	☐	☐
Are the contents presented in a structured manner?	☐	☐	☐
Is there a glossary or definition of terms?	☐	☐	☐
Are tips and elaboration presented?	☐	☐	☐
Are there tips and elaboration only on request?	☐	☐	☐
Are good examples given?	☐	☐	☐
Are the concepts demonstrated step by step?	☐	☐	☐
Can one practise the demonstrated concepts?	☐	☐	☐
Does the software do what its OLH says it can do?	☐	☐	☐
Does the system make work easier?	☐	☐	☐
Does OLH present tasks you may need?	☐	☐	☐
Does OLH encourage a problem-solving approach?	☐	☐	☐
Overall, do you consider the contents sufficiently detailed for this type of software tool?			
Subtotals:	☐	☐	☐

Design of Screen Interface	N/A	Yes	No
Is screen well designed and orderly?	☐	☐	☐
Do graphics add to functionality, not merely decorative?	☐	☐	☐
Does the navigation allow non-linear access?	☐	☐	☐
Do sequential access tools look different from linking tools?	☐	☐	☐
Are navigational buttons in same place on all screens?	☐	☐	☐
Do buttons and sensitive spots you have tried link correctly?	☐	☐	☐
Do icons or buttons clearly communicate their function?	☐	☐	☐
Is text used in tables clear?	☐	☐	☐
Are metaphors (if used) understandable?	☐	☐	☐
Is the smallest chunk of information easily accessible?	☐	☐	☐
Can windows be minimized, or resized, or moved?	☐	☐	☐
Can windows be kept open while working?	☐	☐	☐
Can a window be printed?	☐	☐	☐
Can all information on a topic be printed?	☐	☐	☐
Can information be saved?	☐	☐	☐
Overall, do you consider screen and interface design good enough for this type of software tool?			
Subtotals:	☐	☐	☐

Search Capabilities	N/A	Yes	No
Does OLH support linear access (browsing)?	☐	☐	☐
Does OLH present a list of topics?	☐	☐	☐
Are the topics organised by tasks?	☐	☐	☐
Can one scroll or page through all topics?	☐	☐	☐
Does typing/clicking a letter reach the topic?			
Is an index in alphabetical order provided?	☐	☐	☐
Can one scroll or page through the index?	☐	☐	☐
Does typing/clicking a letter reach the index item?	☐	☐	☐
Does the system support "How do I?" searches?	☐	☐	☐
Does OLH support searching by keyword?	☐	☐	☐
Is search performed on topic titles?	☐	☐	☐
Is search performed on contents?	☐	☐	☐
Can one search for two keywords ("print" "error")?	☐	☐	☐
Is (or/not) search possible?	☐	☐	☐
Can one search for partial string ("mount")?	☐	☐	☐
Overall, do you consider search capabilities sufficient for this type of software tool?			
Subtotals:	☐	☐	☐
Totals:	☐	☐	☐

Appendix G. On-line Help Evaluation Checklist

Figure 1: On-Line Help Evaluation Checklist

can be tallied and a percentage score obtained: the higher the score the more functional the on-line help. Three open-ended questions allow the evaluator to limit or exclude areas from the evaluation that are not pertinent to its purpose.

The new instrument can provide a general picture (albeit a non-representative one, due to the lack of testing) of the strengths and weaknesses of the help system under evaluation. In the following implementation phase the new instrument is applied within the context of a needs analysis.

3.6 Implementation

The tools evaluated with the OLHEC instrument are Hypercard, First Class Client, Deskscan, Aldus Superpaint, and Macro Media Soundedit Pro. Two help systems rated around the 75% mark, two around 50%, and one around 25%. From the exact scores and the answers to the open-ended questions, it can be seen that the higher-ranking products generally are considered sufficient, with only some desirable features missing. Lower ranking items lack at least one of the three categories—contents, interface design, or search capabilities. Searching is the most problematic area for all tools evaluated. One tool facilitates problem-solving by providing a glossary; since users often do not know the right technical term to search for, chances of finding the needed information is greatly improved.

The project was completed by providing an assessment of the software's difficulty levels and assigning priorities based on both difficulty of the tool and the availability of adequate on-line help, for the development of additional on-line learning resources for the ISM-1 course.

3.7 Further study

To complete this study, the *On-line Help Evaluation Checklist* still needs to be subjected to statistical analysis to test the validity and reliability of the instrument, this should include an inter-rater reliability analysis.

Considering how many of the evaluators have commented on the fact that they either do not use on-line help or never find what they are looking for, further research on evaluation, accessibility, and usability of on-line help products is needed to uncover why the perception persists that on-line help is of so little use to most computer users.

An interesting area for further research is the possibility of widening the definition of on-line help from a self-contained unit within a software program to any help information related to a compter problem, irrespective of the information's location. The instrument's applicability to the evaluation of materials found on the World Wide Web, such as frequently-asked-question archives, company fax-back services, and technical databases could be reviewed, tested, and possibly redesigned. The three sections of the instrument could be applied independently, with one part of the questionnaire used for evaluating contents in databases; the second part used for evaluating the functionality of the navigational and design aspects of browsing tools; and the third section used for testing the probability of search engines providing the desirable functions that help users solve computer problems without flooding them, incidentally, with irrelevant subject matter.

4 Conclusion

This study has uncovered that on-line help is a tool typically accessed only by those trying to complete a task using a computer application or software tool, when completion of the task is somehow hindered by a real or a perceived problem. The learner does not have the time to embark on a lengthy search, nor should she be required to peruse many available resources or practice examples, unrelated to the task at hand, just to come to an understanding of the logic and design features of the tool. She should not be required to invest a large amount of effort to resolve the problem at hand. Since the primary reason for using a software tool is to accomplish a task, the on-line help can be seen as a secondary source of information that helps in the continuation of a task once a person perceives that he or she has a problem. To solve a problem or accomplish a task with on-line help, the user must initiate access, which he will only do if the help is perceived useful. Even when help is available, the problem-solving information may be missing, in the wrong place, or simply incomprehensible. This study hopes to provide examples of problem-solving through on-line help and provide a tool for assessing the adequacy of on-line help tool currently available.

Acknowledgments

The study was supervised by Dr. Betty Collis of the Faculty of Educational Science and Technology at the University of Twente, and was supported by the faculty's Department of Educational Instrumentation, and by the University of Victoria granting study-leave support.

References

Collis, B., and Verwijs, C.: "Evaluating Electronic Performance Support Systems"; ETTI 32, 1 (1994), 23-30.

Duffy, T., Mehlenbacher,B., and Palmer, J.: "On line help: Design and evaluation"; Ablex Publishing Corporation, Norwood, NJ (1992)

Henin, I.: "Online help for ISM-1"; Masters thesis. Universiteit Twente Enschede (1995)

Lazonder, A.: “Minimalist computer documentation”; Proefschrift; Cip-Data Koninklijke Bibiliotheek, den Haag (1994)

Lazonder, A., and Meij, van der, J.: “Error information in tutorial documentation: Supporting users’ errors to facilitate initial skill learning”; International Journal for Human Computer Studies, 42, (1995), 185-206.

Meij, van der, J.: “Does the manual help? An examination of the problem-solving support offered by manuals”; IEEE transactions on professional communication; 39, (1996), 3.

Shneiderman, B.: “Designing the user interface: Strategies for effective human-computer interaction”; Addison-Wesley Publishing Company, Reading, MA (1987)

Stevens, G., and Stevens, E.: “Designing electronic performance support tools: Improving workplace performance with hypertext, hypermedia and multimedia”; Educational Technology Publications Englewood Cliffs, NJ (1995)