

Redesigning and Evaluating the Census Bureau's Intranet Site for Usability: Methods and Lessons Learned

Elizabeth D. Murphy, Susan Ciochetto, and Carol I. Bateman

U. S. Census Bureau¹
Washington, DC 20233

elizabeth.d.murphy@census.gov
susan.m.ciochetto@census.gov
carol.isak.bateman@census.gov

Abstract

The U. S. Census Bureau's Intranet site needed to be redesigned to make it easier for employees to find information on the site. As previously designed, the site reflected the organizational structure of the Census Bureau, but new employees did not necessarily know which division or branch was responsible for what information. The objective was to redesign the site using functional, logical categories for the information content. This paper focuses on the methods that were applied to the redesign effort, including card sorting and usability testing with an eye-tracking component. The use of card sorting is recommended along with various usability-testing methods for intranet evaluation and redesign; but the use of eye tracking must be carefully targeted to avoid spending excessive time on data analysis.

Introduction

The motivation for redesigning the U.S. Census Bureau's Intranet site came from upper management's recognition that new employees were having trouble finding information at the site. The original intranet site had grown out of a grass roots effort with units within the Census Bureau developing individual Web sites. The intranet looked like a patchwork quilt; each patch had its own look and feel and contained diverse content. No two individual Web sites contained even the same basic elements. Because the design of the site reflected the hierarchical structure of the Census Bureau, new employees had to know which organizational unit was responsible for the information in question in order to find it. Even then, there was no guarantee that the information would be available on the site. The objective was to redesign the intranet site using functional, logical categories for the information content and a common look and feel across individual Web sites. Such a framework could be reasonably expected to better support successful information retrieval for all employees.

The Intranet Redesign Team used several methods to evaluate the content and layout of the proposed design. A user survey and a web-log-usage analysis established the list of content items that were important to users. Card-sorting techniques helped the team in organizing content items to develop an information architecture or sitemap that included high-level categorization of the content items.

The results of card sorting and other content evaluation techniques enabled web designers to produce a layout based on the proposed sitemap. Design review sessions and several iterations of usability testing, one including an eye-tracking component, helped in further refining the design. The result is a Web site designed and tested around the interests and needs of users. This paper provides details of the methods and lessons learned from each method. At the end, we reflect on the value of using the various methods and assess the overall process of redesigning the Census Bureau's Intranet.

¹ This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the authors and not necessarily those of the U. S. Census Bureau.

Content Evaluation Methods

The following methods contributed to developing the content of the redesigned intranet site:

Card Sorting

The Intranet Redesign Team began with an analysis of the content of the original intranet site. Working with a contractor, the project team identified and grouped over 140 terms, including some new items suggested by team members as well as many from the original intranet site. The project team then asked the staff of the Usability Laboratory to obtain some input from users to confirm/disconfirm the logic of the team's grouping of topics. The project team wanted to see how actual intranet users would organize the topics irrespective of the Census Bureau's hierarchy. The usability team chose card sorting as the method to answer this question because it is a method used widely to elicit groupings of terms from users themselves (e.g., Lazar, 2001; Gergle, Brinck, & Wood, 2002; UsabilityNet, 2003). The groupings are interpreted as representing the sorter's mental model of the domain from which the terms are taken. Cluster analysis was used to merge the sorts created by the individual participants. Designing the information architecture to resemble a composite mental model derived from users was intended to support the user's ease of finding information.

Usability staff conducted a card-sorting exercise with 10 internal users of the original intranet site. According to the card-sorting literature, five to 10 participants is sufficient to extract users' conceptual models of how they expect Web site content to be organized (e.g., Lazar, 2001). These users came from various organizations within the Census Bureau and represented a range of experience working at the Census Bureau, from a few months to decades. Some were managers.

Ninety-nine² terms were placed on cards, which were shuffled and spread out on a table. In separate sessions, participants sorted the cards into groups, based on the instruction to sort the terms for functional (logical) similarity. Once they had sorted the cards into piles, participants were asked to name their groups of terms with a heading or category label.

Results of the sorting exercises were input into IBM's freeware tool, EZSort (Beta version 1.8); and the results were analyzed by IBM's cluster-analysis tool, EZCalc (version 1.3). EZCalc presents the combined groupings in the form of a hierarchical tree structure. Deciding on the high-level groupings is not strictly straightforward but is guided by the computed results. The application generates several sets of results, one that puts greater emphasis on rated similarities between pairs of terms and one that places greater emphasis on rated differences. A third set of results represents a balance between the other two. Following the recommendations of the tool developers (Dong, Martin, & Waldo, undated), the usability team extracted the high-level groupings from the results for similarity and the low-level groupings from the other results.

Card Sorting: Round 2

The Intranet Redesign Team also conducted a second, less formal round of card sorting using team members as participants. The team used the results from the first card sort and from the usage analysis as input to the process. Each team member was given an Excel spreadsheet with the names of categories and the content items within each category. A round table discussion determined categories with an excessive number of content items or ambiguous category names. Using Excel, each team member made an attempt to group items into smaller categories using a previously decided maximum of three levels. Some grouped items were condensed or eliminated. Content items that seemed out of context within the grouping or categories were put aside.

Items out of context, content items from ambiguous categories and items representing groups that were resolved in discussion, were then written on "post-it notes" for posting on a white board. The team then arranged and discussed items or groups until more specific categories were created. The team then clarified

² The IBM analysis tool's limit was 99 terms. This tool is no longer available at the IBM Web site.

high-level category names. Based on the categories and grouped items, a final sitemap was created and distributed for review as an Excel spreadsheet. The process was repeated until the team was satisfied.

User Survey

The Intranet Redesign Team conducted a user survey to get a better understanding of what sites users of the intranet used most and what improvements or suggestions they would like to see on the Census Intranet.

The survey asked for the following information:

- Q1. Division or Office Name:
- Q2. What is your most important purpose for using the Census Intranet?
- Q3. (a-e.) List the Census Intranet sites or services that are most important to you (please list no more than 5).
- Q4. When you use the Census Intranet what is the first site you use or start from?
- Q5. What type of information would you like to see that is not currently available on the Census Intranet?
- Q6. Please offer any other comments or suggestions for how we can improve the Census Intranet.

Other than the selection of division or office name, all questions were open ended. The survey was web based and accessible to employees by a link from the original intranet home page. An email message was sent to all employees announcing the survey.

Data were collected into a Microsoft SQL server database. The analysis of questions was divided up between project team members. Each team members was asked to convert data for a specific question into an Excel spreadsheet and to become familiar with the answers. The team then performed a general review and search for keywords to group similar items into broad categories.

Over 500 employees responded to the user survey. The team compared the respondent's selected division or office to the official employee list in the Commerce Business System database to ensure divisions were well represented in the survey results.

Usage Log Analysis

The purpose of the intranet site usage log analysis was threefold: 1) to determine the web pages that were viewed most often; 2) to confirm/disconfirm the card-sorting results; and 3) to identify any content items that had been omitted from the card sort. The Intranet Redesign Team used the results to determine common categories and to identify content items that should be highlighted on the site.

The usage log analysis was based on web server logs for a three-month period of time. Web statistics were tracked using the SAS Webhound tool (SAS Institute, 2003). The team wanted to use recent information but cover a period of time long enough to avoid spikes due to a single announcement or event. However, the decentralized nature of the Census Intranet required data to be collected from four different web servers. This required importing daily logs from each server into SAS datasets. To keep processing to a minimum, the team felt three months was adequate. (Note: The four servers do not account for all web servers in use at the Census Bureau. The four that were chosen represent major content, including administrative services and human resources, and were available to the team for analysis.)

The SAS Webhound tool produced individual reports for each web server. A report showing the twenty-five most requested pages was used for analysis. The project team cross-referenced reports from each web server to determine the most requested pages overall. The team then developed a list of intranet content items that were most important to users. These content items were used in the second card sort and in producing the sitemap.

The team also reviewed a list of the words most frequently used for searching from the Census Bureau Home page and the Human Resources Web site. This review identified content items that were labeled insufficiently or that were difficult to find under the current design.

Design Evaluation Methods

This section discusses the methods used to evaluate design prototypes for usability. As shown in Table 1, methods used for content and design evaluation were interwoven over the course of the redesign effort.

Table 1: Timeline of Intranet Redesign Milestones with Content and Design Methods Highlighted

Date	Milestone Activity
Jul 2003	Deputy Director requested intranet redesign.
Oct 2003	Contractors hired to provide initial design alternatives
Oct - Mar 2004	Design alternatives developed in graphical form. Team design review sessions
Mar 2004	Initial card sort
Mar 2004	Stakeholder review of preliminary design
Jun 2004	Intranet user survey
Aug 2004	Usability study conducted using design alternatives and eye tracking
Aug 2004	Web site usage log analysis
Aug 2004 - May 2005	Prototype developed. Team design review sessions
Jan 2005	Team card sort
May 2005	Team design review session with business/program areas
May 2005	Usability study using working prototype
May - Jul 2005	Prototype modified. Team design review sessions
Jul 2005	Usability review session

Design Review Sessions

The Intranet Redesign Team conducted design review sessions during various stages of development. Design reviews can be performed on simple designs called wireframes, static graphical images, or working prototypes with limited or full functionality. A design review session allows designers and developers to obtain feedback before resources are expended on a design that does not work as intended or on functionality that is unnecessary. Participants in design reviews sessions can vary to include both stakeholder reviews for testing and buy-in as well as expert reviews to assess usability, accessibility and implementation issues.

A stakeholder review took place early in the design process. The Intranet Redesign Team identified a list of participants who were responsible for intranet content and could represent their area of the Census Bureau. Seventeen people attended two sessions each. The idea was to obtain feedback on the design, identify implementation issues early and gain buy-in. The first session described the project, explained the sitemap, and introduced the design in the form of wireframes. The second session reviewed the design in detail and obtained feedback in a question-and-answer format. A tape recorder was used to record feedback.

The Intranet Redesign Team also conducted several iterations of expert reviews within the team itself. Team members varied at times but most of the time consisted of web developers and major content providers of the intranet, including members from the Information Technology, Human Resources and Administrative Service areas of the Census Bureau. Additional members were added for different sessions when needed and included members from the business or program areas of the Census Bureau. The purpose of the expert review was to compare the design requirements to local standards, guidelines and best practices. Many of the design review sessions occurred after major changes were implemented as a result of usability studies. Sessions consisted of a review of the site with feedback provided in a round table discussion. Some sessions included a web developer to implement changes instantaneously and allow for immediate feedback.

Usability testing

To test the look and feel of the redesigned intranet pages, the Usability Team conducted two rounds of usability testing. The first round took place after the card sorting exercise; the second round was near the end of the redesign effort, prior to launch of the redesigned intranet.

Round One Usability testing

Working with the Intranet Redesign Team, the contractor developed two different sample layouts (“wireframes”) and sample screens (“composites”). These prototype screens had the same content but differed in the colors and graphics used and the general look of the page (“traditional” versus “contemporary”). The purpose of the usability testing was to determine whether one of the prototype designs was “better” than the other on various dimensions, including accuracy and user satisfaction.

Figures 1 and 2 show the home pages of the two designs.

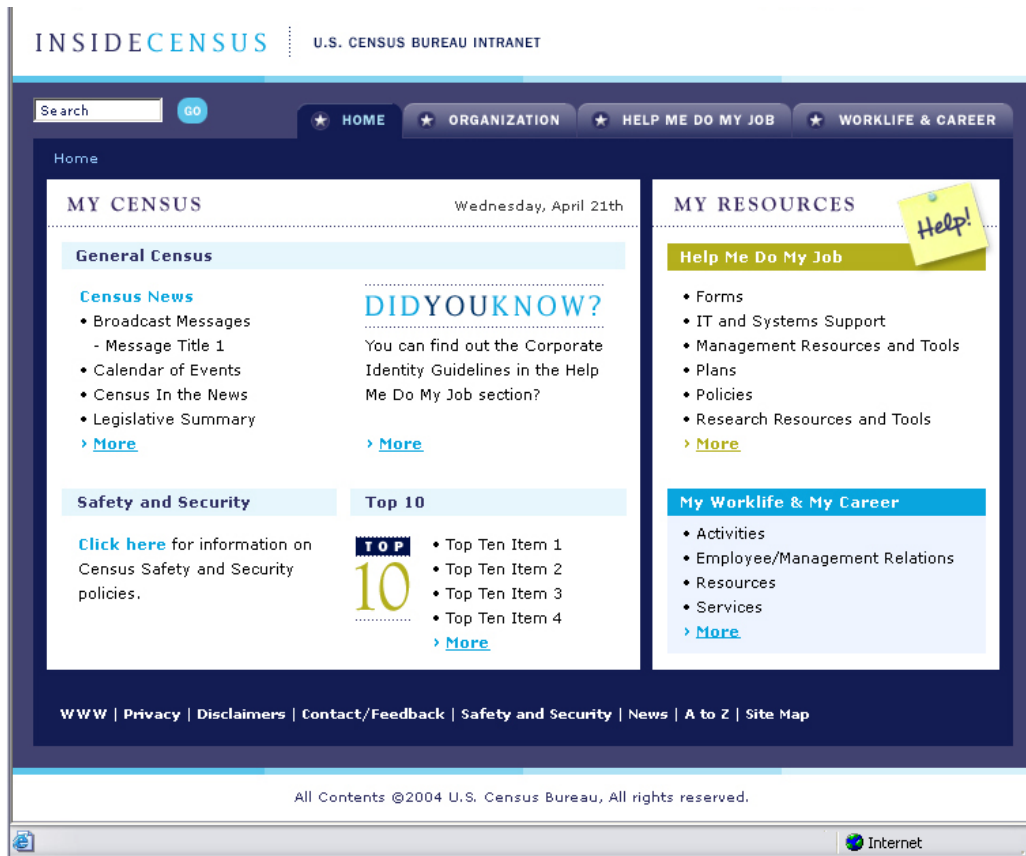


Figure 1: Home page for the "traditional" prototype

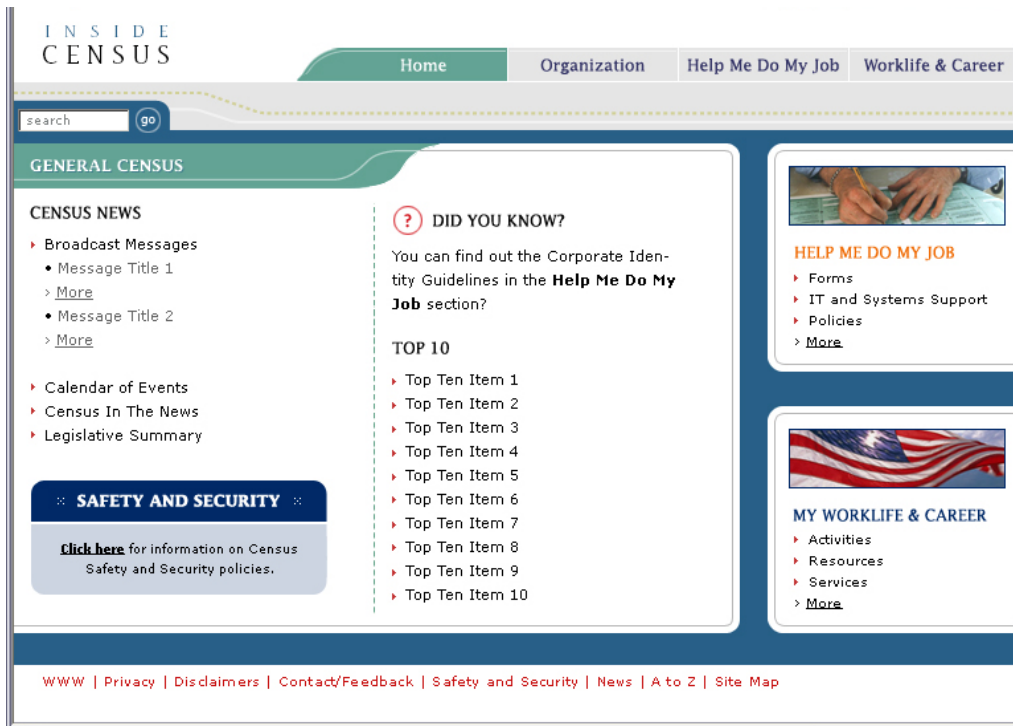


Figure 2: Home page for the "contemporary" prototype

The Usability Team used these two sets of prototype screens for usability testing. These graphics were static, and the user could not interact with them. However, substituting for the computer, the test administrator could bring up the individual sample screens as the test participant proceeded through the tasks.

In the first round of usability testing, the Usability Team chose to supplement traditional think-aloud usability testing with an eye-tracking component to obtain more precise measures of where the user was looking. This was important because where the user looks changes so rapidly that it is impossible to understand what they are looking at just by observation or self-reporting. Also, since the Usability Team was comparing two different designs, precise measures were needed as a basis for comparison. The Eye-gaze Response Interface Computer Aid (ERICA) was the system used to record eye movements (Eye Response Technologies, Inc., 2003).

Since this was the Usability Lab's first experience in using eye tracking, a literature search for information on the measures used by other researchers was necessary. The literature tentatively associates the following eye tracking measures with various human-performance outcomes, such as search efficiency and processing time:

- **Scanpath length.** As participants search for the target during a task, their eyes perform a succession of rapid movements (saccades) and fixations. During a fixation, visual attention is focused on a particular area of a display for approximately 200 to 300 milliseconds (Pan et al., 2004): A sequence of saccades between fixations defines a scanpath. The length of a scanpath is equal to the total distance between all fixation points, starting with the first fixation and ending with the last fixation. Longer scanpath lengths are interpreted as indicating more searching on the part of the participant (Goldberg & Kotval, 1998, 1999). Shorter scanpaths are thought to reflect more efficient search strategies (Goldberg, Stimson, Lewenstein, Scott, & Wichansky, 2002).
- **Average number of fixations.** A fixation is a brief period of time during which the eyes do not move. The number of fixations is taken as an indication of the number of units of meaning (text,

icon, graphic, etc) processed by the user during a task (Goldberg & Kotval, 1998). Consequently, a finding of fewer fixations before selecting the target suggests that the design is requiring the user to process fewer items in order to find the target.

- **Average duration of fixations.** The temporal length of fixations is another hypothesized indicator of processing demands. Longer fixations are associated with longer time to interpret the fixated items. Displays that are more tightly packed with content are associated with longer fixation durations as compared to less dense displays (Goldberg & Kotval, 1998).
- **Time in Look Zones.** Look Zones are rectangular regions of the screen, such as headers and bordered content areas. Eye tracking software records gaze paths between Look Zones, providing a record of eye movement sequences during a session. The software also records time spent within each Look Zone (“dwell time”). Research suggests that longer dwell times indicate greater relative informativeness of Look Zones (Pan et al., 2004) but lower search efficiency (Goldberg et al., 2002).

Participants were recruited from within the ranks of Census Bureau employees at the headquarters facility in Suitland, Maryland. From those who responded to the call for participants, the researchers selected 16 people to represent a cross-section of divisions and demographics.

Testing took place in the Census Bureau’s Usability Laboratory in August 2004. Prior to the test session, participants’ eyes were calibrated to the ERICA eye tracking software (Eye Response Technologies, Inc., 2001). It was possible to calibrate most of the participants, and the eye tracking software was used to record their eye movements during the first part of their test sessions.

Eye tracking data were collected as follows: The participant placed his or her chin in the chin rest and his or her forehead against the headrest of an apparatus meant to prevent head movements. Focused on the user’s eye, a computer model LCL-902K camera computed the position of the user’s gaze. The GazeTracker software recorded this position information for later analysis. The ERICA software was used to analyze the data from the camera. Both the ERICA and GazeTracker software programs ran passively on the user’s computer throughout the session (Eye Response Technologies, Inc., 2003).

During the testing session, the test administrator remained in the test room with the participant to manage the presentation of the various static screens. In the first phase of a usability testing session, the eye tracking software was used to record where on the screen the participant looked as he/she completed seven tasks (eight if the participant was a manager). Next, two more tasks were completed using the think-aloud protocol, i.e., the participant provided a running commentary while performing each task. After completing the tasks, the participant completed a tailored version of the Questionnaire for User Interaction Satisfaction (QUIS). The QUIS is a standardized instrument developed by the University of Maryland (Norman, Shneiderman, Harper & Slaughter, 1998). The participant was then debriefed on his/her reactions to the user interface.

Round Two Usability Testing

Usability testing of working prototype screens took place in May 2005. Testing was again held in the Census Bureau’s Usability Laboratory. The eight test participants represented a range of experience and levels of responsibility (e.g., technical staff, management). Participants completed a set of tasks that required them to locate target items on the prototype screens. These sessions were conducted without eye tracking but using a think-aloud procedure. The test administrator sat in the testing room with the participant to facilitate probing and discussion. Using this method, the Usability Team identified user interface design issues and then discussed ways to resolve them with the developers.

Usability Review

As a final review prior to launch, the Intranet Redesign Team asked the Usability Team to conduct a usability review in a focus-group setting. Originally, the Usability Team planned to work with two small groups of actual users: six managers and six non-managers. However, because of scheduling conflicts the actual groups contained a mix of managers and non-managers. The participants were volunteers

representing a wide range of internal organizations, with varying lengths of employment at the Census Bureau, from a few months to over 30 years. The redesigned user interface included the top two navigational levels of the Census Intranet. The purpose of this review was to gather feedback from users prior to launch.

For the most part, the Usability Team used tasks that were identical to those used in previous usability testing sessions. However, the team needed more tasks for managers. As a basis for formulating these tasks, usability staff members interviewed several managers about their use of the intranet. This was helpful in identifying tasks that management-level participants would find reasonable and familiar.

The environment for the review was a computer-equipped classroom. Participants were asked to perform a series of tasks designed to mimic their own daily use of the intranet. Tasks varied in complexity, i.e., the number of steps required for completion. After each task, the facilitator led a discussion of the perceived usability of the site for performing that task: How easy was it? Did the user interface help or hinder the participant in performing the task? Do the participants have any suggestions for improving the user interface? Members of the development team observed these sessions and participated by asking questions and by clarifying the rationale for various design decisions.

Lessons Learned: Content Evaluation Methods

Card sorting

The Usability Team found that formal implementation of this method was useful for gaining insight into the user's perspective. Deciding on the groupings was guided by the EZCalc results but also involved a certain amount of expert judgment. Although there were often commonalities across the participants' ratings, there were also many instances of unique groupings. Some terms clearly "belonged" in one group, while others were ambiguous enough that they could logically belong in more than one group. Differences between users' job responsibilities naturally affect their view of how to sort the set of terms. Thus, there is a need for a systematic method for combining the results of the sorts across users. Expert interpretation is still needed, however, because the different methods of cluster analysis produce different results. Even when the individual groupings are combined, it makes good design sense to provide alternative paths to the same item, especially for ambiguous items.

Similarly, informal card sorting with the redesign team was helpful for the refinement of generalized categories and for clarifying definitions. This was useful for items where users were not familiar with various terms and erroneously grouped them. These items usually represented technical terms or jargon. Additionally, the refined card sorting with the design team helped to ensure the categories and content items adapted well to layout or design specifications. For example, keeping groups and categories to three levels was important to ensure users would find items quickly. Lastly, the team's enhanced knowledge helped them prioritize items within the hierarchy based on other input such as log usage analysis. One warning note is that the team must be careful not to let their enhanced knowledge of the subject influence the outcome.

User Survey

The user survey was helpful in gaining a general idea of user requirements. The surveys also involved actual users in the design process. Involving users may prove beneficial to increase buy-in later on and provide for an easier transition when implementing changes.

The analysis of open-ended responses overwhelmed the Intranet Redesign Team because of the sheer number that they had to process. This proved too time-consuming and required too much subjective analysis. Multiple-choice questions are recommended with fill-in options for choices not provided.

Usage Log Analysis

The web server log usage analysis was useful for verifying the actual use of items that users had identified as important in the user survey. That is, this analysis served as a reality check on self-reported data. It was important that the team agreed on a specific time period that best represented actual usage. Note that it is possible that a specific announcement or event can skew results.

The team found it difficult to compare logs from many different servers because the Webhound software produced individual reports for each server. The team had to carefully compare counts across reports. Individual reports gave the impression that web pages were more important than they actually were because they showed up higher in the individual report but had a lower count when compared across reports. Using an analysis tool that compares logs across servers would be beneficial.

Lessons Learned: Design Evaluation Methods

Design Review Sessions

Design review sessions are valuable to obtain feedback from a smaller representative sample when testing designs, pilots or prototypes. They can also be used to obtain managerial or stakeholder buy-in. Careful attention should be made when choosing participants. Members should be representative and consist of executives and senior managers through professional and clerical staff. By limiting the stakeholder review to content providers, the Intranet Redesign Team found it difficult to obtain feedback on the general design. Participants concentrated on how their own content or work would change.

The Intranet Redesign Team changed members as stages of the project progressed. This allowed for diverse input during design and allowed team dynamics to fluctuate so that no one opinion dominated. It was particularly valuable to vary team makeup for expert reviews.

Usability Testing with Think-Aloud Protocol

Multiple rounds of usability testing are necessary, especially when prototypes are changing frequently. It would be useful to have a means of tracking usability recommendations. Observation of actual users using the intranet in their daily work would have been useful as a basis for developing tasks for test participants.

Eye Tracking

Since this was the Usability lab's first foray into eye tracking, it would have been a good idea to do selective eye tracking. Using the full functionality of the system generated huge volumes of data. Analysis of the data took far longer than expected, even though a summer intern provided additional skilled assistance. The Census Bureau Usability Lab is investigating a different system that does not require the participant's head to be immobilized. In the future, the researcher should plan specific analyses using the eye-tracking data and should not attempt to interpret all the data.

Usability Review

This method was used to obtain quick feedback late in the development cycle. It showed that users were now able to find items that previous participants had been unable to find, mainly because the items had been brought out of obscurity. The Usability Team does not recommend mixing managers and non-managers for this kind of exercise. Managing the flow of discussion proved to be somewhat of a challenge in these mixed groups. If the groups must be mixed, it would be better to have everyone do the same tasks so that everyone can participate in the discussion.

Epilogue

Redesigning the Census Bureau's Intranet involved redesigning both the content and the user interface. Design of the content necessarily influenced the design of the user interface and *vice versa*. When the team completed the first round of usability testing, for example, it became apparent that further content development would be necessary before conducting a second round. It was often the case that further input was needed from users before continuing with the design of the content or the user interface. Thus, as reflected in Table 1, a natural flow evolved of going to the users, doing design, and going back to the users.

For purposes of designing or redesigning a Web site, the Intranet Redesign Team recommends that usability of both the content and the user-interface design be considered early in the planning process. The extra time and effort required must be scheduled accordingly to avoid costly delays in the project. Enough flexibility must be built into the schedule to allow for iteration of methods, such as occurred in this effort. The payoff for early investment of time and effort comes in the diminished need to retrofit the design later when it is difficult or even impossible to make changes. Content and user interface must be designed in tandem, not in isolation from each other.

The availability of a usability lab at the Census Bureau was beneficial. Although many of the methods can be applied informally by the project team itself, the use of the lab was much more efficient, and the added expertise was invaluable in justifying design decisions. A well-organized and efficient intranet site that satisfies users and developers is well worth the investment in usability studies. However, each method must be evaluated on its own merit with careful consideration to the overall work involved when analyzing results.

The usability methods applied during the Census Intranet redesign resulted in a design that works well for the Census Intranet users who participated in the final usability review and satisfies Web site developers. Involving the user community with design decisions increases buy-in and aids developers in understanding the need for certain design elements. The use of various usability methods throughout the design process is necessary to achieve effective results. The true test will come as Census Bureau employees at large begin to use the redesigned intranet and gain experience in its layout and topical organization of content. The usability team recommends another user survey about six months after launch of the redesigned site.

References

- Dong, J., Martin, S., & Waldo, P. (undated). "A User Input and Analysis Tool for Information Architecture." Retrieved on 3/23/04 from http://www-3.ibm.com/ibm/easy/eou_ext.nsf/Publish/410 (no longer available).
- Eye Response Technologies, Inc. (2003). *ERICA System Reference Manual*. Charlottesville, VA: Eye Response Technologies, Inc.
- Gergle, D., Brinck, T., & Wood, S. D. (2002). *Designing Web sites that Work: Usability for the Web*. New York: Morgan Kaufmann/Academic Press.
- Goldberg, J. H., & Kotval, X. P. (1999). Computer Interface Evaluation using Eye Movements: Methods and Constructs. *International Journal of Industrial Ergonomics*, 24, 631-645.
- Goldberg, J. H., & Kotval, X. P. (1998). Eye Movement-Based Evaluation of the Computer Interface. In S. Kumar (Ed.), *Advances in Occupational Ergonomics and Safety* (pp. 529-532). IOS Press.
- Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye Tracking in Web Search Tasks: Design Implications. *Proceedings of ETRA '02* (New Orleans). New York: Association for Computing Machinery.
- Lazar, J. (2001). *User-Centered Web Development*. Boston: Jones and Bartlett.
- Norman, K. L., Shneiderman, B., Harper, B., & Slaughter, L. (1998). *Questionnaire for User Interaction Satisfaction*. College Park, MD: University of Maryland, Human-Computer Interaction Laboratory.
- Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., & Newman, J. K. (2004). The Determinants of Web Page Viewing Behavior: An Eye Tracking Study. New York: Association for Computing Machinery.
- SAS Institute, Inc. (2003). WebHound documentation. Retrieved from <http://support.sas.com/documentation/onlinedoc/webhound/> on 8/26/2005.
- UsabilityNet. (2003). "Card sorting." Retrieved from <http://www.hostserver150.com/usabilit/tools/cardsorting.htm> on 6/08/05.

Acknowledgements: The authors appreciate the long hours of data collection, data analysis, and logging of videotapes by several of our student interns during various phases of this work. Thanks to Aaron Miller (George Mason University) for his help with the card-sorting and eye tracking studies; to Michael Rosen (University of Central Florida) for his invaluable contributions to the analysis of the eye tracking data; and to Amira Abdalla (George Mason University) for leading the most recent usability testing and assisting

with the usability review. And thanks to our Census Bureau colleagues, Elizabeth Nichols, Robin Goldberg, and Sarah Brady, for their valuable comments on drafts of this paper.