Data exploration interfaces
Meaningful web database mining by non-statisticians

John J. Bosley
Bureau of Labor Statistics
Bosley_J@bls.gov

Kathleen A. Straub
Human Factors International
kath@humanfactors.com

Introduction

The advent of the World Wide Web -has enabled government and private statistical agencies to make vast stores of data available to anyone with internet access[1]. Historically, individuals accessing these resources were primarily academic, government or other statistically proficient professional data users. However, today any member of the public with a browser-equipped computer connected to the Internet has direct, interactive access to several large-scale databases. This "democratization" of access to data has the potential to implicitly elevate the general statistical sophistication of the public at large. To the extent that this potential is fulfilled, the increased understanding of how society and the economy can be portrayed statistically (including an historical dimension) should produce social and political benefits including more informed participation in community affairs and public decision-making. The realization of these benefits, however, will be limited by the extent to which public data users are able to confidently understand and manipulate the available statistical data, and to interpret its relevance to the information they are seeking. In short, statistical websites—as any others—need to be usable. Federal statistical agencies are aware that the transition of their user base from a small, discrete community of specialized professionals to the general public entails a responsibility beyond simply opening the doors of their data warehouses and inviting citizens to "come in and browse." There is a growing awareness of the need to build web gateways to federal data whereby non-specialists can find and get the information they want.

To date, however, these efforts to improve public data access have been largely isolated initiatives by an agency or group of collaborating agencies, with little overall coordination. As a result, despite a growing literature, usability findings rarely provide designers clear, prescriptive guidance that could be applied to their own interface design projects. We begin here to bridge this gap by developing an aggregated summary, in the form of (re-)design heuristics, of findings and observations from multiple usability tests of statistical websites. In these tests, representative general public end-users tried to find data to answer everyday questions that have statistical answers

[1] In addition, there are an increasing number of commercial websites that provide aggregate paid access to statistics collected by both public and private organizations (e.g., EconoMagic at http://www.economagic.com. In principle, these secondary providers should be responsible for the usability of their products. However, since some 'statistical retailers' simply provide a direct links to federal statistical sites, poor site design at the source "bleeds through," causing problems for users

Our heuristics are intended to guide interface designers focusing on the development of statistical data-mining interfaces intended for use by the general public. Some of the single-agency sites that have been reviewed[2] in the process of deriving these heuristics include:

- Crime statistics (Bureau of Justice Statistics): http://149.101.22.40/dataonline/, a sub-site of http://www.ojp.usdoj.gov/bjs/.
- Labor statistics (Bureau of Labor Statistics): http://stats.bls.gov

In addition, the design heuristics are also informed by observations from usability testing of the following multi-agency initiatives, which enable users to locate and extract data from multiple federal data repositories simultaneously:

- FERRET (Federal Electronic Research and Review Extraction Tool): A joint Census Bureau and BLS initiative: http://ferret.bls.census.gov/cgi-bin/ferret
- FedStats: the Interagency Council for Statistical Policy Project : http://www.fedstats.gov

In addition, the creation of these heuristics is informed more broadly by usability tests of statistically oriented websites which are oriented toward the retrieval of retrieving document-like charts or tables of aggregate.

- Energy statistics (Energy Information Administration): http://www.eia.doe.gov/
- Biological statistics: (National Biological Infrastructure Initiative site at the US Geological Survey) http://www.nbii.gov
- Science manpower and other resources statistics (National Science Foundation): http:/\www.nsf.gov/sbe/srs

While this third group of sites does not support the degree of flexibility or capacity to accommodate highly specific data requests that are the focus of the proposed heuristics, users of these and similar sites are observed to experience difficulties which are broadly congruent with the highlighted challenges. As such, designers of similar sites may also find a subset of the proposed heuristics applicable to their work.

Statistical Interfaces

Bosley & Conrad (2000) report meta-analyses of a series of usability tests focused on data access websites including websites at BLS and Census Bureau. Based on that work, they have presented a working model of the data access task outlining three principal, sequential sub-tasks[3]:

---

[2] The sponsoring agencies vary in the degree to which they have implemented the findings and recommendations that the tests produced. Therefore, these websites as they currently exist may not reflect the benefits of the tests performed.
[3] For simple data access scenarios, some or all of some steps may be omitted or skipped.

1. Users <u>specify the attributes</u> of the data of interest. Interaction tools such as check-boxes with pick-lists or drop-down lists may be provided to support this specification process. The system responds by returning a set of descriptors of accessible data that closely matches the user's specifications. In many cases, the system will return multiple possible matches, because each agency's storehouse of data contains thousands of multi-attribute items. This result set does <u>not</u> usually contain any actual data, but rather <u>labels</u> of data (sets) in the agency's data warehouse.

2. Users <u>evaluate the set of system responses against their initial attribute set.</u> In some cases, the user's specifications will exhaustively describe a data series that is stored in the accessible data, so that the system returns just one "hit" that exactly matches the user's specifications. More commonly the system will return a list of possible "hits" whose attributes partially match user specifications, so that users must scan these "possibles" and use judgment to make a final choice. Often at this juncture the user will need to get more descriptive information (metadata) than brief database labels provide, to make informed choices.

3. The user <u>selects the best-matching data set</u> from the array of potential "hits" that the system has delivered, and requests the actual data thus described. The user then must determine if the data itself meet expectations. The query results may be questioned if the data grossly violate user expectations. For example, a user may have used similar data from the recent past, and now finds the current data values so different that they appear implausible. In such a case, the user may iterate the query process from either step 1 or step 2. In the end, the task is complete when the data "satisfice" the user.

Although the sites examined in this analysis differed markedly in the scope of data that each linked to, the nature of the data, and output format options, the common set of major usability problems diagnosed across the interfaces included navigational errors, unintentional entry into navigational "blind alleys," and complete or partial failure to find data called for by test scenarios. Bosley and Conrad observed that these difficulties appeared to result from three problem <u>types</u> or as "root causes," including:

- Insufficient guidance and instructions for using the interface's functionality
- Complex or "cluttered" screen layout, leading to uncertainty and confusion, especially concerning procedural sequencing
- Deficiencies in labeling of data sets and lack of sufficient information to indicate how data are organized in a database

Subsequent testing of data access websites at BLS and by others working on other agency websites provide validation for the utility of model as a <u>problem diagnostic</u> to pinpoint critical systemic design problems that give rise to user frustration.

Despite great variability across the sites they examined, Bosley and Conrad (2000) identified a crosscutting set of usability problems that undermined the self-evidence and ease of use for those sites. Moving forward from this initial success and informed by subsequent usability testing efforts, we transform the Bosley and Conrad problem diagnostics into a set of design heuristics for interactive statistical websites[4] . These are intended as heuristics in the traditional sense: Procedures, typically applied to problem-solving activities, that if applied appropriately and proactively yield a high probability of success while effectively reducing the cognitive burden entailed by the task process (Tversky & Kahnemen, 1973). That is, ideally, these (and other) heuristics would be construed as fundamental shortcuts for the design process over which the rules apply rather than as a basis for reactive evaluation. Note critically, that there is a trade-off for the efficiency benefits of employing heuristics: The marginal risk of short-cut failure. However, this type of outcome typically only occurs when heuristics are inappropriate or haphazardly applied without consideration for the intended domain of the heuristic process.

What constitutes, then, an interactive statistical site? What characteristics must sites share—despite superficial differences—in order for the heuristics outlined here to apply meaningfully? The commonalities in site design for such a site are embodied in overall site architecture as detailed, individual screen layouts and content presentation modes that make it easy for inexpert users, including first-time or one-time users, to find the data they want quickly and accurately. Table 1 lists an overview of characteristic behavior demonstrated by inexpert users in usability testing environments, paired with general descriptions of infrastructure design improvements that might reduce the difficulties experienced by the users.

Taken together, these improvements point to a broader need for designers to provide "scaffolding," or guidance and performance support for users interacting with statistical websites. An additional challenge arises in that, at the same time, the enhanced design should not hinder more expert users' performance[5]. Sites that entail these characteristics are typically interactive at a very "fine grain." They will be both information- and instruction-rich. These sites enable user choices by selection among pre-specified alternatives using familiar vocabulary rather than free-text inputs. This not only reduces cognitive burden by the well-known superiority of recognition over recall, it also prompts inexpert users entering an unfamiliar semantic space. Well-designed sites will consistently confirm actions and reassure users that they are "on track" through meaningful feedback at progressive stages of data query specification.

---

[4] Although we focus on web-based interactive, statistical web-sites, the same usability challenges obtain for data-mining interfaces independent of delivery platform. other than the web.

[5] In fact, many of the tested sites contain "shortcuts" and other features for sophisticated users. Detailed discussion of these features, however, is beyond the scope of this paper.

Table 1 **User Performance Characteristics Map to Supportive Design Features**	
Inexpert User Performance Limits	**Supportive Design Features**
Lack of familiarity with how to search large data repositories efficiently	Step-by-step procedural guidance within and between screens
Lack of understanding of technical data definitions and data types/tokens that differ only slightly	Rich, easily accessible descriptions of data in lay language (extensive metadata[6])
Lack of familiarity with multivariate structure of data repository, plus overwhelming quantity of accessible data	Flexible navigation with abundant "You are here" cues (situational awareness support)
Desire to focus small scope, personal questions (e.g., local jurisdictions) rather than surveying the "big picture"	Support for efficient and accurate specification of a small but relevant set of data
Inability to cope with gaps and other irregularities in data and/or coping by adopting erroneous expectations and beliefs	Abundant warning against wrong turns, blind alleys. Consistent feedback about <u>correct</u> progress toward user goal
Inability to accurately interpret the data or observe and deduce meaning from data patterns	Presentation of output in a familiar format (table, simple chart), accurately and extensively labeled

Statistical Interface Design Heuristics

We propose a working set of 10 heuristic design rules for developing interactive statistical data-mining interfaces in Table 2. They pattern into three broad areas that are labeled in **bold** typeface. Subsequent reference to the heuristics in interface reviews may refer to either the broad categories or to a specific heuristic, by number. The remainder of this paper reports on the application of these heuristics as a framework for critical review of interactive statistical sites. We hypothesize that end-user testing of these sites will identify usability problems parallel to those that are uncovered through systematic application of the reviews using the heuristics. As such, this effort represents an important first step in the external validation step of the heuristics.

[6] Although it is beyond the scope of this paper, interested readers are directed to Dippo and Gillman (1999) for a comprehensive presentation of metadata and the related issues.

Table 2
Proposed Heuristics for Interactive Statistical Interfaces

Orient the user to the available body of data
1. *Give an overview of the available data*
Provide general orienting information about the data that can be accessed using the interface. Highlight indicators of data scope, restrictions, deliberate omissions and other important characteristics.

2. *Support situational awareness within the available data*
Use text or graphics to propagate the data structure across levels. Tell users when they enter a disjoint partition of the data. Make it easy for users to return to the initial state (screens which link to major divergent paths?)

3. *Display and clearly define metadata*
Embed sufficient metadata with tools to get users started. Provide easy, "just-in-time" access to definitions of technical or unfamiliar descriptive terms encountered as querying proceeds. Avoid use short , cryptic labels for data sets or variables.

Design the interface for interacting with the data
4. *Put adequate and clear instructions on the interface*
Tell users explicitly how to work with the interactive elements on the interface. Make the association between an interactive element and user guidance clear and available. Set defaults in all data specification "widgets;" defaulting to the broadest specification is recommended.

5. *Link users to frequently requested analyses*
Provide links to frequently requested numbers or datasets. Store common queries for novice users to modify to suit their needs. Build shortcuts for advanced users' quick data access.

6. *Use simple interaction schemes to accomplish complex query-building*
Use logical task sequences or natural language instructions to support advanced Boolean query syntax. Enable users to add or exclude data selections, and express optional inclusion ("or") by making a series of clear, discrete choices.

7. *Summarize outcome of complex data specification for review and confirmation*
When users apply multiple filters, especially across multiple screens, display the final specification for review and confirmation before a user submits a data request.

Help users anticipate, interpret and evaluate results
8. *Offer choices of easy-to-interpret output formats*
Offer users choices among well-known, understandable outputs like tables or simple graphs. Use graphics, or actual examples of output formats, as well as text to describe output options.

9. *Design output formats to facilitate quick and reliable query validation*
Make output labels consistent with variable selection options in query. Make labels clear and highly visible on statistical tables. Support keeping row and column labels visible as user explores table contents (large tables).

10. *Help users avoid searching for non-existent or non-available data*
Warn against, or actively prevent, requests for missing or unavailable data. Notify user when query will return a null result.

<u>Site Review 1: FERRET</u>

The Census Bureau and Bureau of Labor Statistics support a data query tool that provides aggregate access to survey databases including their own and the National Center for Health Statistics (NCHS). Here we step through the process of attempting to complete the task:

Usability Task: How many elderly individuals <u>living in</u> the state of Maryland died of stroke in 1994?[7]

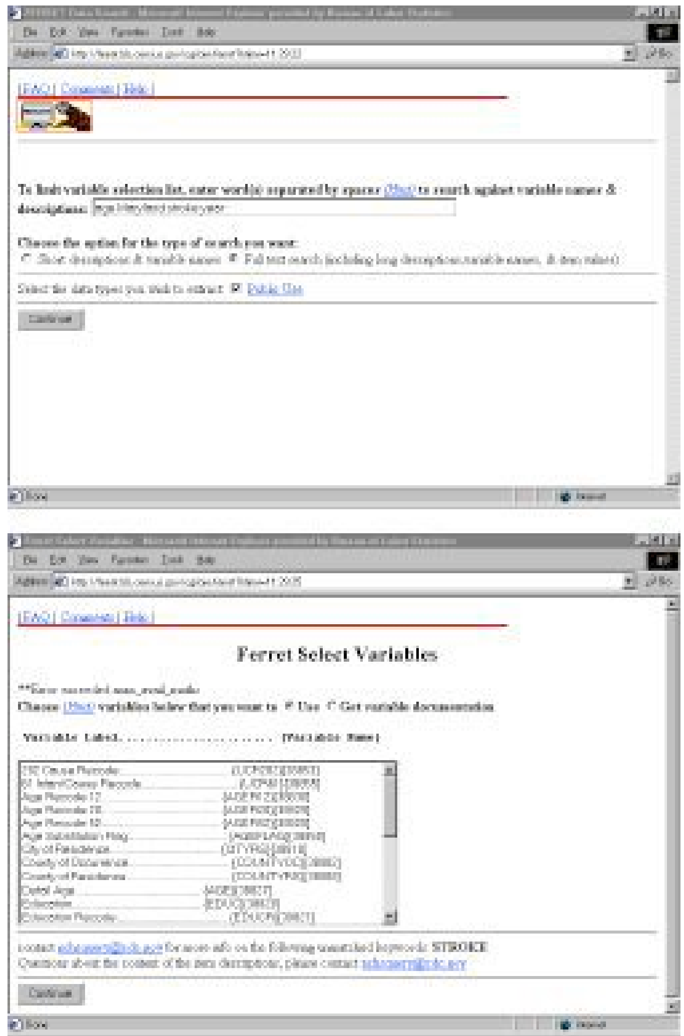

Figure 1: Initial FERRET screens

Figure 1 shows the initial screens for formulating a data search using FERRET. Here the user is to enter key words representing the question that the user wants answered. This approach violates several of the proposed heuristics, including #1, #3, #4, #5 and #6. Although, there is not a clear violation of #2 <u>on this screen</u>, evaluation of subsequent screens will indicate whether or not there is compliance. Incidentally, unless the "Full text search" option is selected, many of the key words will not be found in the short descriptions and variable names. The user has to discover this by trial and error. Once the user submitted the search string, the second screen in Figure 1 appears. This screen also violates many of the design heuristics. While it does a better job of providing a data overview, it fails to provide information about the data structure (#2) The metadata, including variable names and labels, are cryptic and difficult distinguish[8] (e.g., the three age recodes.) This screen largely ignores the entire set of heuristics dealing with design-for-

---

[7] For the sake of completeness mortality data is specified in this question even though this data is available for only one year.

[8] In the service of fairness to the developers, FERRET staff does not have control over this cryptic terminology, however. Names and labels were specified when the database was set up and FERRET can only display what is available to it. This fact, however, should serve to

interaction (#4, #5, #6 #7), as well. Compliance with the last three output-oriented heuristics remains to be evaluated at a later stage. Note that "stroke" was an unrecognized key word, constituting an additional violation of heuristic #3.

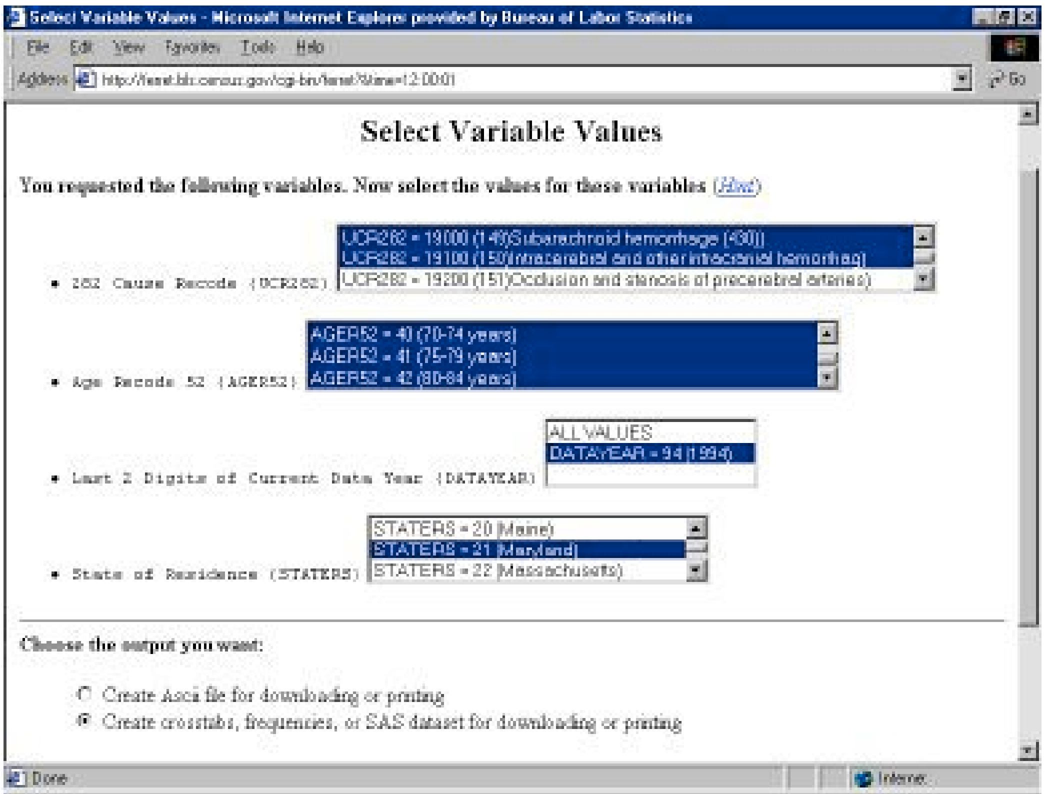

Figure 2: FERRET data specification page #2

As can be seen in Figure 2, the subsequent screen allows users to winnow data within the variables indicated in their free text query entry. To save space, value selections for the usability task are indicated on this screen shot, although in reality the screen presents initially with no default values selected.

---

encourage statistical interface designers to engage in a systems level usability review prior to embarking on detailed design.

To complete the usability task, the following data selections were made using the accompanying rationale:

- *282 Cause Recode*: in an attempt to get cause of death
- *Age Recode 52*: an arbitrary choice to specifying for the age range "elderly."
- *State of Residence*: in expectation of being able to select the state of Maryland.
- *Last 2 Digits of Current Data Year:* indicate date

This screen consists of secondary data specification. Once users arrive at this point, the design becomes more compliant with the orienting heuristics least #1 and #2, albeit within a highly restricted range. Since the metadata or selectable variables can only be viewed a few items at a time, compliance with #3 is still marginal The screen complies poorly with heuristic #4: The user is left to infer that multiple selections are possible by using the "Control" or "Shift" keys .

Compliance with heuristic #6 has vastly improved on this screen. Complex specification can be built by making discrete selections if visible variable values (the variables "of interest" were set on the prior screen.) Both good and bad usability results obtain. For instance, again in violation of heuristic #3, the use of cryptic labels in some lists forces the user to make some complex choices involving what is probably unfamiliar, technical language. For example, the user needs to know that stroke must be mapped to two specific medical categories (*Subarachnoid hemorrhage* and *Intracerebral and other intracranial hemorrhage)* to capture the range of data intended on the lay definition for the word 'stroke.' In addition, some labels, for example, "All values," force the user to guess what the result would be if that value were chosen. Compliance with heuristic #8 is also questionable: Only two output options are offered. Further, one of these is labeled in three different ways—cross tabs, frequencies, and SAS dataset—terminology that is likely to be impenetrable to non-analysts. On the other hand, this screen demonstrates increased compliance with several heuristics. The (albeit initially arbitrarily-chosen) age recode variable has narrowed age categories so that the user can construct a reasonable data range for the subjective concept of *elderly.* In addition, this screen does a reasonably good job of complying with heuristic #7 offering the user feedback on the variables selected, while also supporting the additional value-selection operation. This page also implicitly complies with heuristic #10, in that the selectable values are limited to those that exist.

The first screen in Figure 3, which occurs first in the task sequence, provides users an opportunity to both review current data specifications and output specification. This screen provides a solid demonstration of the spirit of heuristic #7, offering users a summary confirmation and opportunity to modify their complex data request. However, this positive design feature may be offset by poor compliance with heuristic #3: the "metadata" provided are obscure code values that a non-specialist cannot interpret, largely undermining the benefits of the good design-for-verification. That such trade-offs could occur supports both taxonomy of the heuristics and the complex nature of the useable design. Compliance with heuristic #8 to facilitate output format choices is marginal. This, however, is due in part to the lack of complete instructions for the specification widgets, a violation of heuristic #4.

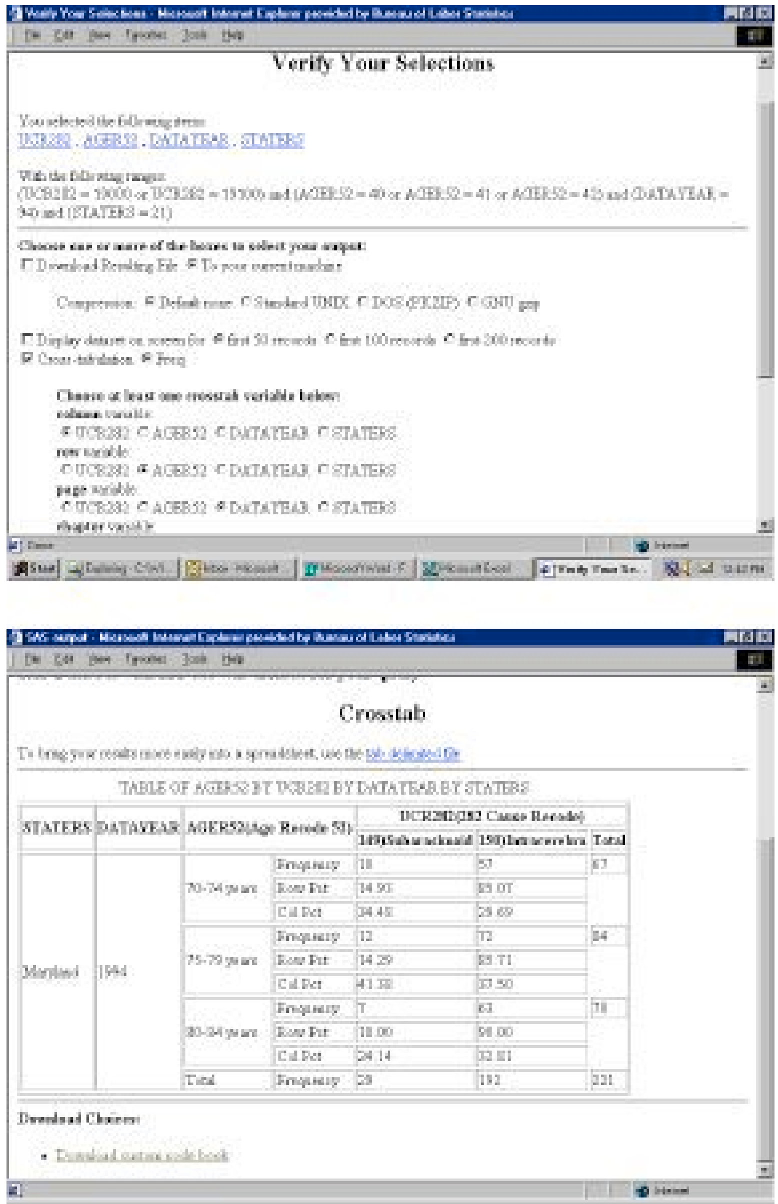

Figure 3: FERRET query review and data output screens

The second screen in Figure 3 screen shot shows the results of this data access process, in cross-tabular form. Substantial extraneous information (e.g., table computation/processing time) that appears above the table has been omitted from the figure. While the complete output is viewable in a single screen, the rigid formatting conventions of the cross-tabulation make it challenging to scan the resulting table.

Further, although the problem is not exemplified in this figure, indicating or nominating variables in the appropriate order on the data query construction screens poses an additional significant user challenge. The choices must be ordered so that single-value variables (in this case, *year*, for example) do not subdivide the tabulation in a confusing manner. In many cases solving this problem entails several back-and-forth maneuvers to reorder variables, imparting a significant cognitive load on the user in terms of process memory and recall of variable specification language. More concretely, if “state” had been labeled as the row variable in the first screen of the figure, then the tabulation would have been broken down into sub-tables, with the “Maryland” row label repeated multiple times.

As shown in Table 3, the heuristic review of FERRET reveals examples of both heuristic violations and demonstrations of successful application of the proposed design heuristics.

Table 3 **Heuristics demonstrated and violated on FERRET**		
Page in task flow	**Demonstrated**	**Violated**
Overview/Entry Page		1, 3, 4, 5, 6
Data Specification		2, 3, 4, 5, 6, 7
Data Specification 2	1, 2, 6, 7, 10	3, 4
Query Review		3, 4
Data Output		9

Unfortunately, the effect of the frequency and number of violations is exacerbated by their occurrence early in the data-mining process. The predicted resulting effect is that early violations will outweigh supportive design details encountered later in the task flow. This summary predicts that formal usability testing with non-expert participants would uncover significant performance challenges.

Site Review 2: CDC WONDER

The Centers for Disease Control (CDC) webtool "WONDER" can be used to find health-related statistics on the CDC website. This tool supports user registration so that a user's prior analyses can be stored and retrieved via password-protected site areas [9]. Thus, the site complies with heuristic #5 at an individual level, although it is less compliant with the heuristic's spirit that the interface provide general level of access to popular or frequently requested analyses.

Figure 4 presents the initial screen within the tool's core functionality. While <u>all</u> of these functions are accessible on a single page, as observant readers have likely noticed, the multiple scrolls required to interact with all of the variables imposes an early processing/memory load on who need to develop a model of where they can specify or have specified what. For illustrative

[9] This heuristic-based review will not include the log-in screen and a textual "overview" screen so as to focus on the key functionality of the access tool. However, those pages demonstrate good compliance with the "user-orienting" subset of heuristics (#1-#3), as the reader can see by visiting the WONDER home page at http://wonder.cdc.gov/ .

purposes, the interface will be used to perform the same task that was used in the FERRET review.

Usability task: How many elderly individuals <u>living in</u> the state of Maryland died of stroke in 1994?

The correct dataset (mortality data) has been selected from a list on a prior screen. That list strongly complies with heuristic #2, since selection of a particular database excludes all others from the data search (disjoint partitioning.) Thus the user "knows" that specifying data on the interface in these illustrations will refer to mortality data and not some other dataset. The list also conforms to overview heuristic (#1) to the extent that a linked page gives easy access to a clear description of each dataset in the list.

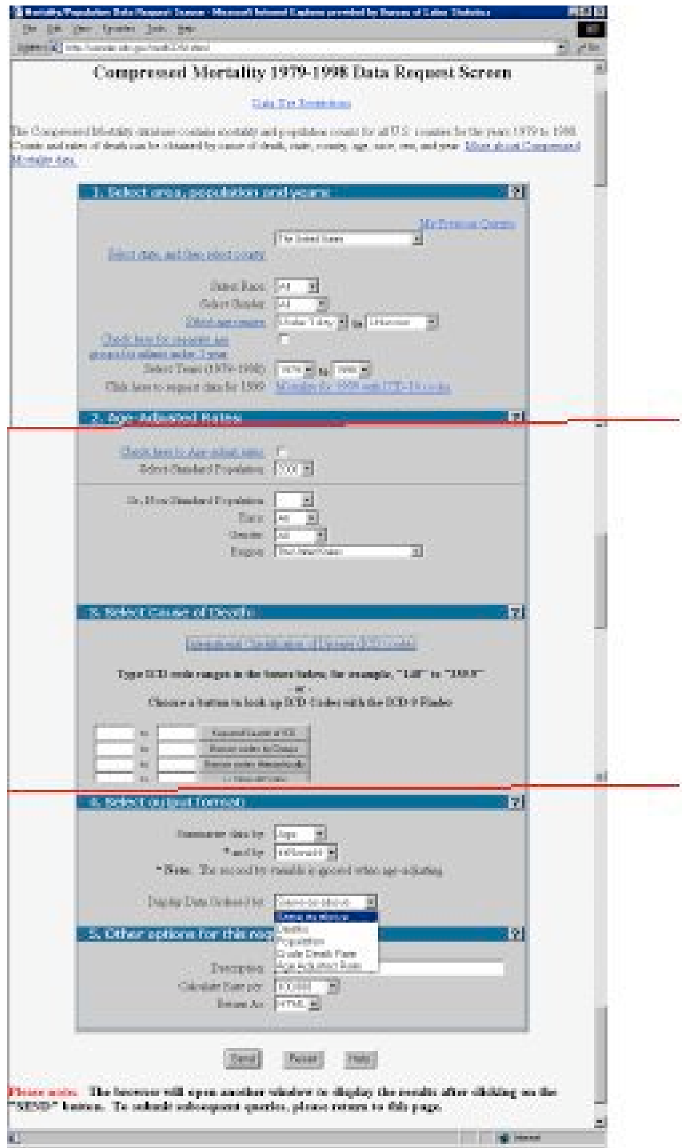

Figure 4: CDC Wonder orienting page

The user's steps to complete the task are numbered sequentially, beginning with this first step named "Select area, population, and years." The interface shows good compliance with the "interaction design" heuristic set (#4-#6): There are ample and clear instructions, links to prior analyses (at least for this user), and specification proceeds by a series of discrete steps. Specification fields have appropriate default values. Compliance with the final interaction design heuristic #7 remains to be assessed as the data search process goes forward. Step 2 of the numbered task process entails further data specification. Here, WONDER users move seamlessly between the levels of variable specification by scrolling up and down on the same screen. Recall that on the

FERRET model, users were propelled to a second screen to further specify their dataset. In this case, the second model places considerable cognitive burden on the user, who has to recall actions taken on prior screens until some feedback is provided that summarizes and shows the results of those actions explicitly [10].

Throughout steps 2 & 3, the interface demonstrates compliance with the subset of interaction design heuristics. The link "Check here to Age-adjust rates" serves both to provide clear instruction for use of the juxtaposed check-box and to provide the user direct access to a pop-up metadata definition of "age adjustment" as a procedure. This simultaneously complies with heuristic #3 and #4. The link labeled "International Classification of Disease (ICD) codes" under step 3 serves the same dual explanatory function for the specification widgets that appear below it. The ICD list is extensive and highly technical, thus there is a critical need for such strong metadata support is essential for non-expert users because of the failure to avoid cryptic data labels, a violation of Heuristic #3. Although it is difficult to capture an image of the interface that fully illustrates the simplicity and flexibility of the output format specification of CDC-WONDER, one of the several drop-down lists that provide this flexibility is opened in the illustration of step 4. The entire screen or interface complies with heuristic #7, since by scrolling, the user has direct access to review, confirm and revise the data selections made up until the data are requested. The retrieved data are presented in a separate window which preserves and maintains the availability of the selections made on the specification interface.

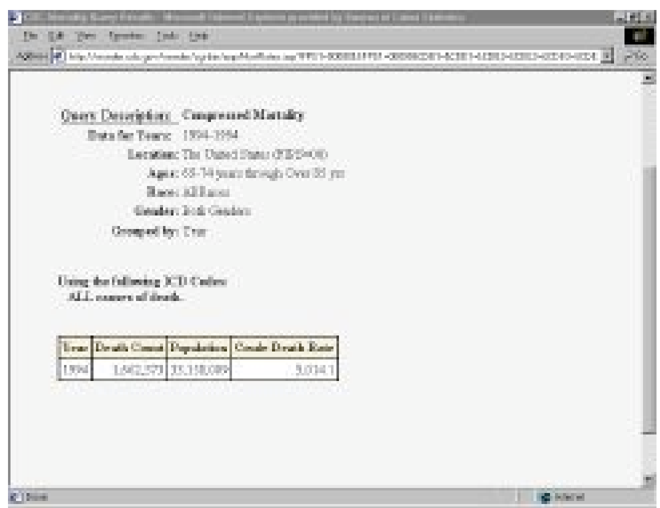

Figure 5: CDC-WONDER data output

Figure 5 presents the CDC-WONDER data output screen. Here the fact that WONDER demonstrates compliance with heuristics #9, allowing users to quickly and easily verify that the data returned corresponds with that which was requested, is absolutely critical since the returned data does not match the data request[11]: In general, however the process of verifying that the

---

[10] Although long scrolling pages represents a widely touted usability "Don't," in the specific context of interactive statistical interface design, the scrolling-page layout offers better user support than the successive-screen specification-building model.

[11] The data that are returned are for the entire United States—not just Maryland and include deaths from all causes—not just "stroke" as defined from the ICD codes in terms identical to those used in the

requested data was generated (or not) by the system is straightforward, and the detailed page design provides a visual hierarchy which supports scanning for details. Although the identification of geographic areas as FIPS (Federal Information Processing Standards) codes may be unfamiliar to the ordinary user, it is presented as secondary information to the more familiar taxonomy (Location = United States) code, suggesting that this label may be intended to provide quick support for advanced government analysts.

As shown in Table 4, the heuristic review of CDC's WONDER reveals significantly more demonstrations of successful application of the proposed heuristics than violations. In addition, these positive design characteristics occur early in the task flow. In comparison with FERRET's operation on the same dataset, WONDER gives the user much more guidance as well as flexibility and access to essential metadata. Based on this heuristic-driven review we would predict that individuals from the general public participating in a formal usability test of WONDER would enjoy significant successes and that any usability consistent problems uncovered would represent minor issues.

Table 4
Heuristics demonstrated and violated on CDC's WONDER

Page in task flow	**Demonstrated**	**Violated**
Overview/Entry Page	2	
Data Specification	1, 2, 3, 4, 5, 6, 7	3
Query Review	7	
Data Output	9	

FERRET example. This outcome violates an even stronger prohibition articulated previously by Levi & Conrad (2001): Don't lie to the user. At the time of this review, an internal server error resulted in the presentation of misleading output. Unfortunately, the system failed to provide any warning or notification that there was any problem on the server side. Thus, while this explanation offers helpful closure to the unexpected failure of the interface, its discovery required significant sleuthing on the part of the first author. It is not realistic to expect users to exhibit similar behavior.

Site Review 3: Bureau of Justice Statistics

Our third review focuses on the Crime and Justice Data Online statistical interface developed by the U.S. Department of Justices, Bureau of Justice Statistics (accessible at: [http://149.101.22.40/dataonline/)](http://149.101.22.40/dataonline/). This site provides the public direct access to significant and interesting crime statistics collected by BJS from both law enforcement agencies throughout the United States and by the FBI.

Figure 6: Initial screens for BJS Data

The initial screen for Data Online, shown first in Figure 6, links to the three major partitions of the available data. The first two links, "Crime trends from the FBI's Uniform Crime Reports," and "Homicide trends and characteristics," are the subsets of the data of greatest public interest and as such, are the focus of this review. The third area, "Law Enforcement Management and Administrative Statistics," tends to be accessed more professionals concerned with administrative issues within law enforcement.

This initial page demonstrates general compliance with the data orientation heuristics (#1-#3) and with the applicable interaction design heuristics (#4 and #5.) The user is shown explicitly where this page is situated by the path representation at the top of the screen (heuristic #2). This situational information is augmented and carried through to subsequent screens, where it emphasizes that selecting one of the segments on this screen excludes exploration within either of the other two.

The second screen, which situates the user in the selected data set, essentially replicates the format of the initial screen supplemented by additional detail to the metadata (heuristic #3).

Taking a slightly different tack from the previously discussed interfaces, BJS Data Online promotes heuristic #8 (Offer choices of easy-to-interpret output formats) to the data specification stage, demonstrating creative and proactively design by providing users <u>samples</u> of the available data output formats in the data specification phase of the task. Thus users know <u>exactly</u> how the output will appear, and whether it matches their specification goals before they invest in specifying the variables. These simple output formats naturally lend themselves to conformity with heuristic #9, “Design output formats to facilitate quick and reliable query validation.” The simplicity of the underlying datasets as well as the openly-displayed metadata also makes it easy for the design to comply with heuristic #10, guarding against lost task effort invested in searching for data that does not exist on this website.

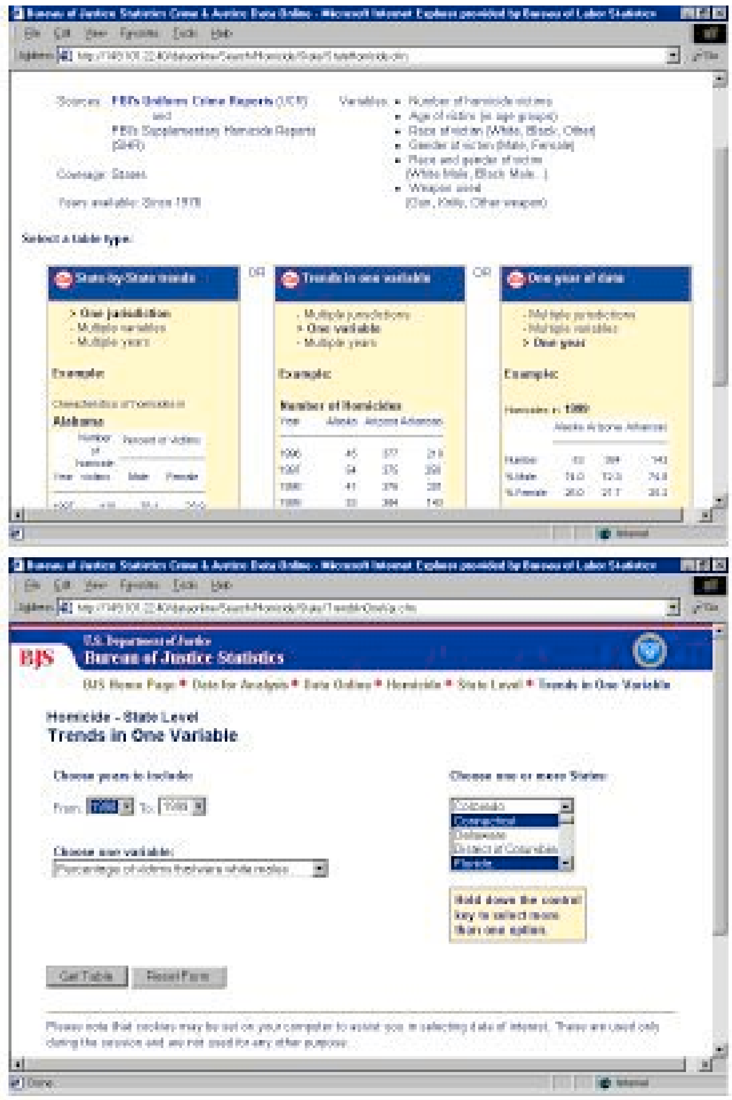

Figure 7: Data specification screens on BJS Data Online

Selecting a table formats propels the user to a screen of data selection widgets that includes <u>just that set of selection options that matches the chosen output format</u>. This dynamic configuration of a specification interface to match the desired output is a powerful way of implementing compliance with heuristic #6 and incidentally avoiding the need to comply <u>explicitly</u> with heuristic #7.

The second screen in Figure 7 shows the specification screen corresponding to the “Trends in one variable” (middle) output format above, with some output options selected complete the user task:

Usability Task: What percentage of homicide victims in Connecticut and Florida were white males during the decade of the 90’s?

Note the clear and visible instruction to support assimilation of the multiple-select option within the data selection widgets. This clearly demonstrates proactive compliance with heuristic #6 to provide clear and visible instructions. Further, labels for actions such as “Get Table” are transparent and predictable, demonstrating compliance with heuristic #4.

One potential difficulty for users of this interface is the requirement to remain aware of both which of the three disjoint dataset the user is currently in and which dataset will have the data relevant to the current query task. In this example, this entails recalling from prior screens that the source for these data is the FBI’s Uniform Crime Reports and Supplementary Homicide Report. Carrying that metadata information on to this screen could improve the design from the standpoint of heuristic #3, generally.

Since the structure and scope of the data available within each sector is simple to describe and define, like WONDER this data specification process omits the separate for reviewing and confirming data specifications step (heuristic #7) as unnecessary.

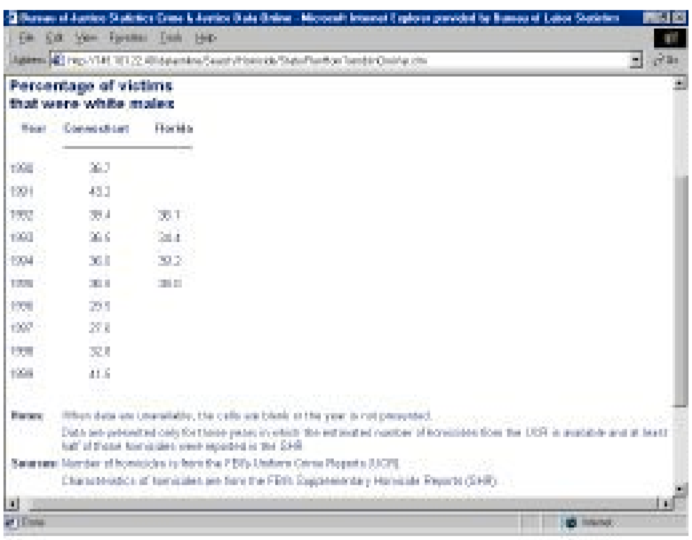

Figure 8: BJS Data Online Result Table

Figure 8 shows the system output. Note that clear and complete metadata are again supplied at this stage. Further, the table supports quick, reliable validation of a match between the data requested and those supplied (heuristic #9.) The output (and by extension the site design) indirectly handles compliance with heuristic #10, avoiding search for non-available data, by <u>omitting</u> data that are missing or do not meet quality criteria that the notes clearly outline. For the heuristic review task, the user was able to complete a data specification <u>despite</u> the absence of some of the specific data elements, since the specification terms did not extend to the level of detail represented by a single cell in the table.

As shown in Table 5, the review of BJS Crime and Justice Data Online almost exclusively reveals demonstrations of the proposed heuristics. Thus, in a formal usability testing review, we would predict that the overall success rate for this site would pattern more closely with CDC’s WONDER than with FERRET. Based on the heuristic-driven evaluation we would predict only minor or cosmetic issues would be uncovered in the lab, with the single exception that users might periodically demonstrate difficulty completing a task because they had unknowingly become mired in the wrong partition of the available but disjoint data sets.

Table 5 **Heuristics demonstrated and violated on BJS Crime & Justice Data Online**		
Page in task flow	**Demonstrated**	**Violated**
Overview/Entry Page	2	
Data Specification	1, 2, 3, 4, 5, 6, 7	3
Data Specification 2	7	
Data Output	9	

Concluding Observations

The heuristic-driven reviews presented here provide a preliminary foundation for validating the set of proposed heuristics as a useful framework for designing and analyzing highly interactive statistical interfaces. The review results lead us to hypothesize that of the three examples, end users would have the hardest time performing tasks using FERRET, and a relatively easier time performing a very similar task with CDC WONDER and the BJS Crime and Justice Data Online interface. In fact, independent usability tests of the FERRET interface discussed here resulted in significant frustration for non-expert participants while tests of BJS Crime and Justice Data Online uncovered no pervasive usability problems. For this reason, the FERRET interface is being extensively reworked. While no comparable data is available for WONDER, a formal end-user test to evaluate whether the usability results are generally congruent with the review findings is under consideration.

Interestingly, the review of FERRET served to crystallize two additional meta-considerations for the application of heuristics in design and interface evaluation. First, the cumulative usability of an interface entails not only the relationship between the frequency of positive and negative instantiations of design heuristics, but also a consideration of the temporal sequence and local density of the occurrences of these tokens within the task flow. That is, any benefit gained by thoughtful design late in the task flow can be fundamentally offset by heuristic violations occurring early in the task flow. Second, effective design constitutes a global understanding of the intended context for heuristics and appropriate application of the complete set. Even armed with clear, concise guidelines, interactive statistical interface design is not a simple activity.

In summary, the work presented represents a first step in refining and validating our proposed set of design heuristics for interactive statistical database interfaces. Through additional evaluation, testing and review, we expect to develop converging and more comprehensive evidence that, appropriately applied, these heuristics can serve to provide interface developers clear guidance in the initial design process. In addition, we hope that by providing clear, foundational guidance, more agencies and developers will feel able to confidently design and implement highly interactive websites providing their users with greater capability find and retrieve just that data which is highly relevant data for their purposes easily, quickly, and with a high level of confidence.

Bibliography

Bosley, John J., and Conrad, Frederick G., (2000) "Usability Testing of Data Access Tools," paper presented at the *International Conference on Establishment Surveys*, Buffalo, New York, June 17-21, pp. 971-980.

Dippo, Cathryn, S., and Gillman, Daniel W. (1999) "The role of Metadata in Statistics*" Paper presented at the UN/ECE Work Session on Statistical Metadata.*

Levi, Michael D., and Conrad, Frederick G. (1996) "A Heuristic Evaluation of a World Wide Web Prototype," *Interactions Magazine,* July/August, Vol.III.4, pp. 50-61.

Levi, Michael D. and Conrad, Frederick G. Usability Testing of World Wide Web Sites. manuscript available: http://www.bls.gov/ore/htm_papers/st960150.htm

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1130.

Acknowledgements

We would like to thank Fred Conrad, Peter Doyle, Catherine Gaddy, Jean Fox and Marianne Zawitz for their helpful and timely interactions and feedback in the development of the proposed heuristics and this manuscript. The opinions expressed in this paper reflect the beliefs of the authors and do not represent the opinions of either the US Bureau of Labor Statistics or Human Factors International.

Correspondence:

Please direct questions, comments and reprint requests to:

Kath Straub
kath@humanfactors.com
410.327.7082

Complete Citation

Bosley, John J. and Straub. Kathleen A. (2002) Data exploration interfaces; Meaningful web database mining by non-statisticians. *Proceedings of the IBM Make IT Easy Conference*. San Jose, CA,

For clearer images, download the pdf:
http://www-3.ibm.com/ibm/easy/eou_ext.nsf/Publish/2773