

User Studies Informing E-Table Interfaces

Gary Marchionini

University of North Carolina Chapel Hill
Interaction Design Laboratory

CB #3360 Manning Hall
Chapel Hill, NC 27599
1.919.966.3611

march@ils.unc.edu

Xiangming Mu

University of North Carolina Chapel Hill
Interaction Design Laboratory

CB #3360 Manning Hall
Chapel Hill, NC 27599
1.919.962.8274

mux@ils.unc.edu

ABSTRACT

In this paper, we describe a series of user studies that were used to advance understanding of how people use electronic tables (E-Tables) and inform the development of a web-based statistical table browser for use by non-specialists. Interviews and focus groups with providers, intermediaries, and non-specialist end users; transaction log analysis; and email content analysis were used to develop a user-task taxonomy for government statistical data. These studies were the basis for a prototype web-based interface for browsing statistical tables. Two usability studies with 23 subjects and two eye-tracking studies with 21 subjects were conducted with this interface and paper, PDF, and spreadsheet interfaces. The results of the needs assessment, prototype development, and user studies provide a foundation for understanding E-Tables in general and guiding continued design of interfaces for E-Tables.

Keywords

statistical tables, table browser, needs assessment, usability testing, eye tracking

1. INTRODUCTION

User studies take a variety of forms and are used to achieve a range of goals. These include needs assessment studies to understand problem contexts and inform design, usability tests to assess specific design decisions, and studies of user behavior that use novel interfaces as stimuli. User needs assessments are often conducted to define gross functions and specifications before prototype interfaces are created (e.g., [6, 16, 19]). Needs assessments set the stage for design by circumscribing human capabilities, tasks, and behaviors related to the problem situation of interest. As design solutions evolve, usability testing is used to guide and refine progress. Usability testing has become a crucial element of the design cycle for interfaces to all manner of information technology (e.g., [13,17]), thus many software development companies have usability labs and specialists. On the theoretical side, user studies have also been used to understand human behavior and to help in building models of human-computer interaction (e.g., [1]). This paper reports results from a series of user studies of that include each of these three forms. Our overall goals were to build an empirical base for understanding the nature and use of electronic tables (E-tables), and to design and test a web-based table browser for statistical tables used by the public. The user studies took place over a four-year period as part of ongoing efforts to understand how non-specialists seek and use statistical information available in the WWW and develop a variety of user interfaces to help people do so. We argue that combining results from these different types of user study provides a rich base of understanding with both theoretical issues of statistical information use and practical elements of design. The details of the different user studies are reported individually in the literature and this paper synthesizes the results to frame a discussion of electronic statistical tables as an increasingly important medium for an informed populace and to suggest guidelines for designers.

2. THE CHALLENGE: WEB-BASED STATISTICAL TABLES FOR NON-SPECIALISTS

The hypertext and hypermedia communities developed industrial-strength systems for creating and delivering electronic texts and multimedia even before the WWW became ubiquitous. Authors and readers/users have been studied using these systems and we have substantial theoretical work on the nature and use of these media in

professional and popular culture (e.g., more than a decade of ACM Hypertext Conferences). Today, there is no such broad base of work for statistical data in electronic form. We believe that statistical data on the WWW is a distinct medium that requires specialized tools based on understanding how people think about, retrieve, analyze, and use statistical data. This work aims to examine this belief in the context of building specific interfaces to support use of statistical tables.

Government agencies are making large volumes of statistical data available to the public through the WWW so that people can make better decisions about all aspects of their lives. The most prominent example of this trend is the FedStats portal to U.S. government agency statistical data (www.fedstats.gov). FedStats is produced by a consortium of more than 70 federal agencies and provides access to thousands of statistical data sets on a diverse set of factors such as demographics, health, energy, education, science, economics, transportation, crime, and agriculture. This service and others like it around the globe (e.g., see Canada and Sweden respectively: <http://www.statcan.ca/start.html>, <http://www.scb.se/eng/index.asp>) aim to give non-specialists access to data that in the past have been the purview of privileged specialists who used them for research or policy advice, and journalists who wished to interpret them for the public.

This broadened access to statistical data carries two difficult challenges, one related to design, the other to user characteristics. First, there are few design precedents for tools and interfaces that help non-specialists access and use statistical data, especially in tabular form. A survey of the books available on producing and using statistical tables and graphs in the 20th century yields a few fairly thin volumes on guidelines for producing printed tables (e.g., [4,23]) and a number of thick volumes on the efficacy and design of graphical displays for statistical data. The emphasis in electronic displays has been on visualizations for numeric data rather than on tables of data. There seems to be an assumption that reading and using a table is intuitive. There are powerful systems for analyzing statistical data (e.g., SAS, SPSS) and general-purpose tools for managing and analyzing statistical data (e.g., spreadsheets), but these tools require a significant level of statistical literacy to use and assume that the user already has in hand statistical data sets of interest. There have been innovative interface designs for making statistical tables more usable (e.g., Rao & Card's TableLens, [18]) but most HCI effort has gone into developing graphical representations for statistical data and these techniques are being adopted by specialists (e.g., stock analysts). Additionally, the evolution of the WWW has not yet provided good models or tools for displaying highly compressed and highly structured statistical tables. The primary formats today are crude ASCII and PDF formats for paper-based displays. Alternatives are to download spreadsheet files for import into local applications. What is needed is a richer understanding of non-specialist user needs and behaviors as they use statistical tables and concerted efforts to design E-tables that are easy to retrieve, display, and understand.

The second challenge to an empirically-grounded framework for E-tables design and usage is posed by users themselves. It is well-known that the level of numerical literacy in the general US population is extremely low [14]. Although people are very reluctant to say they have difficulty reading text, many are happy to admit they have difficulty with mathematics or statistics. Statistical tables exacerbate people's insecurities because any specific value within a cell may be highly compacted—e.g., a cell may have a numeric index value (e.g., consumer price index) that itself is based on a complex set of derivations or computations. Thus, to make intelligent use of a value once it is found, people must understand what is behind a value—the variable(s) it represents, how those variables are defined, and how the data was collected. Given the large number of values within a table and the large number of tables available, the finding and interpreting demands are extremely high. Interface designs must somehow deal with both of these challenges.

3. THE TABLE BROWSER PROJECT

Over the past several years a team of information scientists has been working with federal statistical agencies to study how non-specialists access and use statistical data [5, 9]). Building upon this work and as part of the NSF Digital Government Program, we focused attention on statistical tables with the intention of designing and testing

interface prototypes aimed at non-specialists using the WWW. A number of user studies using different techniques were conducted before and during the prototyping and implementation phases. These studies are summarized here with the goal of presenting a baseline for understanding E-table usage and guiding interface designs.

3.1. User Needs and Behavior Assessments

How do people think about and use statistics? To assess user needs, we used interviews, focus groups, transaction logs, and email content analysis. Professionals responsible for gathering and disseminating data in agencies, intermediaries responsible for assisting people with finding and using statistical data (e.g., librarians, state-level liaisons for federal agencies), and end users (e.g., students and researchers at universities, senior citizens) were interviewed or participated in focus groups with the intention of understanding what types of questions are generated for statistical agencies, how people go about finding and understanding data, and what types of tools and services might be helpful to them. Likewise, transaction logs for statistical websites were studied to determine patterns of access and types of queries and terminology used. Additionally, email requests to agencies were analyzed to understand the types of problems people had with statistical concepts and existing delivery systems. These studies are detailed elsewhere [5, 9], but several themes emerged that served as a basis for specifying the functionality of the Table Browser prototype. First, there is enormous variability in the needs people bring to these services. A task taxonomy was developed that includes pragmatic, semantic, and syntactic categories. People seek and use statistics as part of larger goals such as decision making, reporting, and persuading. Specific tasks range from value lookups to identify or verify, to simple comparisons, to trend analyses, to complex computational analyses. In most cases, finding out how to complete these tasks is as crucial for non-specialists as the execution of the task.

Second, we found that vocabulary is problematic for end-users—agency specific or technical terms were not useful (e.g., people tend to think about general terms like income or inflation without recognizing the many variations or interpretations these concepts imply for specialists). For example, income may only include earned income at an agency oriented toward labor (e.g., BLS) but include capital gains at agencies oriented toward business (e.g., Commerce, Federal Reserve). Concepts like average income for one's state have very different values in tables from these different agencies. Another finding is that people trying to make personal decisions tend to think locally about statistics rather than nationally or internationally—they want data about their local community and this is typically not available from national agencies. Thus, in addition to stating what data is available in a system, there is a need for data providers to clarify what is not available.

Third, people had difficulties with the systems used to find and use statistics. Transaction logs illustrated that large proportions of users abandon statistical websites after one or a few requests—clearly not enough to even get to tabular values. People reported dissatisfaction with WWW tables, finding them awkward to use, especially with respect to screen size, scrolling, and lack of explanations. People also had difficulty finding data in government websites due to the volume of data available and the large number of variables involved.

Overall, these studies illustrate that non-specialists have poor understanding of statistical concepts generally, but sophisticated understanding of particular concepts of immediate interest. For example, although senior citizens showed little interest or understanding about statistical concepts for data tendencies (e.g., mean, median) or distributions (e.g., variance), they showed sophisticated understanding of the adequacy of death certificates as sources for morbidity data. These differences in understanding concepts specific to a problem at hand and the general concepts of statistical reasoning are exacerbated by the disconnects between the terminology used by agencies and by people seeking data noted above. This disconnect parallels the well-known ambiguity problem in text-based searching in the WWW, but may be more crucial because tables tend to have far fewer words, even taking into account the notes and reports in which they are often embedded. This point also suggests one of the

reasons E-tables represent a distinct medium in the WWW. The user needs and behavior assessments and characteristics of e-tables gleaned from them helped to inform an e-table interface that aimed to assist non-specialists in understanding and using statistical tables.

3.2 The Table Browser Interface

Based on these studies, we developed an overall architecture for statistical data on the WWW (see [11] for the overall scheme) and developed prototype interfaces for different modules. A dynamic query data explorer was created to allow users to ‘slice and dice’ large data sets into meaningful partitions. A query module was also defined (see [10] for examples). A table browser application was created to display tables located with these browse and search tools and to provide basic manipulation and data elaboration support. Static displays such as HTML and PDF minimize interaction loads by not providing manipulation functions and using standard display options. Downloadable spreadsheet files give users manipulation capabilities but require substantial user knowledge and interaction loads. Based upon our user needs assessments, we aimed to design a mid-level solution that would allow people to have crucial display control and some minimal manipulation control without learning an entire application suite. The Table Browser (TB) prototype is shown in Figure 1. The prototype is implemented as a Java applet that reads XML files that contain the table data as well as substantial amounts of metadata that support understanding and use. The applet source code is available as an open source package from www.ils.unc.edu/idl/etables/source.

The main goal was to provide some basic functionality for interacting with a small set of retrieved tables while minimizing the perceptual and cognitive effort of users. The design goals were to develop a tool that had at least the following characteristics:

- web-based delivery
- “sticky” headings that do not scroll with the data
- highlight capabilities for rows, columns, and/or cells
- data reorganization capabilities (e.g., delete, drag columns, create subtables, zoom)
- metadata notes to aid understanding

Although these seemed like a reasonable set of specifications, realizing them required technical and conceptual innovations. The biggest challenges were related to providing notes that are specific to the tabular entity. To address this problem, we adopted the Data Documentation Initiative metadata model (www.icpsr.umich.edu/DDI) and created an XML Document Type Definition (DTD). The DTD defined small sets of metadata elements for each of the three table entities: table, row/column, cell. For the table, an explanation of the purpose of the table, study/survey description, data file(s) description, and links to agency email addresses were encoded in the XML files. For the row/column elements, footnotes, expanded descriptions of variables, units of measure, and source of data were included; and for the cell elements, cell-level footnotes and verbose variable names were encoded. We then mapped these metadata elements onto different user-control mechanisms. The primary notes were put into a scrollable and resizable metadata notes pane (lower left pane in Figure 1). Verbose variable names, units, and cell level footnotes were provided as tooltips when users hovered the mouse over a row/column heading or cell value. The prototype provides a tool bar with basic functions for choosing focus (table, row, column, or cell), deleting

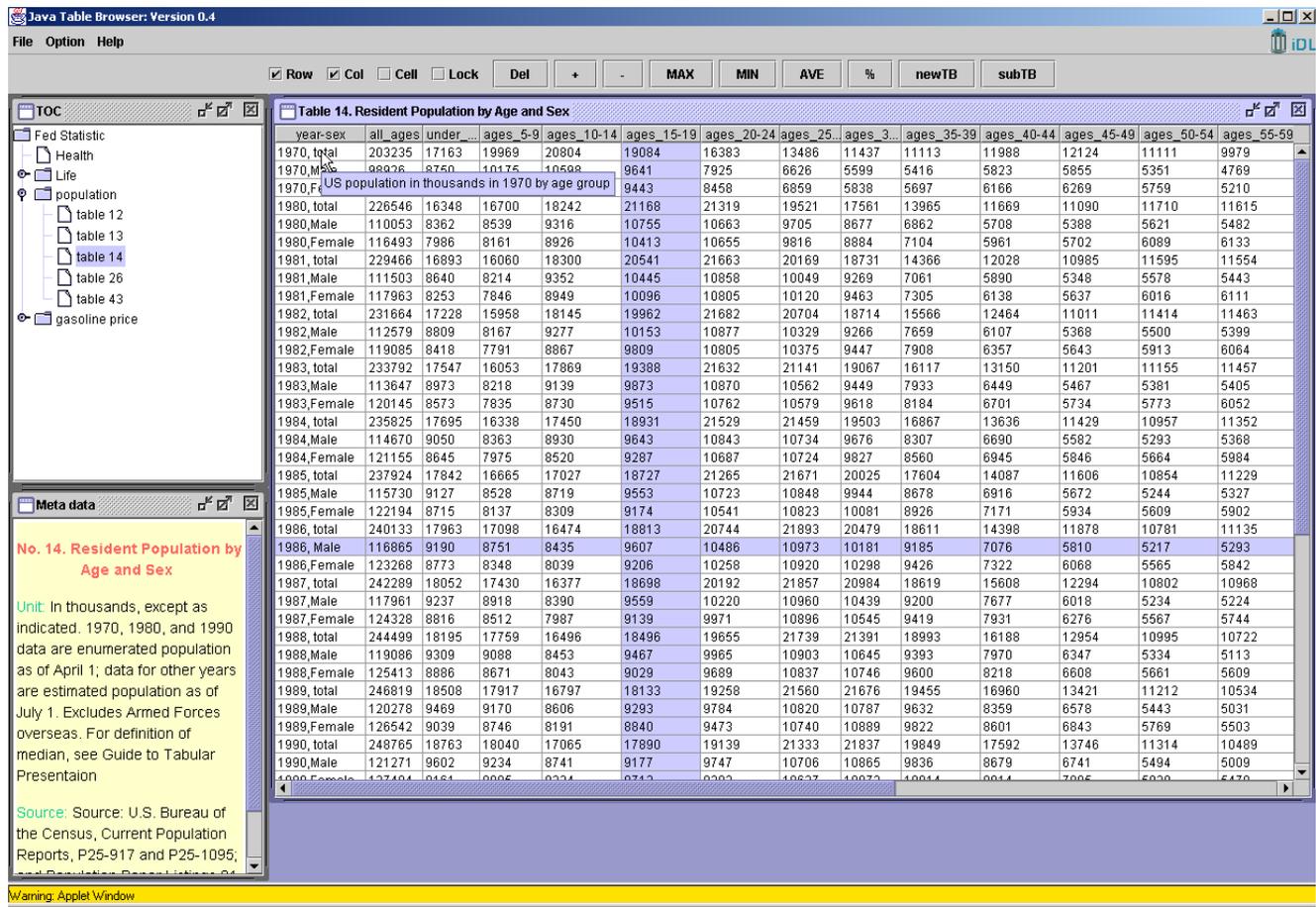


Figure 1. The Table Browser Interface.

selections, locking row headings (column headings are always ‘sticky’ but we gave users the option with rows to maximize horizontal screen usage for the data), a primitive zoom function, simple compute features (max, min, mean values for selected columns, and change values into percentages), and two ways to define a subtable (use the current selected region or define a subtable by selecting variable names from a pick list). The zoom function is admittedly clumsy in the present implementation.

Because the data and metadata must be on the client side and some tables are quite large, the architecture was designed to send the table-level metadata and small batches of data and metadata in separate Java threads (see [15] for details on architecture and implementation). In Figure 1, the user has highlighted a single row and column and is hovering the mouse over a cell that pops up a brief description of what the value in that cell represents. The upper left pane shows a hierarchical list of available tables. Single clicking any of these table names displays the metadata overview for that table in the lower left notes pane. Double clicking displays the table in the main pane.

3.3. Laboratory User Studies with the TB

The prototype was developed and tested over an 18 month period. Four user studies were conducted during the development of the TB. Two of these studies were usability tests to determine the efficacy of specific features of the tool with respect to other interfaces (PDF and spreadsheet) and inform subsequent iterations. Two other studies used eye-tracking to examine low-level behaviors with the TB, PDF and paper formats. It is important to note that these studies themselves reflect the evolution of both the TB interface and our understanding of E-Tables used by non-specialists. In conducting the studies we tried to control the study protocols as much as possible so

that qualitative comparisons could be made across studies, however, no statistical hypothesis testing was planned either within specific studies or across the series of studies. In the individual summaries and synthesized discussion, both qualitative and quantitative data are presented, but no statistical inferences are attempted.

All four studies used a common set of three categories of tasks that were designed to examine specific functionalities. These three categories correspond to Wainer's [24] recommendations for a triad of numeracy measurements. The first category included tasks that asked participants to locate specific values in cells or answers in notes (e.g., What was the population of Utah in 1994?). The second category included tasks that asked participants to do a computation or compare values (e.g., In 1996, what was the difference in life expectancy between 30 year old white males and 30 year old white females?). The third category asked participants to determine trends (e.g., Across all groups, what happens to the difference in life expectancy between males and females as people age?). In all these studies, paid undergraduate and graduate students with substantial WWW experience but a wide range of statistical experience participated. Tables used were related to demographics, longevity, and gasoline prices and came from the Bureau of the Census, National Center for Health Statistics, and Energy Information Administration.

The usability studies were conducted in the Interaction Design Laboratory (IDL) where participants worked one at a time at a workstation with overhead and side video cameras, a scan converter, and a microphone. The three video signals and audio track were integrated onto one videotape with three windows for the different views. The experimental conditions were presented via a WWW interface and all answers were entered via WWW forms that sent the responses to a database.

The eye-tracking studies were also conducted in the IDL using an ASL 504 eye tracker and Flock of Birds head tracker that yield 60 fixations per second (see <http://ils.unc.edu/idl/details/eyetracking/eyetrack.MPG> for a movie demonstration of the eyetracking process). It is important to note as Jacob and his colleagues [8] have pointed out that eye-tracking data provides excellent low-level traces of human behavior but does not stand alone in explaining how or why people use interfaces. Our approach is to combine these data with coarser data from usability tests and other user-data such as interviews.

In all these studies, a demographic questionnaire was administered before the study that included a series of questions about experiences with computers and statistics. Subject recruitment specified WWW experience as a prerequisite so all subjects has substantial computing experience. The statistical questions asked about formal courses, statistical packages used, types of common tables used, and use of FedStats and other web sites with statistical information. Responses to the statistical questions were used to categorize subjects into either low or high statistical literacy groups. In addition to the demographic questionnaire, a brief introduction to the different interfaces was provided before subjects began the main experimental tasks. A post-study questionnaire with five satisfaction items and a debriefing questionnaire were also administered.

3.3.1. Usability Study 1: TB and PDF

This study was conducted in the Fall of 2000 and used the first functional prototype of the Table Browser as well as PDF versions of the same tables (see [2] for details). Important parts of this study were the construction of a pool of tasks in the three categories of complexity, development of the instrumentation and protocols. Twelve participants used both versions of the tables for comparable tasks in a counterbalanced, within-subjects design. Each subject performed two tasks at each of the three levels of task complexity for each of the two interfaces.

There were no consistent differences in speed or number of errors across the two formats, but there were clear differences across the different tasks with time and errors increasing as the tasks became more complex. Subjects expressed enthusiasm for the TB in that eight of twelve expressed preference for the TB in a verbal debriefing and the mean scores of satisfactions for all five satisfaction items were higher for the TB. Half the subjects noted the efficacy of fixed column headings and the ability to highlight cells, rows, or columns, and eleven of twelve participants used the TB to drag columns around in the comparison tasks. Participants had difficulty in both interfaces on a number of fronts: finding the right table to use from a small set of alternatives, using right scrolling, interpreting units such as values in thousands when giving answers, and spending considerable time looking for contextual cues to understand the data. Participants did use the metadata window for context in the TB. This study illustrated that some introduction to the TB was necessary for even these fairly computer savvy subjects. Based upon this study, a number of improvements were made to the TB, for example, the delete function was implemented, headings were made easier to read, and clearer instructions and feedback were provided.

3.3.2. Usability Study 2: TB and Spreadsheet

This study was conducted in the summer of 2001 using the revised version of the TB. One addition in this version of the TB was brief footnote information (e.g., units of measure) available as users hovered the mouse in row and column headings. Eleven subjects used both the TB and Excel spreadsheet versions of the tables in a within-subjects design with counterbalancing for order. Subjects were asked to complete the same tasks as in the other studies except additional tasks were added that required use of footnotes so we could focus on the effects of mouseover functions in the TB. Overall, subjects completed seven tasks with the TB and seven with the spreadsheet. As in the other user study, all sessions were videotaped and pre and post session questionnaires were administered.

There were no differences in total time taken to complete the tasks using the TB or the spreadsheet. However, there were differences for some of the tasks, with three tasks showing faster performance with the TB and three showing faster performance with the spreadsheet. The TB yielded much slower results on one of the footnote tasks (360 vs. 145 seconds) and the spreadsheet yielded much slower results on the other footnote task (340 vs. 131 seconds). As in the other studies, more complex tasks took longer in both interfaces. In this study, subjects were asked to compare the footnote mouseover function to the footnotes in the spreadsheet using four scales. Average ratings for the mouseover presentation of footnotes was higher for all four of the nine-point scales (7.0 vs. 6.0, 7.0 vs. 5.5, 7.0 vs. 5.9, and 6.8 vs. 5.0 on helpfulness, satisfaction, ease, and flexibility scales respectively). Satisfaction ratings for the overall interfaces were also found to strongly favor the TB (7.5 vs. 5.0, 7.4 vs. 4.5, 7.4 vs 5.2, and 7.4 vs. 4.5 respectively). In the verbal debriefing, subjects noted the “sticky” column headings and rich metadata (including footnotes) as advantages of the TB but suggested adding cut and paste and sticky row headings as features.

The two usability studies provided venues for testing specific features of the TB interface as well as its overall effectiveness. The usability studies led to several improvements in the respective next iterations of the TB. The results clearly demonstrate the effectiveness of specific features such as sticky headings, highlighting rows and columns, dragging columns, and the feasibility of multiple levels of notes to aid understanding. People did require time to become familiar with the TB and thus its role as a middle ground solution between static displays and full-functioning spreadsheet or statistical package applications was demonstrated.

3.3.3. Eye-Tracking Study 1: Paper and PDF

To assess the low-level indicators of behavior while using e-tables, we used conducted eye-tracking studies of people using different forms of tables. The first eye-tracking study was conducted in the Fall of 2000 and used PDF and print on paper versions of the tables (see [25] for details). Eight participants used the PDF tables and 2 subjects used the paper versions. Since we were not interested in controlled comparisons of the two media but

what types of behaviors arose in E-tables, we limited the number of subjects in the paper condition so as to get a general baseline for examining how people used the E-Tables. In the eye tracking studies, dependent variables included time to completion, number of eye fixations, and fixation duration. The basic procedure was to have subjects complete the consent forms and demographic questionnaire, perform the eye calibration (subjects are prompted to gaze at nine points in sequence), the task is verbally stated for the subject with opportunities for clarification/repetition, the subject uses the table to answer the question verbally, a new table is loaded by the investigator and the next question is posed (repeated until all eight tasks are completed), subjects complete the post-study questionnaire and debriefing. For the two subjects who used the paper condition, the PDF tables were printed (in larger fonts to account for distance) and taped to the monitor. The entire procedure took about one hour depending on subject performance with the tasks.

As in the usability test, there was a strong relationship between task complexity and user load—in this case, number of eye fixations increased with task complexity. Subjects with high statistical table experience also showed faster task completion than novices and these differences increased with task complexity. Novices required 24%, 64% and 135% more time respectively for the three increasingly difficult tasks. The mean durations of fixations were not different across the three tasks, perhaps this is a stable measure of perceptual performance and a less discriminating measure of cognitive performance than number of fixations. Two of the PDF subjects and the two paper subjects were characterized as statistical experts based on the pre-trial questionnaire. When these two sets of users' performance was compared, the paper users showed higher mean numbers of fixations for five of the eight tasks, but shorter mean fixation durations for seven of the eight tasks. Although the number of subjects is too small for any firm conclusions, this result suggests that for experienced users of statistical tables, scanning of paper may be more perceptually efficient/comfortable than scanning PDF tables.

Plots of eye fixations show strong influence of the tasks, e.g., the comparison tasks showed many horizontal scans. Figures 2, 3, and 4 show typical eye movement patterns for each of the three types of task with the PDF tables without scrolling. In the figures, the fixations are represented by blue dots and numbers and the eye movements by red lines. Figure 2 shows a simple lookup task where the subject quickly locates the appropriate row and column, double checking the column headings as visual search progresses. Figure 3 shows a typical pattern for a comparison task (Compared to the population information from 1997, which two states have decreased in population size since 1970?) that required systematic search of all rows. Figure 4 shows a plot for a trend analysis task (Across all groups, what happens to the difference in life expectancy between males and females as people age?). In both conditions, there was substantial vertical and horizontal eye-travel to check on headings, especially column headings. One of the main results of this study was strong confirmation for the importance of sticky column/row headings as an enormous proportion of eye movement (between 40% and 50%) was devoted to heading scanning. This result also demonstrates one of the important distinctions of E-tables in contrast to E-text—human processing of statistical tables is much more highly discrete than the typical continuous linear (left to right) processing of text and providing good context anchors that are easily accessible is an important design goal. Use of the scroll bars was particularly evident for the PDF tables. Figure 5 shows a plot for a simple lookup that required scrolling (In 1988 what was the resident population of the 15-19 years age group?). Both the novice and the expert users preferred using paper for large tables since they often found themselves scrolling and losing headings or positioning. This was true for footnotes as well as for cell-level values.

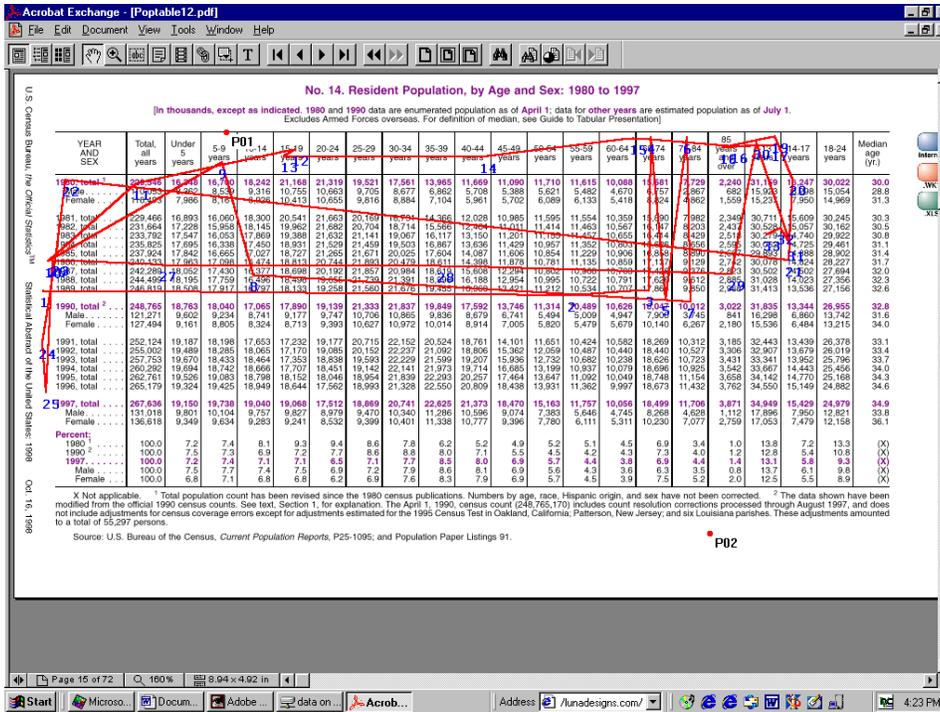


Figure 2. Eye Movement Plot for a Simple Lookup Task.

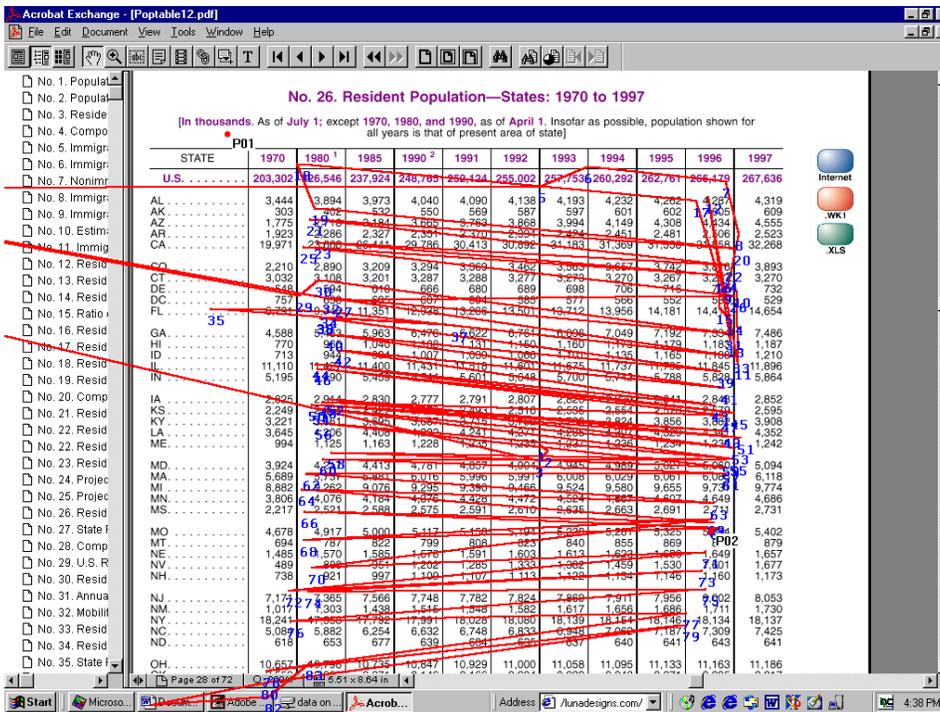


Figure 3. Eye Movement Plot for a Comparison Task

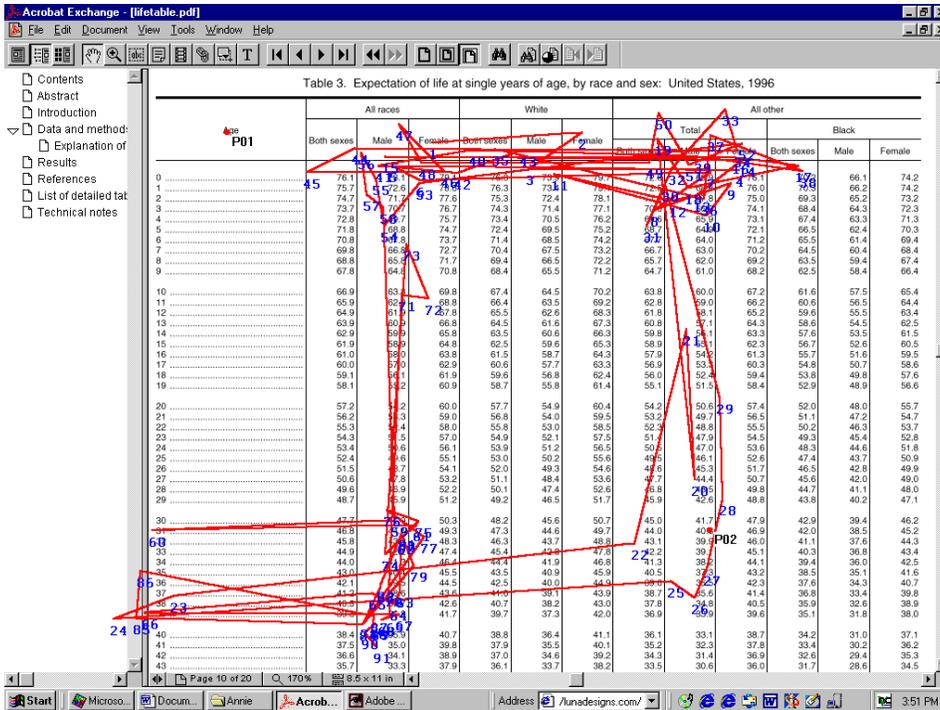


Figure 4. Eye Movement Plot for a Trend Analysis Task

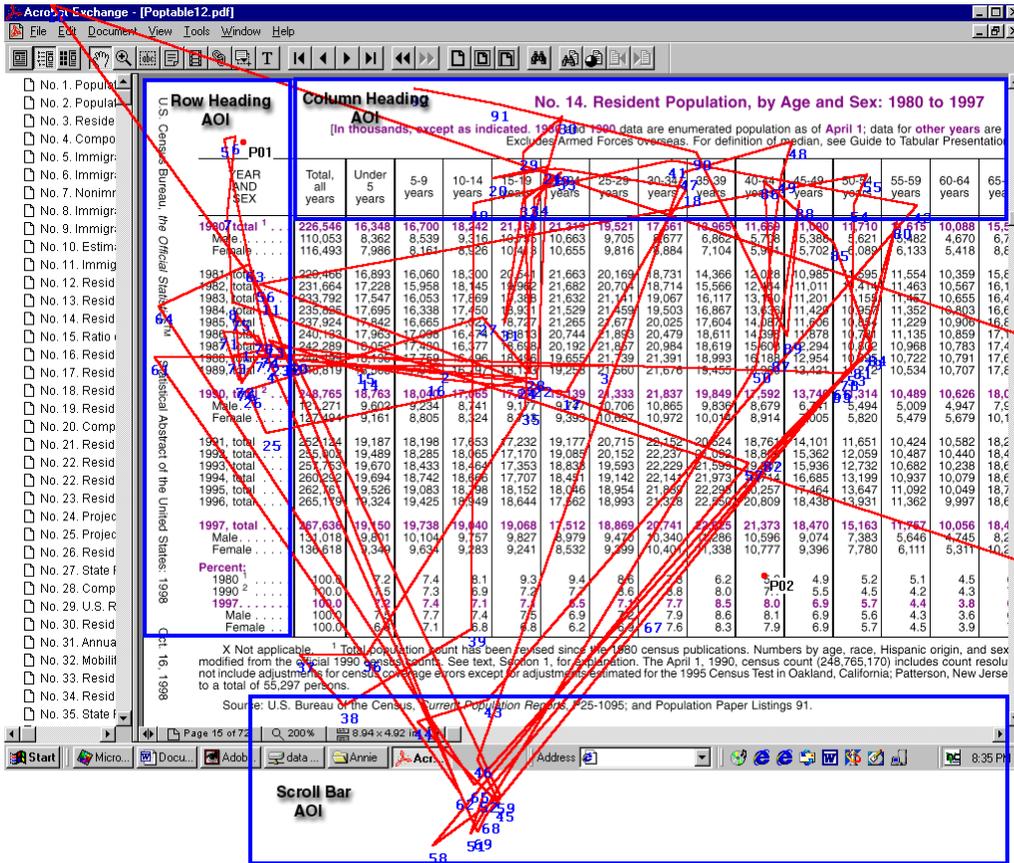


Figure 5. Eye Movement Plot for a Simple Lookup Task that required Right Scrolling

3.3.4. Eye-Tracking Study 2: TB and PDF

This study was conducted in the spring of 2001 using the revised version of the TB and PDF versions of the tables (see [3] for details). Eleven subjects participated, four used the PDF tables and seven used the Table Browser. This sample of participants showed good statistical skills with only two in each group classified as non-experts with statistical tables. Overall, TB users took less time and had fewer fixations to complete the tasks. For the PDF files, non-experts took more mean time and had more fixations than experts. In the TB condition, however, experts spent more time than non-experts but had fewer fixations. Although there were too few subjects to rigorously affirm any differences, these data suggest that the TB may assist non-specialists in overcoming some deficiencies due to statistical experience. As in the other studies, task complexity affected mean performance, with large differences in effort for the complex tasks. Although there were no time or eye-travel differences for the first category of tasks (simple lookups), there were consistent differences across formats in the other, more complex, tasks. As in the other studies, subjects reported consistently higher satisfaction ratings for the TB on all five questions.

A comparison of two fixation plots for the two interfaces used for the same task gives a graphical representation for how distinct user behavior can be and how the interface can influence user behavior. Figures 6 and 7 show the PDF fixation plot on the left and the TB fixation plot on the right for two individuals executing the same tasks. Figure 6 demonstrates the advantage of sticky column headings in the TB interface. The task asked users to compute population increase over a twenty year period and required vertical scrolling in both conditions. The sticky column headings facilitated less orienting eye travel to the row and column headings. Figure 7 shows plots for a task that entailed comparing two remote columns of data. The TB user spent several eye movements in locating and dragging the rightmost relevant column next to the leftmost column, whereas the PDF user had to move back and forth repeatedly to complete the task.

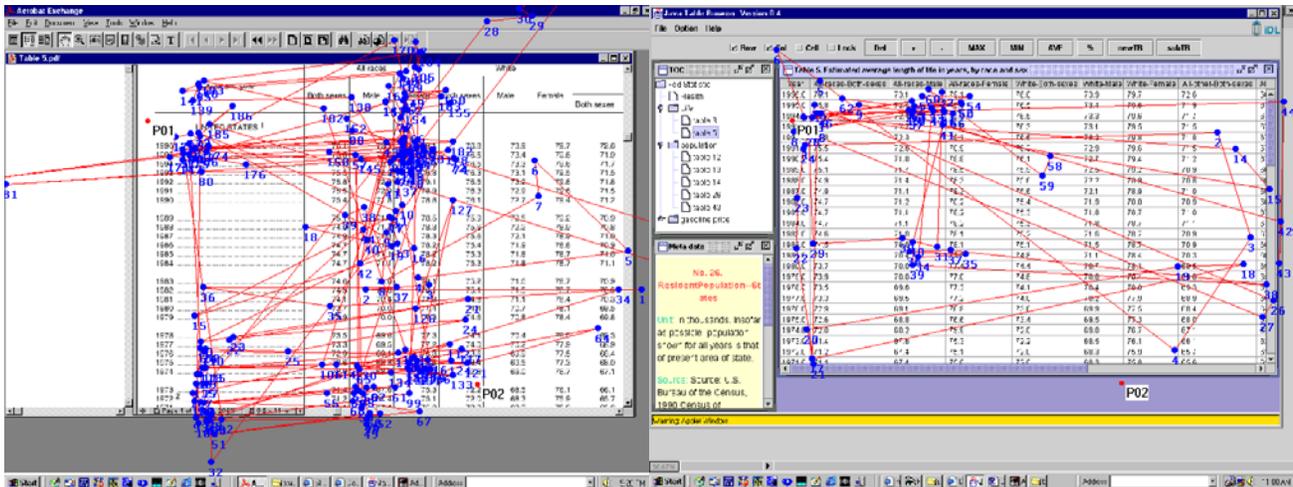


Figure 6. Eye Movement Plots for the Same Trend Analysis Task: PDF on Left, TB on Right

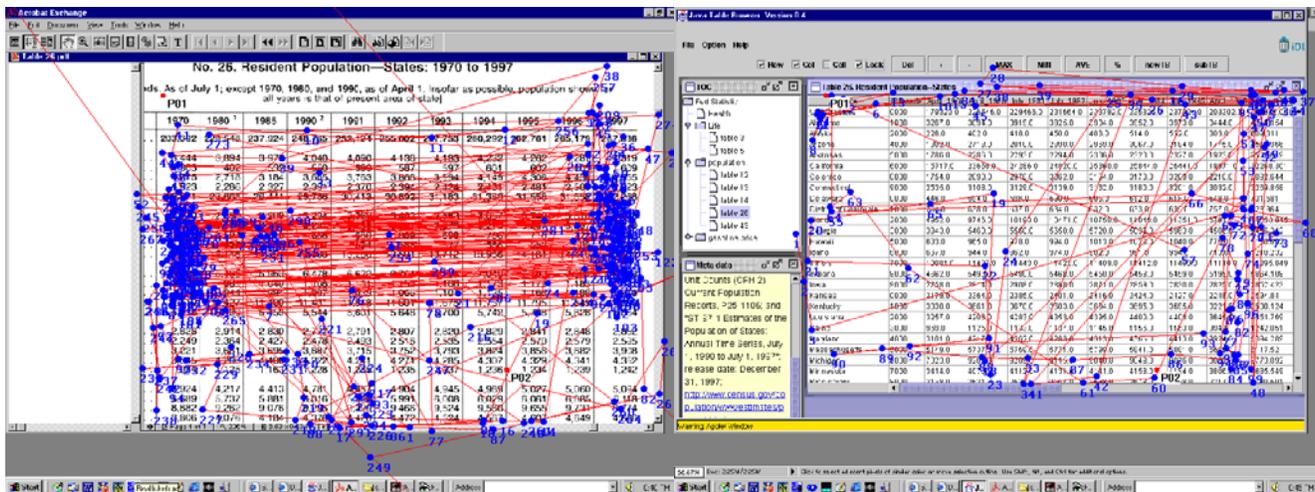


Figure 7. Eye Movement Plots for the Same Comparison Task: PDF on Left, TB on Right

The eye tracking studies provide low-level behavioral data that reinforce well-known human factor principles such as task complexity and user task expertise strongly influencing performance with statistical tables. The studies demonstrate unequivocally the crucial roles that headings play in statistical tables regardless of medium. Most importantly, the studies show the powerful advantages provided by simple interactive features such as highlighting columns or other objects of focus, and repositioning columns. These advantages of interaction and user control must be balanced with the costs of learning to apply this control, however, the eye movement data suggest that substantial can accrue with simple interactive mechanisms and that non-specialists can especially benefit from these mechanisms that make reading and interpreting data more physically direct.

4. TOWARD THEORY AND DESIGN PRINCIPLES

These studies stand individually as interesting snapshots of human behavior with statistical data in today’s practical user interfaces. Taken together, however, they begin to suggest that electronic statistical data is indeed a distinct medium that will require specialized functionalities and interfaces. The particular tool used as a foil here, the TB is but one example of how the design community might begin to provide such tools. One way to think about the overall user experience is to consider what these studies tell us about the nature of statistical content, what people bring to these data, and the types of tasks people aim to accomplish. In many ways, the traditional medium of paper-based statistical tables is so difficult for non-specialists to use that we should be able to add value through the electronic medium.

4.1. Statistical Data Characteristics.

Numbers are not words and statistical tables are not text. Numbers do not stand alone but represent counts or other qualities associated with real world or conceptual objects, states, and processes. Understanding a numeric value is dependent on knowing both sides of these associations--the value and the units of measure or objects counted. Statistics are numbers that are often more abstract than simple numerical counts—e.g., representing distributional characteristics such as frequencies, central tendencies, or variations. When presented in tabular form, numbers and statistics are separated from their associated elements to conserve display space. This alone makes values in tables highly compressed/encoded, and thus cognitively demanding. This also makes searching for data difficult as the words that are typically used as anchors may not be included in the file or may be abbreviated to save space.

Secondly, tables themselves have non-linear structure. Tables have highly partitioned components, each of which may be highly compressed and abstracted. The simplest tables cross two sets of variables and have at least three hierarchical levels: the full table itself, columns and rows, and cells, plus non hierarchical content in the form of various notes. More complex tables offer multiple sets of variables. This structure has implications for human processing and for design. Unlike text, where a word might be decoded using spatially proximate context, the context for a cell in a table is distributed in headings, footnotes, table titles, or accompanying text that is not immediately proximate. People, especially those who are not specialists in the data and tables at hand, must scroll or manipulate windows or turn paper pages, and use gross eye movements to contextualize values. A related effect is the role that metadata play in E-tables. Rather than unwieldy footnotes and textual interpretations in separate pages or documents, the electronic format provides opportunities for more closely coupling contextualizing metadata to the actual data values through interactive control mechanisms (e.g., additional windows and mouseover events).

A corollary feature of tables is that they can be quite large, not fitting on a page or screen, thus requiring innovative techniques to provide the context that users need. In the case of the TB, a set of windows and control mechanisms were customized in the Java applet. Furthermore, transferring large statistical tables in the WWW environment requires novel techniques for packaging the data. In the TB, a multi-threaded approach was used to deliver the interface and chunks of data in separate threads so that users could begin working with a base set of values without having to wait for the entire data set to arrive.

These data structure characteristics surely call for novel techniques for helping people contextualize values in E-tables. The approach we used before people actually get to a table is to provide exploratory, dynamic query interfaces that provide overviews of the data [20, 22]. Once in a table, our approach was to maintain column and row headings and provide a metadata notes pane within the larger table window. Tooltip techniques were also used to provide quick, on-demand notes for various elements. These techniques proved to be effective and were appreciated by participants in our user studies.

4.2. Users

Our user needs assessments showed that the availability of government statistical data on the WWW is rapidly augmenting the traditional set of sophisticated users of government statistics (statisticians, economists, academic researchers, graduate students, journalists, advisory staff for government and business leaders). Non-specialist users from all walks of life are accessing government websites for a wide range of purposes, e.g., seniors investigating health care or economic issues, baby boomers planning for retirement, young adults investigating new jobs or places to move, and K-12 students doing homework. Although these non-specialists typically have low levels of statistical competency, they are often highly sophisticated about their specific needs and are technology savvy (and thus have high expectations about service). They are concerned with finding answers rather than numbers and have difficulty expressing needs in the language used in data sets and difficulty mapping their needs onto the appropriate government agency, survey, and statistical summary. The language issue is exacerbated because statistical tables in particular have fewer word anchors for retrieval and interpretation purposes.

Our approach was to use two techniques for adding explanations once a table had been displayed—the metadata notes pane and tooltips for elaborations for variable names, units, and other heading or cell values. One issue that arose in the metadata notes pane was some confusion about whether a note applied to a cell or column/row or to the whole table. This was of course obvious with the tooltips, but lengthy tooltips for every cell would certainly not be an effective solution. Participants in the studies liked the metadata notes pane and subsequent interfaces might consider highlighting a cell or heading as users hover over the notes, although some more direct labeling

technique might be more appropriate. In our prototype, we added a considerable amount of metadata to each table. Hert and her students analyzed hundreds of questions people posed in the interviews and categorized them so that they could be mapped to the DTD categories. Although we were able to automatically generate XML markup for a few physical values from the raw data tables (e.g., headings), a grand challenge in information management is finding ways to cost-effectively add more subtle and informative tags to data elements.

4.3. Tasks

Non-specialists bring a variety of tasks to statistical websites. Our user studies demonstrate that the task strongly influences performance. Simple lookups or verifications are clearly less complex than trend analysis or comparing sets of values. Because tasks are always embedded in larger purposes, discussions of task must specify granularity. Specific tasks are supported by specific tactics and tools that themselves require subtask actions. From a GOMS-like perspective [1], a good set of low-level interface features can promote progress and fend off frustration for non-specialists focused on higher level goals. That is, if e-tables facilitate efficiencies at the lowest levels of behavior, these efficiencies should build toward higher overall performance. Our studies suggest that specific interface techniques can provide large benefits to non-specialists, especially with more complex tasks. Obvious big wins in the TB interface were the sticky columns, highlighting cell/row/column elements, draggable columns, and easy access to notes. Although spreadsheet applications offer some of these features, they require users to make conscious decisions to activate or change modes. In the more complex tasks with the TB, participants were able to leverage these low-level features to aid performance. As we continue to gain experience with diverse sets of tasks that non-specialists bring to statistical websites, additional interface features will emerge if we are diligent in observing behavior and talking to users.

5. Conclusion

This paper demonstrates how user studies inform both theory and practice. Each of the studies adds to our understanding of how people think about and use statistical tables. They also inform the ongoing iterative design at both the specific feature level as well as the global look and feel level. Our results show that simple techniques for managing display can provide substantial satisfaction and selected behavioral benefits to non-specialists. This examination of e-tables is, however, highly preliminary. Perhaps the electronic format does not matter or is totally swamped by the task complexity effects. The data suggest that the expertise gap may close a bit with better electronic tools. Considerable work is needed to find ways to add additional value to aid understanding of tabular data once it is found.

Fred Brooks noted in his ACM Allen Newell Award acceptance lecture at SIGGRAPH '94 that "the scientist builds in order to study" and it is this spirit that drives this work as we built prototypes and methodologies that themselves are useful but serve to advance our understanding of E-Tables as a distinct phenomenon. Beyond the development and testing of a specific tool, we claim that by using multiple data collection techniques to gather different sources of evidence, we can aggregate rich views of data-task-user phenomena that define the user experience. We claim that E-Tables are a distinct information medium that is becoming increasingly important as more data is available in the WWW and that requires specific user interface tools and techniques. Some of these tools and techniques are included in the Table Browser and it serves as a test environment for learning more about E-Tables and how people can more easily interact with them.

ACKNOWLEDGMENTS

This work was supported by NSF Grant #9876640 that included Gary Marchionini, Carol Hert, Liz Liddy, and Ben Shneiderman as co-principal investigators. The user studies were conducted by Laura Chessman, Airong Xu, Tao Gao, and Xiangming Mu at UNC-CH. Zhen Zhen Deng devised the DTD and added the XML markup to tables used in the studies. Alex Mueller assisted in Eye Tracking study 2 and Deen Freelon assisted in usability study 2. Fred Conrad made useful suggestions in an earlier draft.

REFERENCES

1. Card, S., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum.
2. Chessman, L. *Table Browser Usability Study*. Master's paper, School of Information and Library Science, UNC-CH, November, 2000. <http://ils.unc.edu/idl/details/Chessman.pdf>
3. Gao, Tao. (2001). *Processing Numeric Table Data in Electronic Formats: Eye Movement Patterns and Searching Efficiency*. Master's Paper, School of Information and Library Science, UNC-CH, July, 2001. <http://ils.unc.edu/idl/TaoGao.pdf>
4. Hall, R. (1943). *Handbook of tabular presentation: How to design and edit statistical tables, a style manual and case book*. NY: The Ronald Press Co.
5. Hert, C. & Marchionini, G. (1998). Information seeking behavior on statistical websites: Theoretical and design implications. *Proceedings of the American Society for Information Science Annual Meeting* (Pittsburgh, PA, oct. 25-29, 1998). 303-314. See also <http://ils.unc.edu/~march/blsreport/mainbls.html>
6. Hix, D. & Hartson, H. (1993). *Developing user interfaces: Ensuring usability through produce & process*. NY: John Wiley & Sons.
7. Hyland, P. & Gould, T. (1998). External statistical data: Understanding users and improving access. *International Journal of Human-Computer Interaction*, 10(1), 71-83.
8. Jacob R.J.K. The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems*, Vol. 9, No. 3, April 1991, 152-169.
9. Marchionini, G. (1997). Advanced Interface Designs for the BLS Website. Final Report to the Bureau of Labor Statistics, July 1997. Available at http://ils.unc.edu/~march/blsreport98/final_report.html
10. Marchionini, G., Hert, C., Liddy, L., & Shneiderman, B. (2000) Extending Understanding of Federal Statistics in Tables. *ACM Conference on Universal Usability*. (Washington), 132-138.
11. Marchionini, G., Hert, C., Shneiderman, B., & Liddy, L. (2001). E-tables: Non-specialist use and understanding of statistical data. *Proceedings of dg.o2001: National Conference for Digital Government*. (Los Angeles, May 21-23, 2001). 114-119.
12. Marshall, C. (1998). Toward an ecology of hypertext annotation. *Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia*. (Pittsburgh, June 20-24, 1998). 40-49.
13. Mayhew, D. (1999). *The usability engineering lifecycle*. San Francisco, Morgan Kaufmann.
14. Moore, D.S. (1997). New pedagogy and new content: The Case of statistics. *International Statistical Review*. 65 (2):123-165.
15. Mu, X. (2000). Java Table Browser: Transportation and Presentation of Large Statistical Tables Over Networks. Master's paper, School of Information and Library Science, UNC-CH, May, 2000. <http://ils.unc.edu/%7Emux/homepage/links/masterpaper/XiangmingMu.pdf>
16. Plaisant, C., Marchionini, G., Bruns, T., Komlodi, A., & Campbell, L. (1997). Bringing treasures to the surface: Iterative design for the Library of Congress National Digital Library Program. *Proceedings of ACM CHI '97 Conference*. (Atlanta, March 22-27, 1997), p. 518-525.
17. Nielsen, J. (1993). *Usability engineering*. Boston: AP Professional: Academic Press.
18. Rao, R. & Card, S. (1994). The table lens: Merging graphical and symbolic representations in an interactive focus+context visualization for tabular information. *Proceedings of ACM CHI 94* (Boston, MA) NY: ACM Press. 318-322.
19. Rose, A., Ding, W., Marchionini, G., Beale, J., & Nolet, V. (1998). Building an Electronic Learning Community: From Design to Implementation. *Proceedings of ACM CHI '98* (Los Angeles, CA, April 18-23, 1998). 203-210.

20. Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human---computer interaction* (3rd ed.). Reading, MA: Addison-Wesley.
21. Sibert L.E. and Jacob R.J.K. Evaluation of eye gaze interaction. *Proceedings of the CHI 2000 conference on Human factors in computing systems*, 2000, 281-288.
22. Tanin, E., Plaisant, C., and Shneiderman, B., Browsing Large Online Data with Query Previews, *Proceedings of the Symposium on New Paradigms in Information Visualization and Manipulation - CIKM*, 2000, USA
23. Walker, H. & Durost, W. (1936). *Statistical tables: Their structure and use*. NY: Bureau of Publications Teachers College, Columbia University.
24. Wainer, H. (1997). *Visual revelations: Graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot*. NY: Copernicus Books.
25. Xu, A. Eye Tracking Study on Identifying and Analyzing User Behavior - Eye Movements, Eye Fixation Duration and Patterns - When Processing Numeric Table Data in Paper or PDF Format. Master paper, School of Information and Library Science, UNC-CH, November, 2000. <http://ils.unc.edu/idl/details/AirongXu.pdf>