Gaining User Insight: A Case Study Illustrating the Card Sort Technique

Angi Faiks and Nancy Hyland

In spring 1999, Cornell University Library performed a user study to help determine how users would organize a set of concepts to be included in an online digital library help system. The study employed the card sort technique, in which users impose their own organization on a set of concepts. The card sort technique proved to be a highly effective and valuable method for gathering user input on organizational groupings prior to total system design. The authors present Cornell's experience as a case study with detailed instructions for conducting and evaluating the card sort technique.

In 1999, Cornell University Library (CUL) launched a comprehensive online help system (Gateway Help) for users of its digital library (www.library.cornell.edu). A committee made up of six reference librarians, including the authors of this article, and two information technology programmers designed and built a single point-of-entry system for users seeking help on a variety of library-related topics. This comprehensive online system consists of both library research and technical support components. A significant challenge of the project was that of organizing the numerous files included in the system in a way that would make them most accessible to users. To that end, the reference librarians working on the project created several possible tables of contents (TOCs). The TOC was designed to serve as the principal entry point and initial interface to the contents of Gateway Help and was a crucial component to the system. Although the librarians felt that they should certainly know how to categorize and organize the pages, even within their small group they could not agree on the structure of the TOC. They realized that to create something meaningful for users, they needed the users' input. It was decided to use the card sort technique to cull information on how users would organize the contents of Gateway Help. This technique is easy to replicate. The authors recommend it for any library that is wrestling with how best to organize a set of topics in a Web interface. This article describes the card sort technique and explains how it can be used in organizing Web sites and information systems for improved multi-user access.

Background

Academic libraries today provide numerous resources and services via the Web to

Angi Faiks and Nancy Hyland are Reference Librarians in Mann Library at Cornell University; e-mail: ajf9@cornell.edu and nch9@cornell.edu, respectively. *The authors wish to thank the committee members for their hard work and dedication to the project: Martha Walker, John Fereira, Lance Heidig, Marianne Hansen, Lori Micho, Adam Smith, Tom Gale, Bill Walters and Marty Schlabach. The authors also thank CUL for the opportunity to provide support to users via the Gateway Help System.*

enable users to do much of their research remotely. To assist end users who are not visiting the reference desk, CUL decided to redesign and expand the Gateway Help component of its Web-based collection of resources and services available through the Cornell Library Gateway.

In general, users find online help cumbersome and often have trouble locating help that is relevant to their needs.

In 1999, CUL conducted a study "to assess how Cornell's Library Gateway was being used; to ascertain the Gateway users' satisfaction levels, likes and dislikes; to determine enhancements for the current Gateway; and to identify future improvements for CUL's common entryway."[1] The study found that most Gateway users did not like to use online help, and when they did, they either did not find it helpful or got lost in the process of using it. This finding is not unique. Online help is often the least preferred method of assistance.[2] In general, users find online help cumbersome and often have trouble locating help that is relevant to their needs. Although the committee charged with designing and implementing the CUL Gateway Help realized that it could not change aversion to online help in general, it could try to create a more user-friendly system. Specifically, the committee attempted to offer a variety of ways for users to find or discover relevant help. To this end, it implemented both a search mechanism and an alphabetical index. However, the TOC provided the main access point and served as the structure for the entire Gateway Help interface. The committee felt strongly that a meaningful organization of the contents would play a major role in system design and user satisfaction.

Librarians are skilled in information management and thus often are considered experts in organizing system content. Yet, as Jeffrey Rubin wrote in *Handbook of Usability Testing*, "The design of usable systems is a difficult, unpredictable endeavor, yet many organizations treat it as common sense" and "actual knowledge of usability" is "quite uncommon."[3] A useful method of gathering input on the organization of information is the card sort technique.

Card Sort Technique

Computer science literature shows that card sorting is a "common usability technique that is often used to discover users' mental models."[4] The technique has proved very useful in helping to organize several pieces of information or concepts. It entails providing a group of users with a set of cards. Written on each card is a concept or piece of information from the set that needs to be organized. Users then sort the cards with similar concepts into piles. The cards are scored and the data entered into a statistical analysis program. A statistical cluster analysis can be used to create a composite of all or various groupings of users. The technique is based on the assumption that if users group cards together, the concepts probably should be grouped together in the system.[5] The result suggests how users would organize a given set of concepts, which can be very valuable information when organizing a system or Web site.

Rationale

Good system design incorporates usability testing from inception rather than at the end of the interface design process. Card sorting works well in the early stages because it gives users an opportunity to create a proposed organization as opposed to reacting to one already in place. This provides guidance to the project team as it lays the foundation for the interface. In *Learning the Laws of Usability,* Lucy Lockwood wrote that "you and your design team will create much more usable designs and develop superior solutions to your user interface design challenges if you resist the urge to get concrete early in the development process."[6] Getting "concrete early in development process" can hinder design and make it difficult to adapt and change.

Even if analysis or study shows that designing a system differently would be better for users, a design that already has begun may be too difficult or time-consuming to change. The card-sorting study can provide a means of taking a look at effective organization before system or site design.

Furthermore, user perceptions may be influenced by a preexisting organization. After a structure is presented, it can become fixed in the user's mind. This then influences the structure of the user's mental map of an information landscape. If a user sees concepts grouped together, he or she may be predisposed to consider them as similarly related for the first time and to respond in kind to questions asked about the system (see, for example, Thomas Kuhn's *Structure of Scientific Revolution*). Card sorting imposes no structures and gives librarians a glimpse into how patrons organize information free of the librarians' influence.

Actually, implementing a card-sorting study is a relatively easy task, for both those administering the study and those participating in it. Anyone with a set of index cards and some time can do it. Running the statistical analysis is very helpful, not too complicated, and recommended, but it is not a necessary component. Results, if not too extensive or complex, can be gathered by "eyeballing" the card groupings.[7]

Case Study

In the case of Cornell Library's Gateway Help System, the librarians working on content organization repeatedly tried to improve the Help organization and TOC by reacting to previous iterations of this task. They realized that although they were changing things around and modifying the wording, it was impossible to escape the initial structure, which was not entirely satisfactory. To help obtain a more objective view, as well as to ensure that the online user help system is designed with the user in mind, the committee decided to do a user study employing the card-sorting technique.

Participants

Twelve individuals participated in the study. Because the Gateway Help system is for all users of the university population, the committee chose a random sample from the academic community. None of the subjects were required to be library users or to have any special knowledge of the Cornell University Library Gateway. In the hope that the help component of the Library Gateway would be instrumental in assisting all levels of library and Gateway users, both novice and experienced library and Gateway users were welcome. The study population consisted of five undergraduate students, two graduate students, two faculty members, and three staff members. Because the sample was small, names were chosen from the campus telephone directory, using every eleventh name on every thirteenth page until there were enough participants.

Methodology

Each of the Gateway Help topics was printed out on a label and affixed on an index card. This resulted in fifty individual cards. The cards were numbered from one through fifty on the reverse side, and then shuffled and spread out on a table. A study participant was then invited into the room. Participants were told that the study was being done to aid in organizing items on a library Web site. Then they were instructed to sort the cards by placing similar cards into piles. Users were asked to try not to make piles of a very few or a great many cards but were given no other instructions on how many cards to put in a pile. They were encouraged to ask questions or to request further clarification of concepts at any time. After the subjects finished sorting, they were asked whether any piles bore any relation to one or more of the other piles they had assembled. When the study was completed, a rubber band was placed around each of the individual piles. A second rubber band was then placed around individual piles that participants indicated to be related to other piles. The

TABLE 1
Dissimilarity Matrix

	card 1	card 2	card 3	card 4	card 5	card 6	card 7	card 8	card 9	card 10	card 11	card 12	card 13	card 14	card 15
card 1	0	1	1	1	3	3	1	3	0	1	1	1	1	0	3
card 2		0	1	1	3	3	1	3	0	1	1	1	1	0	3
card 3			0	1	3	3	1	3	0	1	1	1	1	0	3
card 4				0	3	3	1	3	0	1	1	1	1	0	3
card 5					0	1	3	1	0	3	3	3	3	0	1
card 6						0	3	1	0	3	3	3	3	0	1
card 7							0	3	0	1	1	1	1	0	3
card 8								0	0	3	3	3	3	0	1
card 9									0	0	0	0	0	2	0
card 10										0	1	1	1	0	3
card 11											0	1	1	0	3
card 12												0	1	0	3
card 13													0	0	3
card 14														0	0
card 15															0

This illustration is a subset of one user's groupings. Cards 9 and 14 appear in a pile that was not related to any others, so a score of 2 was assigned at the intersection of 9 and 14.

length of time for each participant varied from twenty minutes to one hour. The results were recorded onto a dissimilarity matrix on a spreadsheet and a distance cluster analysis was run on the data.

The cluster analysis, in this case, simply creates a graph describing the cards as they were grouped together.

Recording the Results

Although it is possible simply to look at the various cards to ascertain the way in which a particular user organizes the information in question, it proved to be more manageable to score the cards so that a statistical analysis could be run. This allowed the results to be stored, accessed, and manipulated electronically. Scores were assigned to the cards based on their relationships to the other cards. On a spreadsheet, card numbers were listed sequentially along the x (horizontal) and y (vertical) axes (see table 1). Numbers using a card's relationship to another were typed into the intersection. A dissimilarity matrix was used so that lower numbers were assigned to cards with stronger relationships. The cards from each participant's session were given a numeric score based on how they were grouped. The scores were given as follows: a score of zero indicated no relationship of a card to any other card; a score of one was given to cards that appeared in different piles that the subject reported had some relationship; a score of two indicated that the cards appeared in the same pile, but that the pile was not related to any other pile; and a score of three meant that the cards appeared in the same pile and that the pile had some relationship to other piles.

Using this scoring method, card piles from a single subject were examined and the scores were entered into the spreadsheet. If a particular pile did not show any

relationship to any other pile, it was not grouped by rubber band to any other pile. Each of these cards received a score of two. So, for example, if card number nine and card number fourteen were in this pile, the intersection of these two cards on the spreadsheet would get a score of two. Given a case in which two piles were related, the cards in these piles would be scored twice. First, the cards in each of the individual piles would receive a score of three because they appeared in the same pile. The concepts represented by the cards in this pile were considered more strongly related than those in a pile not related to anything else. Then the cards in one pile that bore relation to another would also get a score of one. In the first pile was card number forty and in the second pile was card number twenty. The intersection of these two cards would get a score of one because the cards appeared in different piles that the subject reported had some relationship. For a total of fifty cards, the scoring process took approximately one-half hour per subject.

These scores were in turn entered into the spreadsheet file created for each subject. Using the statistical software SPSS 6.1, a cluster analysis was run on the data.[8] This resulted in a forked tree graph (dendrogram) that visually indicated which items should be grouped together per each user or a composite of users (see table 2).

Cluster Analysis

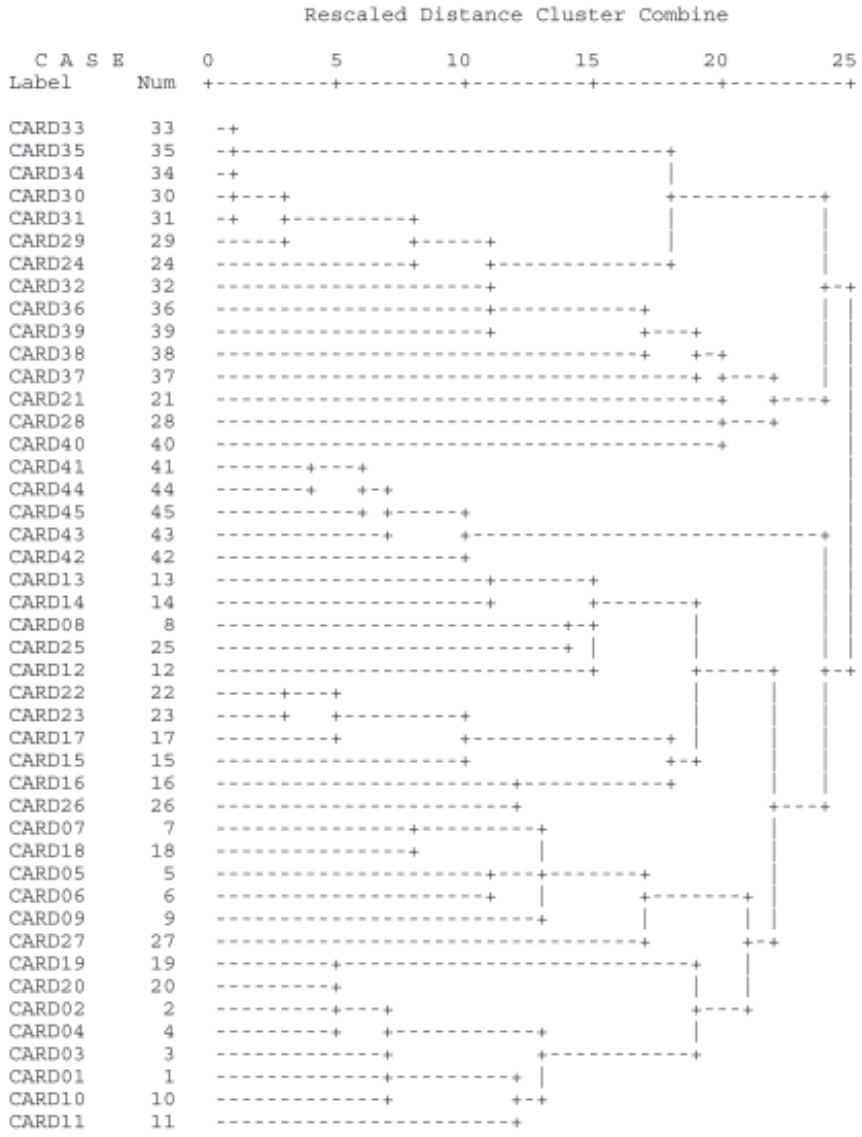
Cluster analysis is a technique often used in the sciences to measure the "distance" between objects. The distance analysis creates a "single-linkage dendrogram" that simply shows branches where clusters close to one another are considered to have a stronger relationship to one another than those on another branch.[9] In this study, a cluster analysis was run on each participant's grouping of cards. Using the tree as a guide, one could easily recreate the piles of cards for an individual participant. The cluster analysis, in this case, simply creates a graph describing the cards as they were grouped together. In this sense, scoring the cards, entering the data, and running the analysis provides a clear and visual depiction of the work done by the users in a format that can be manipulated. True statistical analysis only occurs when working with a mean sample. By "eyeballing" the cards and groupings, one could get an idea of an individual's organizational structure, but an aggregate structure could only be an approximation.[10]

Using the Results

The data were subsequently used to create the new organization and TOC to Gateway Help. As Rickey E. Savage, William E. Hutson, and Richard Cordes suggested, a "composite of the various user's cognitive models can be used to structure the user–computer interface such that a reasonable and close match can occur."[11] Because no one design or organization will match all user mental models, the dendrogram made up from a composite of all study participants was determined to be most reflective of the study's varied population and, therefore, the committee's work was based on this. In certain cases, a composite of a certain subgroups might prove very meaningful. For example, if the system being designed were going to have novice and expert interfaces, a composite of novice users might be used to organize the first interface and a composite of expert users would be used in designing the latter.

The initial approach to reorganization was to take the dendrogram and translate the topics represented into a hierarchical TOC. However, doing so took the committee back to the original problem of not being able to see new ways of organization as soon as a hierarchical structure was imposed. Abandoning this approach, the committee then plotted the clusters in a visual cluster map (see figure 1). This provided a visual representation of the users' groupings. All categories that appeared as highly related were plotted randomly in one cluster. New relationships immediately became apparent from this display. It did not take long to

TABLE 2
Dendrogram Using Average Linkage (between Groups)



come up with an entirely new organizational structure for the TOC and, therefore, the system.

In some cases, the committee decided to alter the organization suggested by a composite of all user studies. For example, the technical information on how to set up telnet software was included with the details on how to search for materials in the online catalog. The committee determined that the technical information should go with the technical setup

and troubleshooting section and that the searching information should go with the research tools section. The results of the study did not dictate the committee's decisions but did provide more objective insight and direction.

Strengths and Weaknesses of the Study

The committee considers the study to be a success and a highlight of its work. Yet, there are both strengths and weaknesses that should be considered before embarking on such a study.

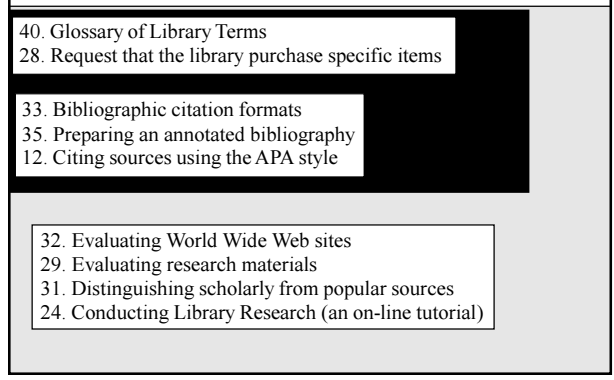
The main drawback was the time commitment. Scoring the cards and entering the data into the spreadsheet was particularly time-consuming, but the study provided a relatively easy procedure for data collection. Generating all of the various categories for sorting had to be done whether the study was to be implemented or not. Creating the cards, administering the study, and entering the data into a spreadsheet took time but was not difficult or mentally taxing. Scoring the cards was the most challenging aspect, but after the procedure was understood, the task became much less complex. After statistical software is set up to run cluster analyses, the analysis itself takes only a few minutes to run.

This type of user study happens outside the actual system and thus is stripped from true context. In other words, after the cards are grouped, what is learned is how an individual or a composite of individuals combines like concepts. However, this does not explain how effectively users will find relevant information in the final Gateway Help system. James E. Palmer and colleagues questioned "whether the card sort technique presents any information about the mapping between real world tasks and the corresponding computer task."[12] A postdesign user study investigating user ease in finding information in the system is recommended.

To what extent does the wording on the cards influence the way the subjects group cards? For example, if ten cards include the words *information* or *directions*, even if the underlying concepts are disparate, it is possible that the mere usage of a word may have an effect on how piles are assembled. The committee recommend that wording on cards be chosen and constructed carefully. Users also should be instructed to think of the concepts behind the words more than the words themselves.

Another limitation to using the card-sorting system for Web site design lies in the singular nature of physical cards: a user is not able to put one card into two places if a concept falls into more than one category. Hypertext, conversely, allows for one page or concept to reside in multiple places. If a representative sample had been used, this problem might have been mitigated because the chance that another user might find the second relationship has increased. Still, it would be impossible to get an exact picture of how the user sees the information landscape because of the limitation of unvarying cards. On the other hand, the static nature of the cards forces the user to select what he or she considers the strongest relationship.

FIGURE 1
One Cluster Section f the Concept p



Furthermore, it is possible to place a concept in more than one place in the final product. The inflexibility of card placement, therefore, should not discourage use of the card-sorting technique.

Perhaps the most significant outcome of the study was the objective and fresh insight it provided.

The card-sorting test, as CUL used it, works best with relatively small groups. A test with more than seventy-five cards not only could overwhelm the test subject but also give results that are difficult to score and interpret. However, it is possible to run a cluster analysis with a much larger group. If the researchers use something other than a single linkage method, it is possible to create graphs other than dendrograms, some of which may be easier to interpret with a large data set. The card sorting is ideal for data sets of about fifty concepts.

The study was very valuable in helping the committee gain perspective on how users would organize the help system if given the opportunity. The fact that the committee was able to combine this input and the expertise of its members added to that value. The study simply provided input. The committee does not recommend strict adherence to the resulting data tree. Rather, it used the results as a guide in creating the TOC while using its collective expertise to decide when to disregard results that did not seem to make the best sense. A system custom fit to every end user is impossible. The "structure of the user–computer interface will probably never match any specific user's cognitive model because no two users will have the same cognitive model."[13] A librarian's information organization expertise is crucial to refine the quantitative information into a system best suited to all of its users rather than specific individuals.

Perhaps the most significant outcome of the study was the objective and fresh insight it provided. Whenever there were competing views on content organization, the committee could refer to the study results for direction and inspiration. The study helped the committee see relationships and patterns that it could not see by critiquing and analyzing its own work. In the end, the organization of help contents changed significantly.

Conclusion

The card-sorting technique proved to have a very positive effect on the design of CUL's Gateway Help for several reasons. First, and most obviously, the committee was able to incorporate user input on the organization of its interface before total system design, resulting in a structure with which everyone was satisfied. Moreover, the simple satisfaction of incorporating the user's point of view had a tremendous impact on the committee's confidence.

Taking the time and devoting the resources to do this study also led to greater confidence in the committee's decisions. When discussing ideas about content organization either within the group or with other CUL librarians, the committee had objective data to point to that could help it arrive at and justify its decisions. It is a given that each individual will have an idea of how the site could be organized best, and it is much easier to arrive at agreement when user input exists. Librarians, whose ultimate goal is to serve the user, should be willing to forego personal views for those of the users, which are garnered from the study.

In summation, the card-sorting technique provides an excellent and relatively simple means of gathering user input into how a set of concepts, such as those in a system or interface, can be organized. This is particularly valuable because the input is gathered prior to total system design and avoids design being driven by a presupposed structure or organization. Although certain limitations are inherent to the study, the card-sorting technique can positively impact system and interface organization decisions and the committee highly recommends it.

Notes

1. Karen Calhoun and Zsuzsa Koltay, *Gateway Focus Groups Report* (Cornell University Library, 1999 [cited November 30 1999]); available from http://www.library.cornell.edu/staffweb/GateEval/contents.html.

2. James E. Palmer et al, "The Design and Evaluation of Online Help for Unix Emacs: Access Mechanisms" (paper presented at the Second IFIP Conference on Human–Computer Interaction, University of Stuttgart, 1987), 461–66.

3. Jeffrey Rubin, *Handbook of Usability Testing : How to Plan, Design, and Conduct Effective Tests* (New York: John Wiley, 1994).

4. Jakob Nielsen and Darrell Sano, "SunWeb: User Interface Design for Sun Microsystem's Internal Web," *Computer Networks and ISDN Systems* 28 (1995): 179–88.

5. Johan Berndtsson, "Designing an Intranet from Scratch to Sketch: Experiences from Techniques Used in the IDEnet Project" (paper presented at the Thirty-Second Annual Hawaii International Conference on System Science, Maui, Hawaii, Jan. 5–8 1999), CD-ROM liii+341.

6. Lucy Lockwood, *Learning the Laws of Usability* (1999); available from http://www.sdmagazine.com.

7. Nielsen and Sano, "SunWeb," 182.

8. SPSS for the Macintosh Ver. 6.1.1, SPSS Inc., Chicago.

9. Samuel Kotz and Norman L. Johnson, eds., *Encyclopedia of Statistical Sciences* (New York: John Wiley, 1983).

10. Nielsen and Sano, "SunWeb," 182.

11. Rickey E. Savage, William E. Hutson, and Richard Cordes, "Using a Composite Cognitive Model for Designing User–Computer Interfaces" (paper presented at the 1986 IEEE International Conference on Systems, Man, and Cybernetics, Atlanta, Ga., 1986), 435–38.

12. Palmer et al, "The Design and Evaluation of Online Help for Unix Emacs," 466.

13. Savage, Hutson, and Cordes, "Using a Composite Cognitive Model for Designing User–Computer Interfaces," 435.