Navigational Issues in the Design of On-Line Self-Administered Questionnaires

Kent L. Norman, Zachary Friedman, Kirk Norman, and Rod Stevenson

Human-Computer Interaction Laboratory
Institute for Advanced Computer Studies
Department of Psychology
University of Maryland, College Park, MD 20742-3255

Introduction

Up until recent years surveys and questionnaires have been administered either by an interviewer or in paper-and-pencil form. The human interviewer has the advantage that a knowledgeable individual assists the interviewee with (a) the sequencing of items, (b) the interpretation of items, (c) the recording of the responses, and (d) the editing of errors. The disadvantage is that human interviewer (a) must be trained, (b) is not always conveniently available, (c) is expensive, (d) makes errors, and (e) can introduce bias into the interviewee's responses. Moreover, the interviewee is somewhat at the mercy of the interviewer who knows the overall scope and content of the survey and reveals the items only one at a time.

The paper-and-pencil form has the advantage of being inexpensive. The disadvantages are that (a) items may be confusing without the aid of interpretation, (b) forms may be hard to follow from question to question and from page to page, (c) omissions and errors of various types can go without detection until it is too late to fix them, and (d) respondents may be less likely to start or once started to finish the questionnaire on paper than they would be under the personal persuasion of a human interviewer. However, paper-and-pencil forms do allow the respondent to browse the scope and content of the survey and the freedom to answer questions in any order.

Recently, the computerized self-administered questionnaires (CSAQ) have become an attractive alternative to human interviewers and paper-and-pencil forms. This is particularly true as access to computer facilities increases and as distribution systems such as the World Wide Web (WWW) become more and more pervasive, accessible, and easy to use (Dillman, 2000).

CSAQs have the potential of being even more inexpensive than paper-and-pencil forms and interactive like human interviewers. They may be less expensive particularly on the WWW because one does not need to print or distribute a physical media. Furthermore, the responses are already in an electronic form and do not need to be coded and/or physically sent. Rather, they can be electronically transmitted to a central server and automatically processed and analyzed. Moreover, CSAQs can be written in a number of interactive ways to assist in (a) sequencing of items and navigation of the forms, (b) giving definitions and clarification of terms, (c) recording responses, and (d) helping to avoid, detect, and correct errors at the point of entry or on completion of sets of items.

The goal of this research is to assist the U.S. Bureau of the Census in developing principles and guidelines for interactive CSAQs. This work draws upon current theory and research in human/computer interaction and cognitive psychology (Norman, 1991; Shneiderman, 1998), techniques of task analysis (Card, Moran, & Newell, 1981), and finally, empirical

research on alternative design implementations. The remainder of this paper focuses on the first two fundamental issues in designing interactive surveys: the sequencing of items and the navigation of forms.

Task Considerations

Before considering surveys and questionnaires on-line it is instructive to consider how people manage paper-and-pencil and interviewer assisted questionnaires.

Paper-and-Pencil and Personal Interviews

First with paper and pencil surveys, the items are by necessity laid out in a fixed, sequential order on pages. The respondent can read through and answer the items in the printed order or jump around to see the overall content of the survey. The respondent can easily get an idea of the length of the questionnaire and where the questions are going. Generally, the respondent will answer the questions in the order that they are printed, but in some cases skip sections and return to them later or jump ahead to questions that they want to complete right away.

In personal interviews, the interviewer controls the order of the items, presents only one question at a time, and waits for the response before going on to the next item. The interviewee can ask how long the survey will take and what it is all about; and generally the interviewer will answer these questions. Sometimes, the interviewee may want to skip questions and come back to them later. The interviewer might oblige but must take the responsibility of remembering to complete those items at a later point.

In either case, it must be remembered that as people are asked questions in a prescribed order or scan questions on a printed form, they must access the information from memory or from some external reference to answer the questions. Many factual questions may be answered from memory and, as such, will be subject to all of the problems associated with retrieval from human associative memory (Collins & Quillian, 1969).

Since retrieval of information is subject to the organization of long term memory, it makes sense to organize the questionnaire in a manner that is congruent with memory access. Figure 1 shows a schematic of memory and a survey with randomly ordered items.

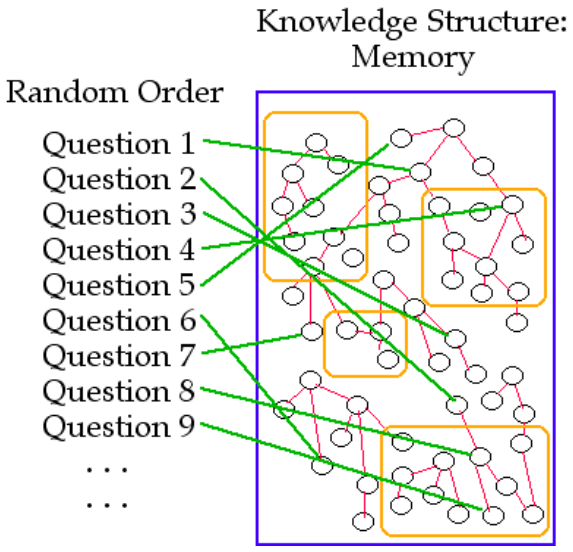

Figure 1. Schematic of a randomly ordered questionnaire accessing information in memory.

The right panel shows a memory structure defined by associations and clusters. The items in the questionnaire on the left randomly access different clusters and associations. It is assumed that access of items within clusters and associations will be faster than access across clusters and not connected by associations. Specifically, it is assumed that

$$\text{Total Time} = nt_r,$$

where n is the number of unassociated items and t_r is the time to answer a question drawn at random.

Alternatively, Figure 2 shows a questionnaire that is organized to access sets of items in memory clusters.

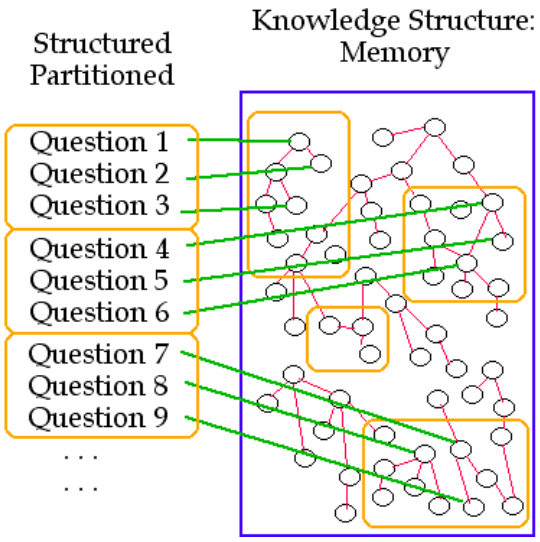

Figure 2. Schematic of a semantically ordered questionnaire accessing information in memory.

If there are n clusters of items, then

$$\text{Total Time} = nt_r + (n_c - 1)t_c,$$

where n_c is the number of items in a cluster and t_c is the time to answer a question within the same cluster. This strongly suggests that when respondents are answering questions from memory, questions should be organized and clustered according to their organization in semantic memory.

On the other hand, respondents may need to consult external sources in order to answer factual questions. They may look at their driver's license, birth certificate, medical records, checkbook ledger, company books, etc. Access to personal records, company records, and other sources will be subject to the physical accessibility and organization of these records as well as the respondent's ability to search and retrieve information from these databases. Both internal and external sources are organized in ways that dictate or influence in some way the order of retrieval of information. When there is a match between the order of questions in a survey and the order of retrieval of information, processing is most efficient. When there is a mismatch, processing is less efficient and may result in errors. For example, if a person is asked to name the days of the week in temporal order, it is very easy. However, if they must list them in alphabetical order, it is not so easy (i.e., Friday, Monday, Saturday, Sunday, Thursday, Tuesday, Wednesday). Similarly, one is more likely to recall the names of one's children in chronological birth order rather than alphabetically. However, records of employee files are more likely to be in alphabetical order rather than by date of birth.

Attitudinal and personal opinion questions can only be answered from one's own personal experiential memory either directly or through a process of subjective assessment. This last type has been subject to much study. Attitudes are not likely to be stored or retrieved in fixed order although they may occur in clusters pertaining to subject (e.g., attitudes about foreigners) or linked by associations. Typically, items in questionnaires are grouped by topic. This may help to set the context for retrieval of attitudes and facilitate the process.

On-Line Design Space

On-line surveys and questionnaires can be implemented in many different ways (Lazar & Preece, 1998). In fact the design space of possibilities is enormous. Consequently, it is extremely important that interface development for computerized questionnaires and the software used to implement them be guided by reasoned principles and empirical research.

Just as surveys and questionnaires implemented in the print media or the personal interview must take into consideration the constraints and dynamics of their environments, the CSAQs must take onboard all of the issues and factors involved in the computer interface. The current design of computer interaction is focused on the graphical user interface(GUI). This interface is characterized by buttons, fields, windows, and mice. When surveys and questionnaires are implemented in GUIs, they must be designed around the functions, facilities, and limitations of these tools. For example, windows that present surveys must be either scrolled or paged. The mouse can be used to select options to answer questions and buttons and scroll bars can be used to navigate through the questionnaire. Many other GUI tools and techniques are used to mediate input/output and control of the user interface.

Interfaces are modified to accommodate tasks and materials. In turn, materials are also modified to assimilate tasks and interfaces. For example, documents, whether manuscripts or surveys, are generally divided into smaller chunks for several reasons. They may be divided into meaningful sections to help define a mind set for thinking about a topic and answering questions of a particular vein. This will be referred to as "semantic partitioning." The smallest unit in a questionnaire or survey is the individual item, whether a question or a rating scale. The respondent may be presented with items one at a time thereby focusing only on that item or presented with groups of items that provide a context or an allowance for comparison. All of these divisions are based on semantics of the questionnaire.

On the other hand, documents may be divided into pieces that fit within a physical space. Paging by length is a physical constraint imposed by paper size or by screen size. Page divisions interact with semantic divisions since they are usually out of synch. Page breaks may arbitrarily subdivide semantic partitions of a questionnaire. Worst of all is when page breaks fall within items of the questionnaire so that part of the question is on one page and the rest is on a subsequent page. If at all possible, this is to be avoided. Figure 3 shows a schematic of divisions of a survey.

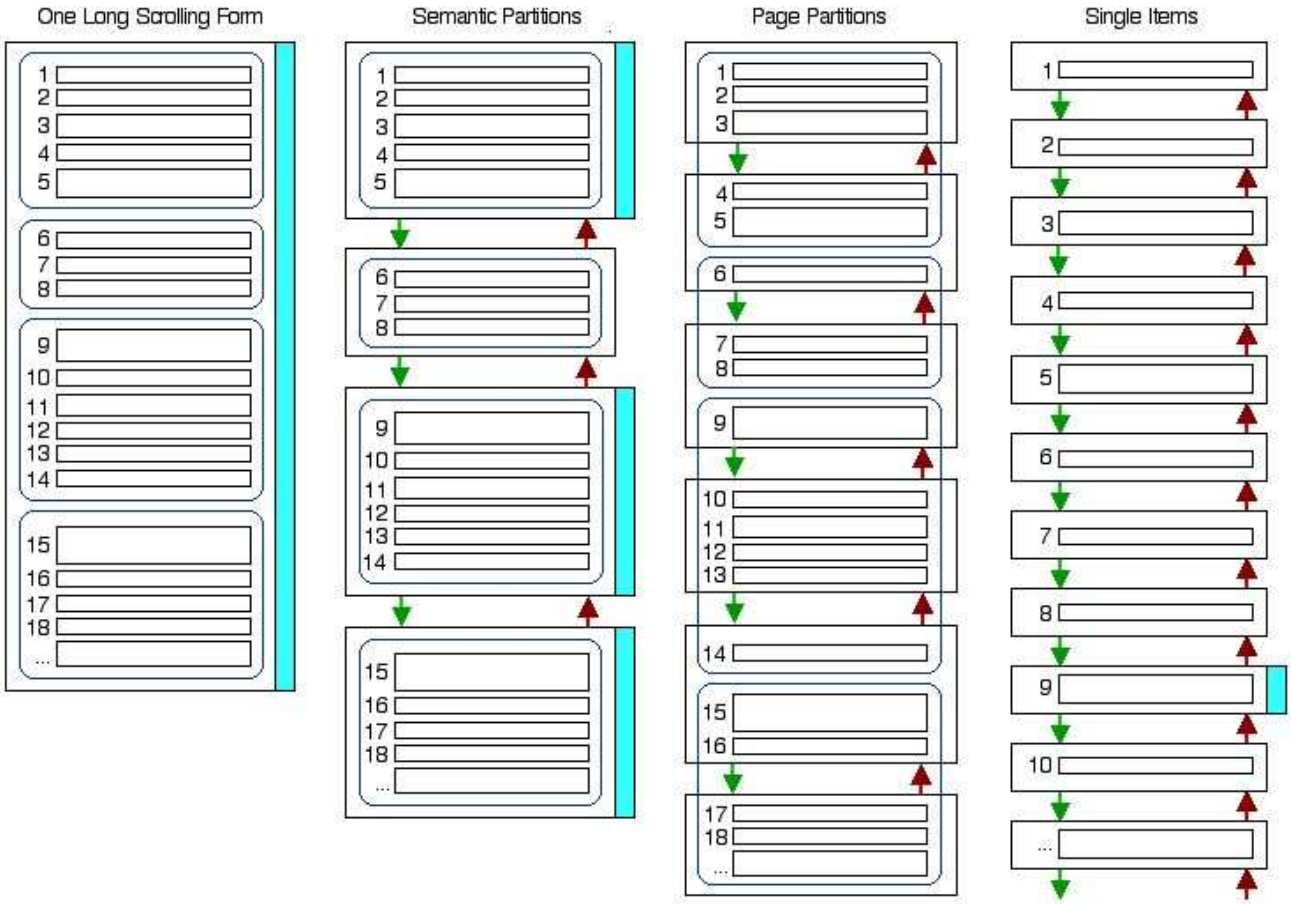

Figure 3. Schematic of a questionnaire presented as one long scrolling form, partitioned into semantic sections, presented as screen sized pages, or displayed as single items.

Scrolling can be used to facilitate movement through sections that are longer than the screen length or window size. Scrolling avoids the problem of arbitrary page breaks in the middle of semantic sections or within the items themselves. The problem with scrolling is that it is generally hard to use as a navigational tool. It requires the user to mouse across the screen to a narrow, hard to hit elevator with small buttons at the top and/or bottom. Research shows that scroll bars with additional paging functions are preferred (Norman, 1995; etc..). such as using the Page Up and Page Down keys.

When surveys are divided into sections, it can help with (a) semantic navigation, (b) thinking about a topic, and/or (c) accessing some type of information database. Semantic sections that fit within the screen size do not need to be further scrolled or paged; but those that don't have the added problems of (a) requiring further navigation and (b) undermining the wholeness of a section by further subdividing it.

Navigation of pages and sections can be sequentially constrained by providing only forward and back buttons; or it can be further enhanced with indexes and even search functions.

Consequently, the survey structure in conjunction with the interface presents a number of pros and cons associated with paging, sectioning, scrolling, and indexing. Much of this is mapped to performance and efficiency issues in terms of the time that it takes to complete questionnaires and the number of mouse clicks required to complete the questionnaire. Confusion in navigation may result in missing information from unanswered questions that were never seen or inaccurate information from misinterpreted questions due to faulty context.

Research and theory over the past 20 years in human/computer interaction has covered a number of issues in screen design and navigation of linear and sequential documents (Shneiderman, 1998). Although much of this work relates directly to the design of CSAQs, there remains a substantial portion of the interface dynamics of CSAQs that have yet to be developed.

Current Study

The current study investigated different ways of partitioning surveys for on-line presentation. A long survey of 76 items was presented in four different ways: (a) as one long, scrollable form, (b) divided into semantic partitions that required scrolling, (c) further sub-divided into screen-sized pages, or (d) presented by single items.

In addition this study investigated the use of an index to sections and/or single items. For each of the above four versions of the questionnaire, an index was added in a frame at the left hand side of the screen. For the long form, semantic partitioning, and screen-sized partitioning versions, the index listed the semantic sections of the questionnaire. For the single item version, the index listed the item numbers.

The four methods of partitioning with or without an index resulted in eight versions of the questionnaire. It should be noted that a number of other design permutations could have been implemented. For example, semantic indexes could have also included the numbers to the items in those sections. This was not done in the present case so that respondents would have to find items by subject matter rather than by number. The single item questionnaire could have listed semantic sections in addition to the item numbers, but since the questionnaire was presented at the item level, semantic sections were not listed within the questionnaire itself.

In the first part of this study participants were asked to first fill out the survey on their own. It was expected that although there were some differences in terms of the number of mouse clicks required to navigate the questionnaires, there would no substantial differences in time to complete the questionnaire relative to the variability of individuals in their time to respond.

Moreover, it was expected that respondents would make little use of the indexes for navigation since they would be filling out the questionnaire in a more or less linear order. It was expected that participants would primarily use down scrolling within long sections and the "Next Section", "Next Page", and "Next Item" links on subdivided questionnaires.

However, in the second part of the study, respondents were asked to enter specific answers for two sets of eight items in the questionnaire. Items in the first set were identified only by the text of the items (e.g., Enter $48 for the amount spent on trash collection for the past three months). Items in the second set were identified by item numbers along with the text (e.g., for Item 54 indicate that you would like to retire at age 65).

This last task is similar to the task of editing incorrect or missing information or the task of starting from an external database and entering information in a nonlinear order into the questionnaire. For the first set of items, it is expected that versions of the questionnaire that have semantic indexes will be faster and more efficient than versions that do not. For the second task, the version with the index of numbered items will be the fastest and most efficient.

Additional data on the navigational tools used and pages accessed should help to define strategies used by respondents and should provide information about how to design better questionnaire structure and navigational tools.

Methods

Materials

A generic questionnaire developed for research purposes by the U.S. Bureau of the Census was further customized for the present study. Items were grouped into seven sections. Four versions of the questionnaire were generated. They were as follows: LF – one, long, scrolling from; SS – semantic sections into six discrete sections; PP – page partitioning with 23 screens; and SI – single items with 76 items. For the long form, respondents had to scroll through the form to see all of the items. In the partitioned forms, a "Next" link was provided at the end of each partition to jump to the next section, page, or item depending on the particular version and a "Previous" link at the beginning of each partition to jump to the previous section, page, or item.

The forms were presented with or without a side index. For the long form (LF) and the semantic partitioned form (SS), the index listed the sections and subsections. The respondent could click on a topic in the index and the form would go to the top of that section. In the case of the single item version, the index listed the item numbers from 1 to 76 without semantic labels. This resulted in 8 versions of the questionnaire. Figure 4 shows the layout for the LF with Index Condition.

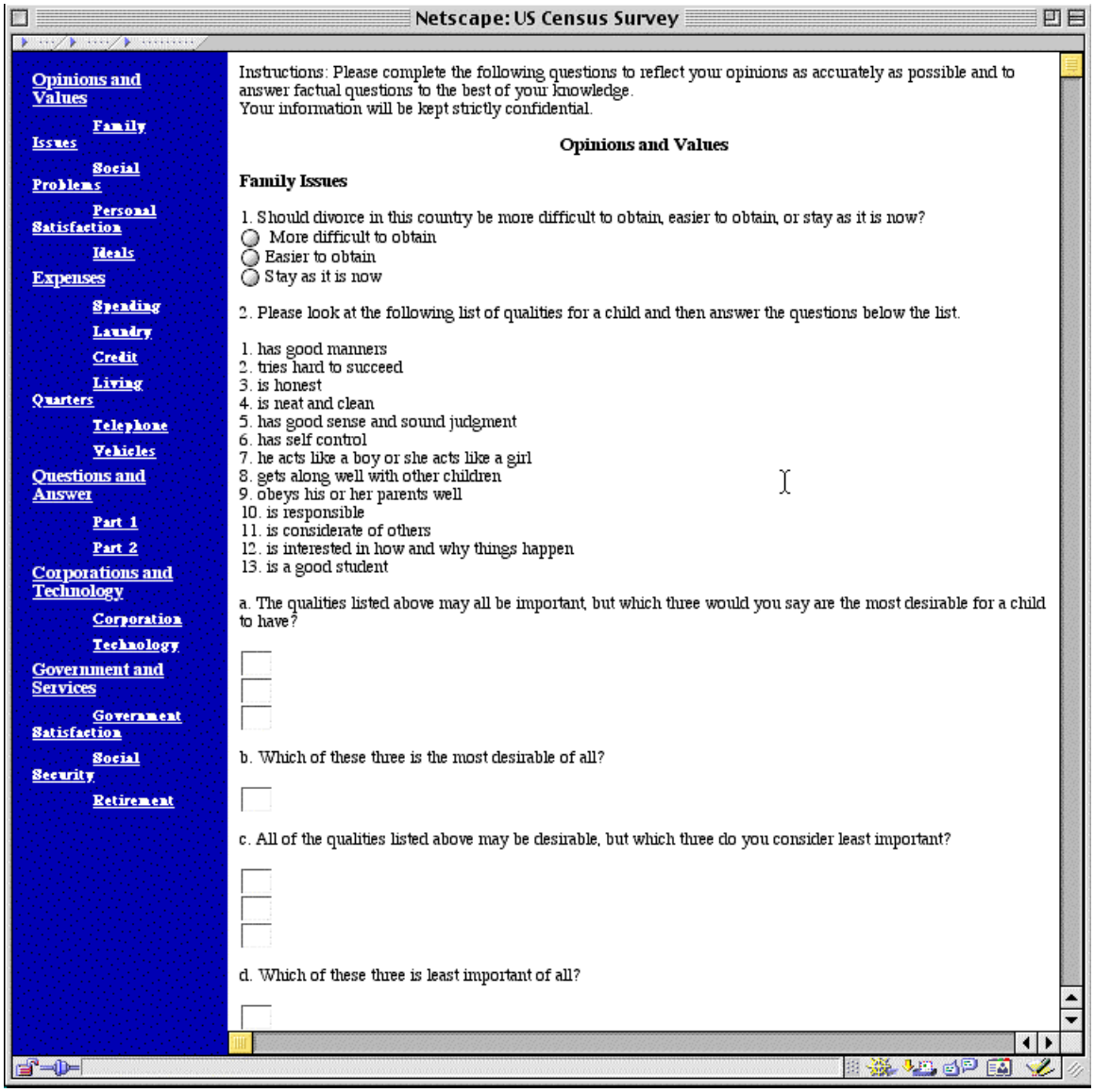

Figure 4. Browser window showing the questionnaire presented as a long form with an index.

Two sets of eight comparable items were selected from representative areas throughout the questionnaire. When the respondent finished filling out the questionnaire on their own, they were asked to enter specific information for these items. For the first set of items (Set T) only textual information was provided to help them find the item. For the second set (Set N), numbers were provided that helped to locate the corresponding items in the questionnaire. These items were listed in the same order that they occurred in the questionnaire.

A follow-up questionnaire was developed to assess the respondents subjective impressions of the interface. Ten items were rated in terms of agreement / disagreement on a 9-point rating scale. (See Table 3).

All materials were generated in HTML and JavaScript. Pages were hosted on a Macintosh™ server running OS 8.5 and

Webstar™ 4.0. Start and stop times were recorded in a separate file from the respondents answers to the questionnaire. Materials were presented to respondents using Netscape™ as the browser on a Macintosh computer with two monitors. The left monitor was a 15 inch display with 640 x 480 pixels. It was used for the presenting, the instructions, the start and stop buttons for timing, and the information to enter for the two sets of items in the editing task. The right monitor was a 17 inch display with 800 x 600 pixels. It was used for presenting the questionnaire. The system was configured such that the mouse moved between two monitors.

Respondents

Respondents for this study were drawn from the university population and were paid $10 for their participation. Subjects ranged in age from 17 to 31 with a mean of 19.92 and standard deviation of 3.01. Seven males and seven females participated in each of the 8 conditions for a total of 140 respondents.

Procedure

Participants were given instructions as to the nature of the task and filled out a consent form on the computer. Then they filled out a short questionnaire on demographics asking their age, gender, and a self-assessment of computer use and expertise. They were then randomly assigned to one of the versions of the questionnaire and were asked to fill out the questionnaire on their own. They were asked to fill it out quickly but accurately and not to skip items unless specifically instructed to do so. When they finished, they clicked on a stop button to record the time. They were then shown information to enter for eight items and asked to go back into the questionnaire to change or enter that information. When finished they were again to click on a stop button. At that point the second set of eight items appeared and they were asked to go back and enter those answers and again click a stop button when finished. Finally, respondents were asked to complete the ten item questionnaire on the interface. When done they were informed as to the nature of the study, allowed to ask questions, and dismissed.

Results

Participants in this study rated themselves as having moderate to high computer skills. On a 9-point scale (1 = no experience, 9 = very experienced) the mean rating of overall experience with computers was 7.01 (s.d. = 1.45); use of the Internet, 6.97 (s.d. = 1.84); and use of the World Wide Web, 6.78 (s.d. =1.84). They also rating themselves as having moderate to high experience in filling out surveys with mean of 6.33 (s.d. = 2.01). No prior significant differences were evident with the exception that participants in the LF No Index Condition tended to rate themselves as having higher computer experience than the other groups.

An index of the accuracy of the information entered was not obtained in the present study, due to the subjective nature of much of the information. However, observations of the respondents and interviews with them indicated that they were conscientious about answering all of the questions.

Table 2 lists the mean time to completion for the eight versions of the questionnaire and Figure 5 displays the results graphically. As expected there were no overall significant differences in terms of time to completion ($F(7, 104) = 1.76$, $p > .05$). However, there was a tendency for the times in LF Index Condition and SS Index Condition to be longer than for the other conditions (Fisher's Protected LSD, $p < .05$). Observations of the pages accessed indicated that respondents tended to fill out the questionnaires in a highly linear manner, scrolling down within the long form, scrolling down within the semantic partitions and advancing to next section, page or item.

Respondents rarely used the index except for testing it out. Indeed the use of the index many have been at times a deterrent for the LF Index Condition and the SS Index Condition.

Table 2
Mean Times to Complete and Edit the Eight Versions of the Questionnaire
(Standard Deviations in Parentheses)

Condition	Completion Time	Edits: Set T	Edits: Set N
LF	1400.57 (462.14)	310.43 (132.91)	201.64 (90.15)
LF with Index	1567.57 (347.03)	328.64 (106.65)	169.14 (34.02)
SS	1263.79 (262.99)	320.71 (131.92)	196.50 (55.05)
SS with Index	1580.64 (320.40)	349.71 (158.55)	190.86 (37.70)
PP	1522.36 (549.32)	338.64 (137.07)	206.86 (74.18)
PP with Index	1428.43 (296.26)	470.21 (146.24)	251.79 (114.46)
SI	1273.71 (360.65)	412.50 (166.79)	226.71 (84.65)
SI with Index	1267.07 (361.81)	434.57 (104.40)	141.93 (31.88)

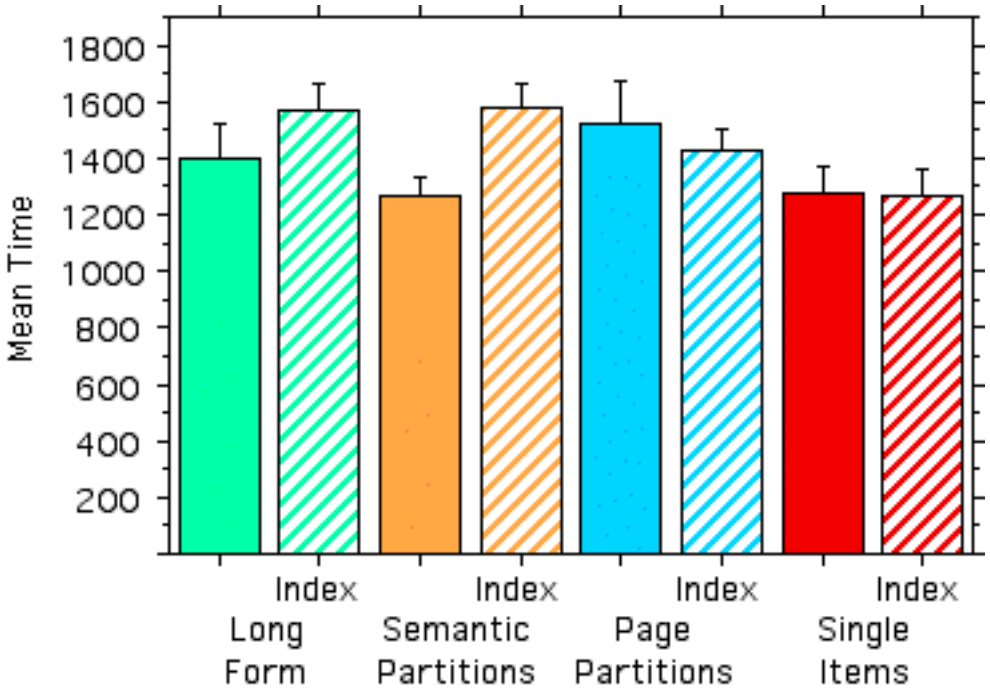

Figure 5. Mean time to complete survey as a function of partitioning and presentation of an index (standard error bars shown).

Table 2 lists the mean times to enter information for the textual set of items (Set T) and Figure 6 displays the results. There were significant differences among the conditions ($F(7, 104) = 2.67$, $p < .02$). The PP with Index Condition and the SI with Index Condition tended to require more time than the LF or SF Conditions. (Fisher's Protected LSD , $p < .05$). It seemed that some respondents who used the index to locate items would get lost in these conditions because the index would not jump directly to the item, but to the beginning of the section. They would then have to advance

through the section or the set of items to locate the right question.

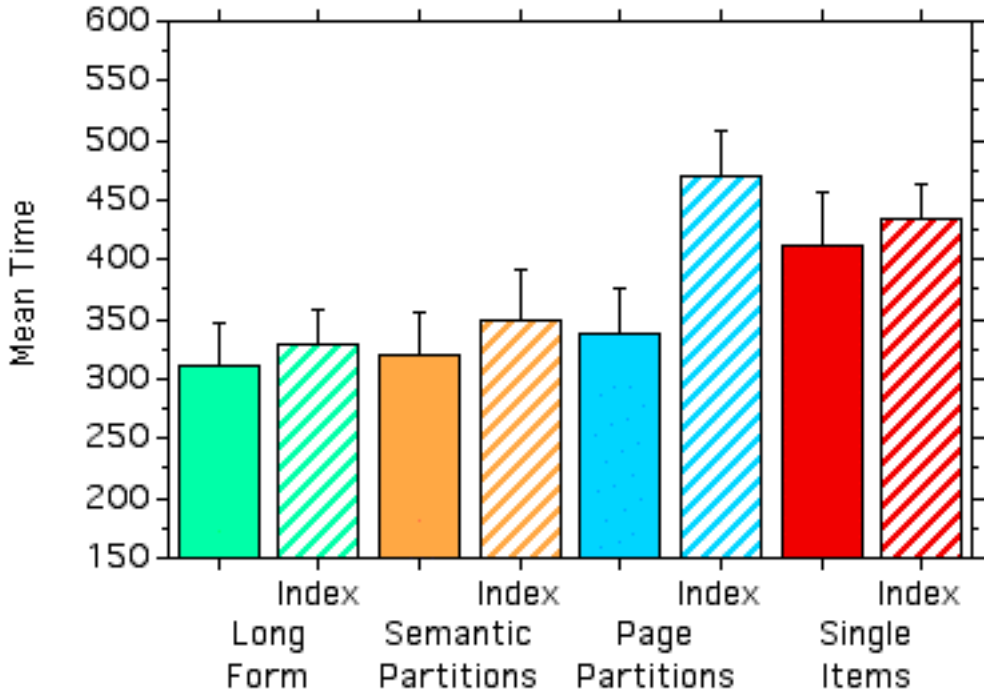

Figure 6. Mean time to complete changes for the semantic set of items as a function of partitioning and presentation of an index (standard error bars shown).

Table 2 lists the mean times to enter information for the numbered set of items (Set N) and Figure 7 displays the results. Again significant differences occurred ($F(7, 104) = 3.10$, $p < .01$). The SI with Index Condition was the fastest (Fisher's Protected LSD, $p < .01$). The obvious advantage was that respondents had merely to click on the item number in the index and the item was presented. All other conditions took two to three times longer to complete. Nevertheless, respondents tended to find the items fairly quickly by scrolling through the long survey or the survey parts. The worst were the PP with Index Condition and the SI Condition (Fisher's Protected LSD, $p < .05$). For the first of these, the index again seemed to work against the respondents by jumping them in a non-linear manner back to the beginnings of sections rather than to the next page or item. For the SI Condition without an index, it merely took more time traversing through single items to get to the next numbered question.

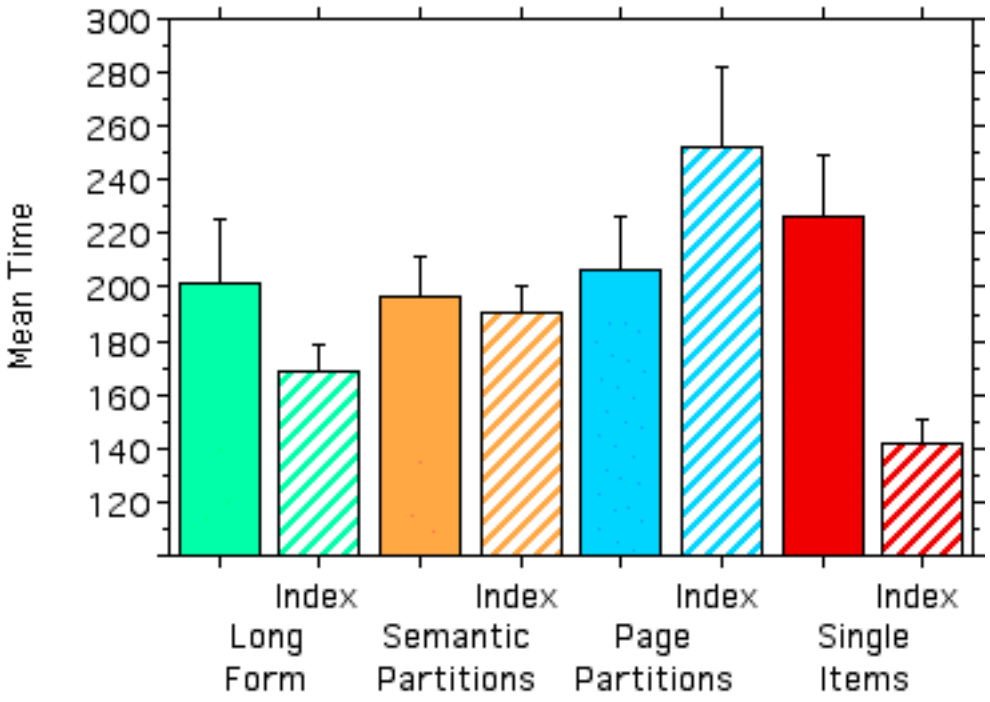

Figure 7. Mean time to complete changes for the numbered set of items as a function of partitioning and presentation of an index (standard error bars shown).

Table 3 lists the mean ratings of agreement with the items on interface assessment survey for each of the eight versions (1 = highly disagree, 9 = highly agree). There were significant differences for the statement "Directions were not detailed enough" ($F7, 104) = 2.15, $p < .05). However, this result seemed unsystematic, as there were no significant differences among the conditions for any other statements.

.

Table 3
Mean Rating of Agreement with Statements (1 = highly disagree, 9 = highly agree)

	Long Form Index		Semantic Sections Index		Page Partition Index		Single Item Index	
1. Eyes became sore	5.36	5.154	4.357	5.786	5.154	4.000	5.857	4.143
2. Easy to make mistakes	5.21	4.692	5.714	5.643	4.538	5.857	4.643	4.500
3. Did not always do what I wanted	3.64	3.571	3.214	3.857	4.857	5.357	4.429	4.214
4. Directions were not detailed enough	4.21	2.571	2.643	3.429	4.077	4.214	4.357	2.857
5. Layout is intuitive.	5.07	5.643	4.643	4.857	5.214	5.000	4.643	5.357
6. Survey is too plain	5.79	4.385	5.000	4.857	4.786	4.786	5.429	3.929
7. Rather fill out survey on paper	3.50	3.929	4.071	4.429	5.214	4.214	3.786	4.214

8. Like if future surveys were submittable online	6.14	4.786	5.714	6.286	5.615	5.571	6.714	6.214
9. Rather have a person ask me these questions	3.86	2.857	2.571	2.857	2.786	3.143	3.357	2.643
10. Feel that submitting census online is easier	7.29	5.500	6.214	7.214	6.214	6.143	6.857	7.214
11. Organization of sections made sense	6.00	5.786	5.786	6.500	5.214	6.143	5.071	6.357
12. Got lost on which question I was on	2.86	3.357	2.500	2.929	3.571	3.214	3.786	2.786

Overall, respondents seemed comfortable with the task and tended to prefer the on-line version of the survey over a paper and pencil version or a personal interview.

Discussion

Interface designs for questionnaires vary in terms of the division of forms into smaller and smaller parts and the navigation methods used to access the parts. At one end are form-based designs that present questionnaires as one long form in a scrollable window. An advantage of the form-based display is that it shows the whole questionnaire. As such it helps to preserve the context of items within the questionnaire and fosters a sense of beginning, linear order, and end of the questionnaire. The disadvantage of the form-based questionnaire is the need for scrolling, which may present problems for some respondents (e.g., confusion, loss of position, etc.). Scrolling, however, did not seem to pose a problem with the current set of computer literate respondents.

At the other end of the design continuum are item-based questionnaires that present only a single item at a time. This design has the advantage of focusing on single questions, but this advantage may be out weighted by the loss of context and the operations required navigating to single items.

The nature of the task of filling out a questionnaire may vary from a straight linear entry of answers to questions to a nonlinear jumping around back and forth from question to question. When entry is linear, the interface should support a smooth transition from one item to the next with minimal action on the part of the respondent. The eight versions of the questionnaire used in this study did not differ substantially in this regard. Forward navigation of items was supported either by scrolling down or clicking on the "Next" link. Consequently, no significant differences were found due to scrolling versus jumping to the next page or the next item in the questionnaire.

However, when the task is nonlinear, the interface needs to support nonlinear access to items so that the respondent does not need to scroll or page through all of the items to get to the required item. In this regard, the results suggest that scrolling one long form may be superior to jumping through many pages in the item-based versions. Respondents could scroll back and forth more effectively than paging with series of clicks on links.

When semantic indexes were provided, they were of little use when respondents were navigating the questionnaire in a purely linear manner. In fact, indexes may have been somewhat of a distraction and a hindrance at times when respondents only needed to go to the next item. In particular, indexes for the

page partitioned questionnaire may have disoriented the respondents when they jumped to the middle of a page where that section started. Numeric indexes were also of little use when entry was linear.

On the other hand, for non-linear access of items when numbers were given, the numeric indexes were highly efficient and task completion times were significantly shorter than all other versions. Item numbers provide a straightforward way to access items in the survey.

Despite the absence of observable increases in performance due to the presence of an index to the semantic sections of the survey, there may be side benefits and clear advantages for other types of tasks. The index may help the respondent see the scope and content of the survey and may be helpful in organizing information retrieval from memory. Other tasks that directly require navigation to sections will clearly be added by the index.

Finally, these results help to guide interface designs in the implementation of on-line questionnaires. They suggest that long scrolling forms are acceptable for at least some users and that an index to sections is not always helpful as one might think.

References

Card, S., Moran, T. P., & Newell A. (1983). *The psychology of human-computer interaction*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory, *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.

Dillman, D. A. (2000). *Mail and Internet surveys (2nd ed.)*, New York, NY: John Wiley & Sons.

Lazar, J. & Preece, J. (1999). Desiging and implementing Web-based surveys. *Journal of computer information systems*, 39, 63 - 67,

Norman, K. L. (1991). *The psychology of menu selection: Designing cognitive control at the human/computer interface.* Norwood, NJ: Ablex Publishing Corporation.

Shneiderman, B. (1998). *Desiging the user interface: Strategies for effective human-computer interaction (3rd ed.)*, Reading, MA: Addison-Wesley.

Acknowledgments

This work was funded in part by a grant from the Statistical Research Division of the U.S. Bureau of the Census (Contract #43YABC917123). We thank Kent Marquis, Beth Nichols, Betty Murphy, and others at the Census for their guidance and direction in this research and Ben Shneiderman and Catherine Plaisant for their input.