

# Chauffeured by the User: Usability in the Electronic Library

Jerilyn R. Veldof  
Michael J. Prasse  
Victoria A. Mills

**SUMMARY.** Librarians have not traditionally been the developers of information retrieval systems. However, with the growth of the World Wide Web and online knowledge management opportunities, academic librarians have begun to emerge as online system developers and designers. As librarians jockey for position as experts in understanding information-seeking behavior in this virtual landscape, they must learn about usability testing. Usability testing reveals how users search for online information and is a key component in determining when a product is easy to use and ready for public use. Both the University of Arizona and OCLC have undergone usability evaluation in various ways and have integrated results of these evaluations into their Web and online product designs.

**KEY WORDS.** Usability testing, product evaluation, system design, World Wide Web, information-seeking behavior

---

Jerilyn R. Veldof is Assistant Librarian at the University of Minnesota, 499 Wilson Library, 309 19th Avenue S., Minneapolis, MN 55455 (e-mail: [jveldof@tc.umn.edu](mailto:jveldof@tc.umn.edu)).

Michael J. Prasse is Consulting User Interface Designer and founder and manager of the OCLC Usability Lab and the OCLC Human-Computer Interaction (HCI) Team, Online Computer Library Center, 6565 Frantz Rd., MC 449, Dublin, OH 43017 (e-mail: [prasse@oclc.org](mailto:prasse@oclc.org)).

Victoria A. Mills is Bibliographic Access Librarian and member of the Access 2000 Project Team, University of Arizona Library, A101 Main Library, 1510 E. University, Tucson, AZ 85721-0055 (e-mail: [vamills@bird.library.arizona.edu](mailto:vamills@bird.library.arizona.edu)).

[Haworth co-indexing entry note]: "Chauffeured by the User: Usability in the Electronic Library." Veldof, Jerilyn R., Michael J. Prasse, and Victoria A. Mills. Co-published simultaneously in *Journal of Library Administration* (The Haworth Press, Inc.) Vol. 26, No. 3/4, 1999, pp. 115-140; and *Information Technology Planning* (ed. Lori A. Goetsch) The Haworth Press, Inc., 1999, pp. 115-140. Single or multiple copies of this article are available for a fee from The Haworth Document Delivery Service [1-800-342-9678, 9:00 a.m. - 5:00 p.m. (EST). E-mail address: [getinfo@haworthpressinc.com](mailto:getinfo@haworthpressinc.com)].

### INTRODUCTION

Librarianship has clearly evolved beyond a profession that only organizes print materials. Digitizing special collections, journals, and data sets and creating and managing online public access catalogs (OPACS) and World Wide Web (WWW) sites are becoming part of library work. As librarians strive to become experts in understanding information-seeking behavior in this new electronic environment, they must learn about usability, one of the most dominant trends in computing today.

Usability means that the people who use a product can do so quickly and easily to accomplish their tasks. Usability focuses on the users and an understanding of what they want and need to accomplish. It requires that users, not the designer of the product, determine when a product is easy to use. It is an excellent way for librarians to begin to find answers to questions, for example: How do people look for information? How would they prefer to get it? How would they prefer to see it displayed, delivered, and processed? For many of us, adopting this kind of a user-centered focus should not entail a major shift in philosophy.

Many libraries are already moving into a user-centered paradigm where we challenge ourselves on all fronts to create services that are user-focused. At the University of Arizona, for example, teams are expected to conduct needs assessments, collect data on our processes, and measure success with tools that solicit customer satisfaction ratings. This is easy enough when it comes to tried and true services which come with a long track-record of complaints: interlibrary loan, reserves, and shelving, for example, were quick to undertake improvement at the University of Arizona Library. Yet, the burgeoning "virtual library" which we were haphazardly building on the WWW somehow escaped scrutiny. This needed to change.

In order to design sites that meet users' needs and expectations, both the University of Arizona and OCLC have done usability testing in various ways. This paper will explore usability and usability research as it relates to Web applications in libraries. It will then show how usability evaluation might be approached, bringing in examples from OCLC and the University of Arizona.

## WHAT IS USABILITY?

Usability is the degree to which a user can successfully learn and use a product to achieve a goal. The International Standards Organization (ISO) formally defines it as “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.”<sup>1</sup> When usability is evaluated and improved upon during all phases of system design, from paper prototype to beta, users will have high success rates and high satisfaction rates with using the final project. The producers of the product will also have avoided the many costs associated with correcting product usability after the product has been released.

### *When Did Usability Become Important?*

In the early days of PC software, the development of functionality predominated. Applications that could do a lot, such as VisiCalc, quickly became best sellers, regardless of their level of usability. However, as more and more applications acquired the same capabilities, the ease of using those capabilities increased in importance. The introduction of the Apple Macintosh in the mid-1980s was a landmark for increasing understanding of the importance of usability in online products. It demonstrated that consumers were willing to pay more for a computer if it was easy to use, even if the functionality of the software had not essentially changed. With the introduction of Windows 3.0 in the early 1990s, the importance of usability increased, and for many companies, the integration of usability testing into software development became reality. Companies began hiring usability specialists and creating new divisions with names like Human Factors Group or Usability Engineering Team. Microsoft and Apple are still industry exemplars in the area of usability. They both have usability professionals that play key roles in product development.

At Microsoft, the Usability Group began in 1988. Usability at Microsoft is defined as “strategies for getting information about users into the development process in a timely way.”<sup>2</sup> The Group’s Usability Specialists do usability testing at all stages of development, focusing on exploration of new ideas early in development and progressing to confirmatory testing (e.g., beta testing) toward the end of the development cycle. The approach is similar to that of a consultant offering usability services to developers and design teams when requested.

At Apple, the User-Aided Design Group also began in 1988, with testing of the documentation for HyperCard. Documentationists had been doing various forms of usability evaluation many years before its application to software, so finding that usability testing in a company began in the Documentation section is not unusual. The Group also uses a consultant philosophy, which allows greater flexibility in the allocation of resources, but also did not necessarily tightly couple usability testing with a project's design team.<sup>3</sup> This philosophy has changed over time with the success of the Group, such that the Group's cost is part of the project's budget. The Group also now provides design skills as well as usability testing.

### *Measuring or Evaluating Usability*

Usability evaluation methods generally fall into one of two categories, those that involve real users and those that do not.<sup>4</sup> Techniques that have little to no actual user involvement are often referred to as discount (or guerilla) usability methods. The term *discount* is not used pejoratively but rather indicates a technique that may be done inexpensively (since users are not always involved) but still return very valuable information. The second category of usability evaluation methods is the more traditional usability test. In a usability test, actual users are observed attempting to achieve some designated goal. Since both kinds of testing have advantages and disadvantages, a combination of both techniques may be most cost effective.

### *Discount Usability Methods*

There are many discount usability methods; only a few will be discussed. The most basic method, called heuristic evaluations, uses a set of heuristics ("rules of thumb") applied by design experts to infer the problems a user may have using a software interface.<sup>5</sup> For example, two heuristics might be: (1) The system should always keep users informed about what is going on, through appropriate feedback within a reasonable time; and (2) Users should not have to wonder whether different words, situations, or actions mean the same thing. With these types of heuristics in mind, several experts independently evaluate the interface, often with a set of target tasks. Expected problems are noted, as well as their anticipated severity and extent. When each expert has

evaluated the interface with all the tasks, they meet and, for the first time, discuss the problems they have identified. A single combined evaluation is created, with the problems identified by more than one expert and given a high degree of severity and extent placed at the top of the list. Heuristic evaluations can be done fairly quickly and inexpensively. However, they generally find fewer usability problems than do usability tests involving actual target users.<sup>6</sup>

Usability inspections involve designers and developers acting and trying to think as users.<sup>7</sup> These inspectors are provided with a product description, user profile, and user tasks (including expected goals). When doing these tasks, the areas where they experience confusion or cannot complete a task are assumed to be areas of likely usability problems. This method is again relatively inexpensive and quick, and it may have the added benefit of getting designers to think more like users. However, trying to think like users is still not the same as having real users test the system.

In the card-sorting technique, a set of cards is created, each labeled with a single concept or potential menu option. Users are then asked to sort the cards into meaningful groups. Users create a title for each group, and combine these groups into meaningful, larger groups. In this way, a menu structure that can be used for a system is created. An example is Nielsen and Sano who developed the SUN Microsystems internal Web home page.<sup>8</sup> In their card sorting task, users were asked to sort into meaningful groups a set of  $3 \times 5$  note cards, each labeled with the name of an information service that might be included in the Web site. Users were then asked to sort each of these initial groups into a smaller set of groups and to invent a name for the group. These results were compiled and analyzed for four users, and a menu structure of fifteen groups was created.

The card-sorting technique can be combined with another quick and inexpensive technique sometimes called the mix-and-match method or icon-intuitiveness evaluation. Icons (both with and without labels attached) or text labels are assigned to each organizational grouping. Users are asked what each icon or label means, and that is compared to the intended meaning of the icon or label. Icons/labels that do not match their intended meaning are redesigned, and a second evaluation is conducted. A set of icons/labels that reasonably match its intended meaning is eventually created.<sup>9</sup>

### *Usability Testing*

Formal usability testing involves real users and more resources and budget than do the discount usability techniques. Usability testing is the observation and analysis of user behavior while users use a product or product prototype to achieve a goal. It has five key components:

1. The goal is to improve the usability of a product.
2. Testers represent real users.
3. Testers do real tasks.
4. User behavior and commentary are observed and recorded.
5. Data is analyzed to diagnose problems and recommend corrections.<sup>10</sup>

For software developers and companies, usability testing is often conducted in a specially constructed lab where testers fitting the profile of an expected user group are videotaped and observed via a two-way mirror. In a typical usability test, testers are provided a brief introduction to the product and asked to complete a series of tasks. They are encouraged to think aloud while doing the tasks, verbalizing what they are doing and why. In addition, testers may be asked to point out areas of confusion and anything they particularly like or dislike.<sup>11</sup> When the tasks are completed, the tester may complete a questionnaire and be interviewed. After the tester has left, the design team will discuss the test, what problems were revealed, and possible solutions. Soon after, an analysis may be distributed. The analysis typically lists the time a problem occurred, a brief problem description, and several suggested solutions. At the end of all the tests, a summary analysis also is distributed.

It is important to note for libraries that good 'utesting' is not as dependent on the physical facilities available as on: (1) The observation of users who accurately represent the target population, and (2) The ability of the observers to detect real usability problems. Hence, using paper-and-pencil to take notes while observing patrons in a reference room can also yield valuable data. Indeed, both techniques combined will yield the best results: lab testing for detailed analysis and blatant usability problems, field testing for "real world" data and fine tuning.

It is also not critical that the item being tested be completely functional. Quite often, a paper prototype (also called low-fidelity) can be

used to achieve results comparable to those for a completely functional product.<sup>12</sup> In such a technique, users are given a set of tasks and asked to indicate what they would do. If this involves a screen change, then a paper version of this second screen is displayed to the user. If the user action would result in a dialog box instead, a paper representation of the dialog box is placed on top of the current screen, and so on.

Regardless of the technique, the fundamental belief underlying all usability testing is that data from actual users is essential to understanding the usability of a product or service. This may seem obvious, but the popularity of usability testing has increased only recently, though its beginnings can be traced to the late 1970s or earlier.

### *Usability and the Library*

Libraries have been slow to tap into this field of knowledge created by usability and human factors specialists like those at Apple and Microsoft. One basic reason for this is that libraries, for the most part, have not been producers of software or computer interfaces. In the past, libraries themselves have been customers for computer/software products. They bought automated library systems, CD-ROMs, and networking software and then customized those products when possible. Libraries might have been beta testers for a system and thus helped in the design, but librarians, with few exceptions, were not the designers.

Librarians may have complained to the companies who sold the products if the product did not work like they or their customers anticipated. They may have joined user groups for these products and asked for improvements in the product along with all the other users. But librarians had very little control over either the design or redesign of these products. What they could control, however, were the instructional tools, teaching sessions, and reference assistance which provided a layer of intervention in between the public and the product. Librarians sometimes became like triage nurses, soothing over frayed patron nerves and providing as much help as they possibly could to make online experiences successful.

Librarians, therefore, have always been concerned with how users seek information in automated environments. Pursuit of this concern played an important role in attempts to provide better service by removing barriers to using various online products for the users and in recommending design changes to vendors. The segment of library

literature that explores this research focuses on end-user behavior in automated systems, online catalogs, networked databases, and CD-ROMs.<sup>13</sup> Although end-user studies and usability studies share some of the same research methodology and tools, the studies differ in several ways. Figure 1 compares end-user studies with usability evaluation. Findings from end-user studies are an invaluable way to begin to approach the design or redesign requirements of a product. They may warn about design flaws or inform about users' preferences or needs, but they are no substitute for the evaluation of usability which will actually determine that a specific product or Web site is more or less usable for its target market.

FIGURE 1. Comparison Table

End-User Studies	Usability Evaluation
Performed by librarians or information specialists who are usually not working with the designers of the product being studied.	Performed by trained testers hired by or working with the designers of the product being studied (who may also be librarians or information specialists).
Conducted primarily to understand users. Problems users are having are identified so that instructional tools, workshops, and reference staff training can be designed to better help users with the product. The eventual improvement of the product may be a by-product.	Conducted to improve the usability of the product. Problems users are having are identified so that the product can be improved. Development of online instructional tools, workshops, and reference staff training may be a by-product.
Results of these studies are often generic and applicable to other OPACS or CD-ROMS. They may involve a variety of products to determine overall problems.	Results of these studies involve a specific product to determine specific problems with product design, although general implications for other products may be gleaned.
Users are the focus. Users are studied to see why they use the system and how they interact with it. Users are usually observed doing their own tasks in the system.	The product is the focus. The product is studied through the behavior of the users. Users are given tasks to complete in the system. They are observed to see how they think and use the system to complete those tasks.
Studies are usually done on the finished product that is already available to the public.	Studies are usually done on a prototype or beta version of the product before it is ready for the public.



### ***Relevance to Libraries***

Usability is particularly relevant to librarians as their roles change to information specialists and system designers. More libraries, for example, are beginning knowledge management projects such as designing electronic journals and online exhibits to their special collections. And with the development of the World Wide Web libraries on a large scale have taken real steps towards becoming "designers." Libraries are now creating the Web gateway to the library, its resources and to the entire world of Internet resources. Libraries finally have control over the organization and design of a powerful information tool. We are responsible for the display and design elements and for the usability of this tool.

This is a new role for us and, unfortunately, it has caught many libraries unprepared. While many libraries jumped right in and began producing Web pages, this production was done with great abandon. Some libraries' Web sites grew with little coordination or systematic planning about the overall look, feel and design of the site. As a Web page was created, it was linked to the library homepage and if it did not fit into a category on that site, a new category was created. Applying established design principles and conducting usability tests on Web sites was just not done in the early days of the Web. The Web was "the in thing," and just having a Web site was considered a wonderful accomplishment. The purpose and quality of what was being done was not scrutinized in detail because we were all so dazzled by the Web itself.

Now that the Web is several years old, some of that dazzle has worn off. In addition, libraries are becoming more concerned with user needs, customer self-sufficiency, and being user-centered organizations. We are facing the challenge of trying to fit customer needs and expectations into our services and systems, rather than forcing our customers to mold their needs and expectations to fit the predefined structure of the library. We are seeing that computers have not necessarily made everything easier and better for our customers. In fact, the multiplicity of products and platforms has added a barrier to the use of information in the library. All of this makes it imperative that, when it is in our control, we must create systems that are easy to learn and use and that remove as many barriers as possible. Usability testing is a way to ensure that we achieve this objective.

In October 1997 a search in the online index *Library Literature* on the term "usability" brought up fifteen citations. A similar search in the computer index *Inspec* brought up 1132 citations. Clearly the computer field is much more tuned in to the concepts of usability and usability evaluation than is librarianship. However, librarians are beginning to be aware of and to take advantage of usability evaluation, as experiences at Simmons College,<sup>14</sup> the University of Washington,<sup>15</sup> OCLC, and the University of Arizona Library will attest. An important book written by Bryce L. Allen from the School of Library and Information Science at the University of Missouri called *Information Tasks: Toward a User-Centered Approach to Information Systems* also supports this direction. Allen cautions, "Most contemporary services seem to be created on the 'If you build it, he will come' principle, where 'he' is the imagined user. Of course, there is no guarantee that users will be willing or able to employ services that are assembled without serious attention to their needs."<sup>16</sup> It is time libraries get serious about usability evaluation. It is time we look outside our profession to see what we can learn from usability professionals.

### *User Evaluation and the Web*

Much of the publishing in the area of Web users and their searching/navigation behavior has largely been in the domain of designers who create guidelines or rules based on their design expertise but not necessarily on Web usability testing. Although these guidelines can provide an excellent starting place for librarians creating Web sites, these guidelines should not be relied upon solely and they should never be substituted for actual usability testing with real users of the sites.

The following is an example of some of these kinds of guidelines/rules taken from *The 7 Keys to Effective Web Sites*:

1. A site must be visually appealing
2. A site must be valuable, useful or fun
3. A site must be current and timely
4. A site must be easy to find and use
5. A site must have intuitive on-page navigation
6. A site must involve the user
7. A site must be responsive to its users.<sup>17</sup>

Not all of these guidelines are necessarily applicable to library Web sites. Often these rules emerge from an assumption (based on minimal studies) that users access the Web to browse around, and not primarily to obtain information. A study at Georgia Tech in 1996 drew from 59,000 users and concluded that 77% of their users described their primary Web activity as browsing. Their use of the Web was not task-specific.<sup>18</sup> Library users, however, may be more task-oriented and thus less interested in bells and whistles. Indeed, some research shows that users who approach the Web for information retrieval search the Web differently than would many of the Georgia Tech respondents. Among other things these users report that sites using such attractions as animated graphics and sound are mere annoyances and distract from their tasks.<sup>19</sup>

This kind of finding, however, is still very new in the Web area. Jared Spool's book *Web Site Usability: A Designer's Guide* is the first published study that has attempted to explore Web usability for those focused on information retrieval.<sup>20</sup> Spool's study had 50 users test usability for specific tasks on nine popular sites on the Web including Travelocity, a site to book airline tickets and make other reservations (<http://www.travelocity.com>); Edmund's, a site to get car and truck prices (<http://www.edmunds.com>); and Fidelity, where users can find information on Fidelity mutual funds and investing opportunities (<http://www.fidelity.com>). In addition to applying findings directly to the individual sites studied, overall results revealed five major implications for Web site design, some of which surprised even the testers:

Implication 1: Graphic design neither helps nor hurts in the search for information.

Implication 2: Text links are vital. They are more often considered before graphical ones. Predictability of these links is probably the highest indicator of user success: "The better users could predict where a link would lead, the more successful they were in finding information."

Implication 3: Navigation and content are inseparable. Separating content and the navigational structure (creating "shell sites") leads to generic links that then make it more difficult for users to predict what they will find, therefore decreasing success rates.

Implication 4: Information retrieval is different than surfing. Users who are task-oriented click on links that they feel certain will lead them to the information they are seeking, and are more distracted by visual noise on a Web site.

Implication 5: Web sites aren't like software.

A follow-up study by Spool et al. has presented some new information.<sup>21</sup> In the earlier study, novice Internet users often said they would have done better with more Internet experience. The follow-up study was designed to investigate this hypothesis. They found that there was no correlation between a user's understanding of the Internet and the number of Web search tasks correctly completed. However, experienced users had developed what Spool refers to as "defensive mechanisms," behaviors designed to avoid the pitfalls of poor Web user interface design. For example, experienced users would scroll to the bottom of a Web page on first viewing it, while novice users would not. Experienced users would look for ways to get back to the starting point, such as links back to the home page. They would also actually *read* search tips while novice users would not. Experienced users were more likely to criticize the look of a site. And most interestingly, experienced users were *less* likely to see the Web as a good place to find useful information. This may be due to hype versus reality; it would be difficult for any system to be as good as the Web is often said to be on the evening news, and experienced users know this fact.

A user study of the Web by Pollock and Hockley indicates that without some a priori Internet training, users may be very discouraged by their first use of a Web search system.<sup>22</sup> They found that users are surprised at the breadth of Web searching, including its international scope. They were not happy, however, with the quality of the material found, or the difficulty it took for them to find it. Not surprisingly, users expected the computer to understand true natural language (versus a system-controlled vocabulary). The authors conclude with several recommendations:

1. Search engines should concentrate on doing *simple* searches well before moving to support advanced users.
2. Search results should be returned to the user as quickly as possible, with clear progress indicators.

3. Search engines should communicate that searching on the Internet is a process, not a single event. Many searches may be needed before the desired results are obtained.
4. Intelligent support. Suggest variants.

Neilsen and Sano have also published results from usability testing.<sup>23</sup> They focused on the SUN Microsystems internal Web home page. Testing methods included a study using the card sorting technique (used to determine menu structure for the site), an evaluation of icon intuitiveness, and two usability tests. In their conclusion, the authors found that, first of all, people have little patience for poorly designed sites. Users were not likely to return if the site had numerous system errors or 'under construction' symbols. Secondly, users do not want to scroll. They felt Web pages that required scrolling to view the most relevant sections were poorly designed. Finally, users do not want to read. They scan for hyperlinks.

Schneiderman, Byrd, and Croft studied information retrieval interfaces.<sup>24</sup> They concluded with a set of eight guidelines for design of a usable information retrieval interface:

1. Be consistent. If you use 'Sources' at one point in the interface, do not switch to 'databases' later on.
2. Provide shortcuts for experienced users. For example, let experienced users enter a complete search term as au:smith rather than selecting 'Author' from a list and typing 'smith' separately.
3. Offer feedback to help improve the search. For example, a system that suggests alternative terms, such as 'feline' for 'cat' will be more usable.
4. Design for closure. Letting a user know when they have viewed all the options in a menu or results list can minimize patron time at the terminal. Such simple cues as placing the text 'End of Results' at the bottom of the last page of results can increase usability greatly.
5. Error handling. Syntax errors should be prevented where possible; all error messages should be specific and constructive.
6. Permit easy reversal of actions. Every action should be reversible so users can go back to a previous state in a session.
7. Support user control. In a well-designed interface, users initiate action, monitor progress of long searches, and always feel in control.

8. Reduce short-term memory load. Keep a session history, so users can always go back and reuse previous effort.

Librarians are also just beginning to contribute to this literature on Web site design based on usability testing. The University of Washington investigated the "effect of terminology and screen layout/grouping on students' ability to correctly select databases" from the initial screens of their library Web site.<sup>25</sup> They tested 170 undergraduates using three versions of a low-fidelity prototype. They found that "grouping resources and assigning concrete, descriptive labels help undergraduates to make more efficient navigation decisions."<sup>26</sup> Although this first study was a fairly simple test, the authors next plan to conduct more rigorous usability studies using think-aloud protocols, timed-task analysis and a post test debriefing.

Many of these studies obtained remarkably similar results, which can be used to create more usable Web sites. They are helpful for developing heuristics (or design guidelines) which can be used in the development and the testing of Web sites. Findings include:

1. Web searchers benefit most from 'smart' search assistance and good search results feedback.
2. Real-time indicators of search status are very useful.
3. Internet experience can benefit by preventing mistakes and misunderstandings, but not necessarily in finding higher quality information on the Web.
4. Usable Web design includes not making users scroll to perform essential actions (e.g., starting a search).
5. Novice users scan Web pages for links, but rarely read text such as search tips. Do not count on such tips to help the novice user.

The rest of this paper will discuss the various methods for conducting usability evaluations used at OCLC, the University of Arizona's campus Web site, and the University of Arizona Library's information gateway site.

#### ***WEB USER EVALUATION AT OCLC: A HYBRID APPROACH***

A hybrid methodology using a combination of usability evaluation techniques is an approach gaining increased use at the Online Comput-

er Library Center (OCLC).<sup>27</sup> The OCLC approach has four major components: (1) internal review, (2) heuristic review, (3) usability testing, and (4) field-testing. The first step is an internal review of a system prototype by designers, developers, marketers, documentation and other interested parties. This prototype may consist of only a few of the more critical system screens, prototyped using paper and pencil or a software graphics package. It is created only after user characteristics, user needs, and hardware/software constraints have been identified. The focus of this initial evaluation is on determining if the prototype meets functional requirements and is within the hardware and software constraints. The reviewers are specifically not looking for usability problems at this point.

A heuristic review occurs after the prototype has been modified based on the comments received during the internal review. Although a heuristic review of a paper prototype is not as effective as for a more realistic version of the product, it nonetheless can yield valuable information very quickly.<sup>28</sup> The heuristic review is a service offered by OCLC's Human-Computer Interaction Team (HCITeam). The Team is comprised of OCLC staff from Development, Quality Assurance, Graphics, Documentation, and User Support (a.k.a. the Help Desk). The Team produces a report documenting the problems located, their severity, and offering solutions.

The third step is a formal usability test. Test users are from both the patron and librarian user group. Tasks appropriate to each group are used, and analyses by the HCITeam are distributed. The prototype may go through several iterations of test, redesign, and retest before it reaches an acceptable level of usability.

The final step is field testing. At this stage, the product is no longer a paper prototype, but a working software package nearing completion. Data from field tests provide the fine-tuning necessary before a product can be released.

### **WEB USABILITY STUDIES AT THE UNIVERSITY OF ARIZONA**

Usability testing was conducted on the University of Arizona's Web site (<http://www.arizona.edu>) by students in a Management Information Systems class taught by Dr. Suzanne Weisband in the spring 1997 semester. Test users were predominantly students from the University

who were familiar with the campus but not necessarily familiar with the Internet. Following a pre-test questionnaire that gathered demographic information, test users were asked to complete specified tasks on the current University Web site and were closely observed for both verbal and nonverbal actions as they attempted to complete these tasks. The amount of time and number of mouse clicks it took to find the appropriate Web page and the appropriate information from that page were both recorded. The number of times the user had to back up and the keywords used to complete the task were also recorded. Users also completed a post-test questionnaire. The questionnaire asked the testers such questions as what they did to complete their tasks and why, their satisfaction with the organization of the site, their satisfaction with the content of the information they found, and how designers could improve the design of the site.

Some of the results of the testing are summarized below. They illustrate how simple and inexpensive usability evaluations can be used to guide design and eliminate usability problems before the public release of the Web site.

1. Student users tended to skim text and read bulleted information. Because of this they had difficulty finding information on admission to the University. This problem extended even to the Admissions page itself, a text intensive page containing numerous embedded (not bulleted) links.
2. The toolbar was often either overlooked by users or not used because of confusion over what the various buttons represented.
3. The index was rarely used as a tool for finding information. In addition, those who did use the index frequently missed the information they were seeking.
4. The online phone book was identified as having major usability problems including an overly technical style for conveying instructions, a frame design which cut off numbers in some browsers, and the use of the word "query," the meaning of which eluded most users.

Immediate changes were made following these studies which helped make the site more usable without a major redesign. These included:

1. Making the background white.



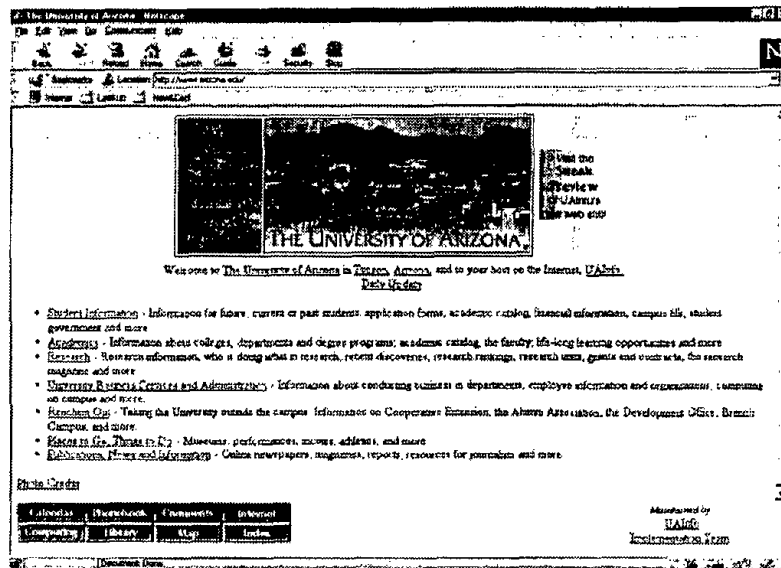
2. Removing the frame (in which the toolbar was located) from the second level pages. The toolbars remained but at the bottom of each page, not in a frame.
3. Moving the UAInfo link from the top of the main page to the bottom. (UAInfo is the name of the University's Web site.)
4. Breaking up the student information page by removing much of the academic information and substituting links that pointed to duplicate information in the academic listing.

Results such as these later helped justify a major redesign effort in which the UA Info Advisory Council enlisted the help of a class focusing on Web design. The Council, in charge of the redesign effort, also employed the card sorting technique to elicit input in the categorization of the new site. Index cards were made of each site that was to be linked from the UA Web site parent pages. Council members asked faculty, staff, and students to sort the cards into reasonable categories, and then to label the groups. These labels and groups became the main organizational structure of the redesigned site. Instead of arranging the site by internal functions with labels such as academics and University Business Services and Administration, the card sorting exercise pointed to an arrangement by type of user: Prospective Students, Enrolled Students, Faculty, Staff, Alumni and Visitors, and Administrators. The old and new UAInfo designs are displayed in Figures 2 and 3.

#### **WEB USABILITY AT THE UNIVERSITY OF ARIZONA LIBRARY**

At the University of Arizona Library the need to redesign our Web site was identified by staff working directly with confused users. Some of the impetus for redesign, however, came from the purchase of an OCLC product called SiteSearch. SiteSearch allows the Library to integrate many interfaces into a single uniform site and to design knowledge management systems to run with this system. A group of four librarians representing public and technical services and one systems analyst (also the Library's Webmaster) were charged with creating a usable Web site utilizing SiteSearch. (For the charge of the team, called Access 2000, see <http://dizzy.library.arizona.edu/library/teams/access9798/access.htm>.) Access 2000 members asked ourselves, "If we were designing a Web site from scratch, how would we go about

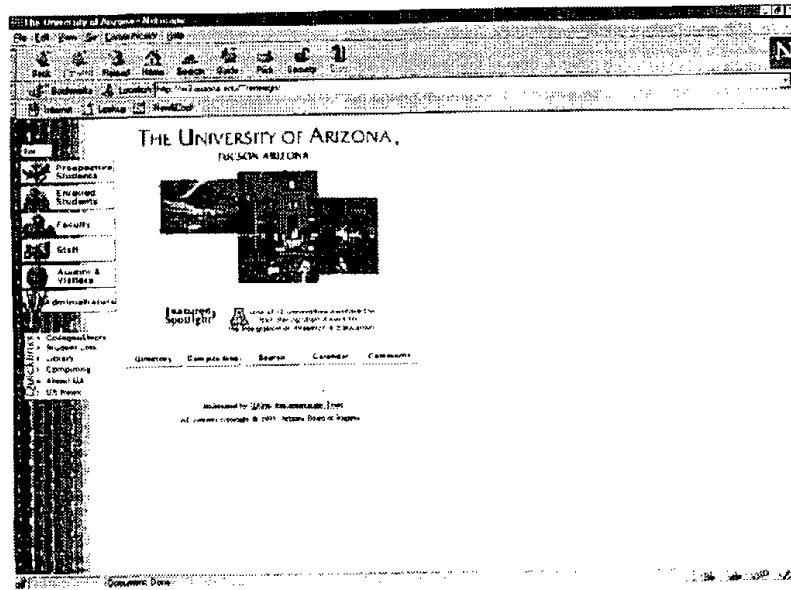
FIGURE 2



doing this?" In order to find these answers, we decided to draw upon usability evaluation methods described above to direct the team during all steps of the redesign process.

Team members began by gathering user feedback from as many sources as possible. Prior feedback through the comments mechanism on the Library information gateway (called SABIO) was read and analyzed. A survey designed to gather baseline data was administered by the team in order to measure success of the redesign. The survey gathered satisfaction levels as well as level of assistance on the current online system. Volunteer staff members administered the survey in the reference areas, asking users to fill out the form in exchange for a fancy pencil. An online version was also mounted on SABIO. Within two weeks the team was able to collect 830 surveys using these strategies. This same process will be duplicated after the new design with the SiteSearch component has gone live. The timeline for this is rather unclear, dependent on variables that are not entirely within the Access 2000 Team's control. It is expected that fall 1998 is a realistic time-

FIGURE 3



frame. In addition, focus groups of faculty, graduate students, and undergraduates were held in order to learn what they think does and does not work on the current SABIO, what components they would like to find in a perfect academic library computer system, and how they would expect to complete various tasks in this ideal system. (See Appendix A for major themes that emerged. A copy of the focus group script is available at [http://dizzy.library.arizona.edu/library/teams/access9798/focus\\_groups/Focus\\_Groups\\_Script.htm](http://dizzy.library.arizona.edu/library/teams/access9798/focus_groups/Focus_Groups_Script.htm).) Using this information, the team built simple paper prototypes to use in the early stages of usability testing. These paper prototypes allowed the team to begin testing before significant time and energy has been put into a fully operational design. In this way team members were not personally invested in these designs and could tinker with them freely. After several usability testers showed that the paper mock-up designs were heading in a good direction, a student was hired to translate them into simple HTML. Although not all of the functionality was in place at this point, the HTML prototype enabled the team and the testers to

begin to get a real “feel” for what the site would look like. Both the paper and HTML prototypes were used in the subsequent usability tests.

The team next designed a usability script to use during actual testing. The script included tasks that would help shape the design of key components of the new site (see Appendix B for the script and tasks). The script was first used on a librarian and two library school students as a way to identify any design flaws in the script and to help the team plan the details for administering the test. The group found that the nine tasks took an average of thirty-five minutes to complete. Initially testers were also asked to repeat briefly these tasks on another library’s system so that Access 2000 members could compare usability across sites with different designs. Total time for the test participant was forty-five minutes that included a brief questionnaire given after tasks were completed (see Appendix C for a copy of the questionnaire). The testers were recruited using a variety of methods—email messages to library student workers, flyers in dorms and the Student Union, personal contacts, and word of mouth. The ten dollar payment did much to encourage participation.

The tests were conducted in a private lab space in the Main Library called the Staff Learning Center. There the tester was able to use a computer for the online prototype design pages and a large table for the paper prototype (often testers had to switch between the two prototypes during a test). The space was private, allowing the tester to think aloud during the test. The main test monitor was in charge of following the script, interacting with the tester, and facilitating the screen displays. A second monitor captured the think-aloud process on paper, and a third monitor recorded the time and the paths the tester took to complete the tasks. In later tests, only the online HTML prototype was used. These tests required only two test monitors. However, it was rare for only the monitors to be present. The usability tests were so instructive that three or four of the Access 2000 team also attended in order to observe and listen. Students were scheduled in one-hour blocks so that the remaining fifteen minutes of the hour could be spent for debriefing with just the observing team members. During this time, notes from the test were reviewed and a summary of the key points learned was developed. Small design changes were also discussed and agreed upon.

The first half-dozen or so usability tests began to show the team

what the major pitfalls were with the current SABIO and the rudimentary mock-up redesign. Test participants had been brutally honest saying such things as “this is really confusing,” “there’s way too much information here, I’d never read it,” and “I never scroll down—don’t put anything important down there at the bottom.” These testers revealed such issues as design complexity, site length, and terminology and identification methods used to identify various functions. Many of these issues were only identified by our “real” users, and not by library staff who were initially tested. One example is the use of the term “catalog.” Staff for the most part used this label appropriately in their completion of assigned tasks while the real users continuously misunderstood or ignored the label and were therefore unsuccessful, at least at first, with many of the assigned tasks.

Quickly the team developed a mix-and-match to get user feedback on the question of what to call “The Catalog” since it was obvious that “The Catalog” meant nothing to many of our library users. The team created a one-page form with several buttons on each side of the page conveying the concept of the catalog, including “What UA Library owns,” “Where to go for books, journals, etc.,” and even “What’s in UA Library.” Students and others researching in the Library were stopped by team members and asked to indicate which button on each side of the form they would click “if you wanted to find out if the Library had *Sports Illustrated*.” They then selected the “best” button between the two. This method took very little time and effort and could be done easily while working at the reference desks. It challenged the team to use terminology that the students understand as opposed to the librarian terminology they normally would use and was used with other icons and terminology as the need arose.

The team also used other quick and easy usability techniques including the card-sorting test discussed in a previous section above. This exercise is another example of a test that, if chunked into small sections, can be done using actual customers in the library without having to undergo the rigors of scheduling, payment, and team participation. In this case testers were asked to sort various pieces of the Web site into like categories. This method was customized to respond to specific questions that arose during usability testing in the same way that the mix-and-match test was able to do.

After the team had conducted its first half dozen usability studies, there were meetings away from the Library to revise the design into

what we are calling "Generation 1." Again, a student coded the paper mock-ups into somewhat functional HTML code and revised the previously coded designs.

Access 2000 team members at the University of Arizona Library plan to continue to use a combination of usability evaluation methods as revisions of the new design emerged. The team found this iterative design method of test-revise-retest to be critical for developing a usable site. We expect that by inviting our users to co-design a system with the Library the final product will truly help our customers be highly successful, self-sufficient library researchers. The new design should be fully live for the fall 1998 semester at <http://www.library.arizona.edu>.

### CONCLUSION

"The web is a whole new ball-game, and we're still learning how to play. We don't yet know how to design for finding information," says Spool.<sup>29</sup> Librarians, therefore, still have an opportunity to make a significant contribution to this field and take a leadership role in the organization and development of the Web for the universities and communities for which we work and the field at large. Ongoing development and application of usability testing in libraries is one way to ensure that we are developing electronic services and resources that are chauffeured by the user.

Members of the Access 2000 team will be publishing a follow-up article with the results of these usability evaluations in the near future. A working directory of documents is available at <http://dizzy.library.arizona.edu/library/teams/access9798/>.

### NOTES

1. ISO DIS 9241-11, *Ergonomic Requirements for Office Work with Visual Display Terminals. Part 11: Guidance on Usability* (London: International Standards Organization, 1994), 10.

2. Michael Wiklund, ed., *Usability in Practice* (Cambridge, MA: Academic Press, 1994), 330.

3. *Ibid.*, 94.

4. A very useful description of a variety of usability evaluation techniques can be found on the Web as an appendix to a forthcoming book, Susan Fowler, "Appendix

B: Usability Tests" in *GUI Design Handbook* (McGraw-Hill, forthcoming), currently available at: <<http://www.fast-consulting.com/preface.htm>>

5. Jakob Nielsen. *Usability Engineering* (San Diego, CA: Academic Press, 1993), 155-165.

6. Clare-Marie Karat. "Cost-justifying Human Factors Support on Development Projects." *Human Factors Society Bulletin* 35 no. 11 (1994): 1-4.

7. Michael J Kahn and Amanda Prail. "Formal Usability Inspections." in *Usability Inspections Methods* (New York: John Wiley & Sons, 1994), 141-171.

8. Jakob Nielsen and Darrell Sano. *SunWeb: User Interface Design for Sun Microsystems's Internal Web*, 1995. <<http://www.sun.com/sun-on-net/uidesign/sunweb>>

9. Nielsen, *SunWeb* and Fowler. "Appendix B. Usability Tests."

10. Joseph C. Dumas and Janice C. Redish. *A Practical Guide To Usability Testing* (Norwood, NJ: Ablex Publishing Co., 1993), 22.

11. K. Anders Ericsson and Herbert A. Simon. *Protocol Analysis: Verbal Reports As Data* (MIT Press: Cambridge, MA, 1985).

12. Carolyn Snyder and Ericca Lahti. "Practical Tips for Paper Prototyping: Things We Learned the Hard Way." in *Proceedings from Usability Professionals Association Annual Conference (4th)* (Dallas: Usability Professionals Association, 1995).

13. There are hundreds of end-user studies in library literature during the 1980s and 90s. A representative sample of these studies is: Christine L. Borgman. "Why Are Online Catalogs Hard to Use? Lessons Learned from Information-Retrieval Studies." *Journal of the American Society for Information Science* 37 (November 1986): 387-400; Karen R. Diaz. "User Success in a Networked Environment." *RQ* 36 (Spring 1997): 393-407; Carol A. Hert. "User Goals on an Online Public Access Catalog." *Journal of the American Society for Information Science* 47 (July 1996): 504-518; Charles R. Hildreth. "The Use and Understanding of Keyword Searching in a University Online Catalog." *Information Technology and Libraries* 16 (June 1997): 52-62; Bruce Leach. "Identifying CD-ROM Use Patterns as a Toll for Evaluating User Instruction." *College & Research Libraries* 55 (July 1994): 365-371; Diane Nahl and Carol Tenopir. "Affective and Cognitive Searching Behavior of Novice End-Users of a Full-Text Database." *Journal of the American Society for Information Science* 47 (April 1996): 276-86.

14. Elizabeth T. Morley. "The SilverPlatter Experience." *CD-ROM Professional* 8 (March 1995): 111-112.

15. Karen Eliassen et al. "Navigating Online Menus: A Quantitative Experiment." *College & Research Libraries* 58 (November 1997): 509-516.

16. Bryce L. Allen. *Information Task: Toward a User-Centered Approach to Information Systems* (San Diego, CA: Academic Press, 1996), 1.

17. David Sachs and Henry Stair. *7 Keys to Effective Web Sites* (Saddle River, NJ: Prentice-Hall, 1997), 16-17.

18. Alison J. Head. "Web Usability and Essential Interface Design Issues." *National Online Meeting Proceedings* 18th (May 1997): 157-163.

19. Jared M. Spool. *Web Site Usability: A Designer's Guide* (North Andover, Massachusetts: User Interface Engineering, 1997), 10-11.

20. Spool. *Web Site Usability*.

21. Jared M. Spool, "Internet-Savvy Users," *UIETips* (November 1997). UIETips@uie.com
22. Annabel Pollock and Andrew Hockley, "What's Wrong with Internet Searching," *D-Lib Magazine* (March 1997), www.dlib.org.
23. Jakob Nielsen and Darrell Sano. 1995. SunWeb: User Interface Design for Sun Microsystem's Internal Web, www.sun.com/sun-on-net/uidesign/sunweb.
24. Ben Shneiderman, Don Byrd, and Bruce W. Croft, "Clarifying Search: A User-Interface Framework For Text Searches," *D-Lib Magazine* (January 1997), www.dlib.org.
25. Eliassen, "Navigating Online," 515.
26. Ibid.
27. Michael J. Prasse, "User Interface Design: Draft 1.1," in *Online Computer Library Center ISO Standards Draft Document Version 1.1*. (Dublin, OH: OCLC, 1997), online document.
28. Jakob Nielsen, "Finding Usability Problems Through Heuristic Evaluation," in *CHI '92 Conference Proceedings: ACM Conference On Human Factors In Computing Systems* (New York: ACM, 1992), 373-380.
29. Spool, *Web Site Usability*, 12.

## APPENDIX A

### Main Themes from the Focus Groups

- Common interface
- Common commands
- Simple & advanced search
- Context sensitive help
- Simultaneous database searching
- Simultaneous format searching and selection
- Simplify—(not multiple layers, or too many graphics) or steps to find something
- Know where you are
- Management tool
- Identify formats first, then search



## APPENDIX B

### Script for UA Library Usability Study

As you go about answering the questions I am going to ask, we would like you tell us what you are thinking every step of the way. We are interested in your reasons for what you are doing on the system and we really need for you think aloud for us.

So, when you select an option, please tell us what you are going to pick and why. For example, you might say, "Gosh, I don't know what this button means, so I'm going to look for something else," or you might say, "Oh, this is obvious, I'll go here."

Nothing you can do will be wrong since we are trying to design a better system for users just like yourself. We want to find out what makes sense and what doesn't, what is hard for people, and what isn't.

Do you have any questions?

Scenario:

You are in the library seeking various kinds of information using the library's computer system.

Tasks:

1. Show us how you would find a book about affirmative action.
- 2a. Find a scientific article about the cloning of sheep.
- OR
- 2b. Find a business article about Hispanic owned businesses.
3. Does the library own Sports Illustrated?
4. Can you find a web site about Yaqui Indians?
5. Can you find whole issues of electronic journals in physics?
6. Can you find the spring schedule of classes for this University?

## APPENDIX C

## Usability Questionnaire

Please take a minute to answer the following questions:

1. Physically, how do you feel right now?  
exhausted 1 2 3 4 5 6 7 full of energy
2. Mentally, how did you feel while working on these tasks?  
completely confused 1 2 3 4 5 6 7 everything made sense
3. Do you feel the ease of use of this site to be  
unacceptable 1 2 3 4 5 6 7 excellent
4. How confident are you that you found all the relevant information?  
not at all confident 1 2 3 4 5 6 7 very confident

Note: This questionnaire was designed from a Post-Task Questionnaire from Jared M. Spool's *Web Site Usability: A Designer's Guide*.