BUILDING USABILITY INTO ELECTRONIC DATA-COLLECTION FORMS FOR ECONOMIC CENSUSES AND SURVEYS[1]

Elizabeth D. Murphy, Elizabeth M. Nichols, Amy E. Anderson,
M. Diane Harley, and Kimberly D. Pressley
U.S. Census Bureau, Washington, DC 20233

Abstract

Because the economic census[2] is the primary source of information about the Nation's economy, collecting detailed economic data from business establishments is one of the Census Bureau's key activities. In planning for the 2002 Economic Census, the Census Bureau intends to offer all respondents an electronic alternative to paper forms. Providing an electronic alternative requires translation of over 650 paper formats into Computerized Self-Administered Questionnaires (CSAQs). Because of the differences between paper and the electronic medium, and because of the capabilities made available by the electronic technology, the translation process is not simply a one-to-one mapping. Consideration must be given to the cognitive and usability issues associated with electronic questionnaires to ensure that respondents who choose to report by CSAQ find it to be an effective and efficient means of providing their data. As an aid to developers, the Census Bureau developed a style guide documenting numerous design decisions. This paper describes the process of developing the user-interface style guide with input from cognitive and usability testing. Informed by user-interface design principles, standards, best practices, and the results of low-fidelity-prototype testing, a style guide of this kind is a tool for building usability into electronic forms for data collection.

Keywords: usability, style guide, Computerized Self-Administered Questionnaires, cognitive testing, usability testing

Introduction

The economic census provides a detailed portrait of the Nation's economy once every five years. Data collected in the Economic Census are essential for government, business, and the public[3]. For example, business establishments use Census data to make business decisions and to evaluate their performance. Trade associations and news media study economic indicators and projections based on Census data. Legislators at local, state, and national levels use Census data in the preparation and evaluation of new laws. Consultants and researchers use Census data to analyze changes in industrial structure and location. To meet these many and diverse needs, the data collected must be as complete and accurate as possible.

In previous economic censuses, the Census Bureau has mailed out paper questionnaires to collect economic data from business establishments. In 1997, an electronic form was offered to a selection of large companies in the retail sector. According to plans for 2002, the Census Bureau intends to offer all business respondents the option of reporting their data on paper or by electronic means. The electronic forms are called Computerized Self-Administered Questionnaires (CSAQs). Electronic technology offers the opportunity to streamline the process of collecting data. However, any change in data-collecting techniques must result from a detailed design-and-development process that includes user testing.

There are more than 650 versions of the economic census form, each customized to particular industries. To convert the paper forms into a set of consistent electronic questionnaires, developers need detailed specifications on such topics as font style and layout. These specifications must cover the hundreds of design decisions that would otherwise need to be made "on the fly" by individual developers. A style guide documents design rules to a level of detail that ensures consistency across developers. Since the universe of possible design rules is enormous, the design team must consider and rule out many options before reaching consensus on the content of a style guide.

Development of the 2002 Economic Census Style Guide

The purpose of a style guide is to foster consistency across forms and across the various programmers who may be working on different pieces of a software development project, in this case, a CSAQ. To ensure quality and consistency

in the design of more than 650 electronic forms for the economic census, a cross-divisional team was assembled to develop a style guide for Windows-based, economic CSAQs. Working in a top-down fashion, the team determined that the style guide should include design rules for layout and screen design, navigation, graphics, error handling, help, feedback from users, and other internal feedback mechanisms.

Once the initial set of topics to be covered had been identified and grouped, members of the team signed up to develop various sections of the style guide. The team met weekly to discuss and review the draft content for the various sections. Content evolved gradually as team members, individually and collectively, consulted existing usability principles, user-interface design standards and best practices in creating design rules and sample screens. The project's developer advised the team on alternative approaches that were technically feasible. In some instances, cognitive testing and usability testing were conducted to help the team decide between alternative design rules.

Usability Principles

By definition, principles are stated at a high level of abstraction and offer general, top-down guidance. The style guide team consulted the following kinds of usability principles, which were collected from two major sources of expertise in usability: Nielsen (1993) and Shneiderman (1998):

- Ease of navigation through the user interface: Users should be able to move easily from their point of origin to their destination. They should not have to work their way down menu hierarchies and back up again to get to where they want to be. Users who need to complete forms for multiple units should be able to use shortcuts, such as alternative keyboard commands, to navigate quickly within the user interface.

- Ease of training/learning: The user interface should reflect an understanding of the user's thought process and the logical flow of the user's work. Required training should not be lengthy.

- Consistency of "look and feel" within and across applications: Consistency is the key to building usability into an automated system. Everything from terminology to system feedback should be phrased and located consistently from screen to screen. System behavior should be consistent from one instance to the next. Consistency in look and feel allows the user to build up expectations with the reasonable hope that they will be met. Consistency within and across electronic questionnaires allows the user to build expertise quickly.

Although principles define general criteria for a design, developers need much greater specificity about the attributes of design elements. Thus, the style-guide team consulted more detailed user-interface guidelines and "best practices," e.g., conventions used by other CSAQ-development teams. This continued the top-down approach to identification of topics and style-guide content.

User-Interface Design Guidelines

In the absence of any standards for the design of electronic questionnaires, the team consulted available user-interface design guidelines (e.g., Compuware, 1998, Microsoft, 1995). The team also consulted an internal Census Bureau report on research into user-interface guidelines. The guidelines and the research report helped the team identify topics to cover, provided a general overview of approaches to the design and layout of visual elements, and gave ranges of specific guidance on some design elements.

Best Practices in User-Interface Design

Although the literature on human-computer interaction was the primary source of best practices, the team often referred to design techniques used successfully in other Census Bureau economic CSAQs, such as the Annual Survey of Manufactures (ASM), the Company Organizational Survey (COS)[4], and the Quarterly Financial Report (QFR). For example, the COS uses what we have termed a "stacked design" in formatting the questions on status, payroll, and employment (Figure 1):

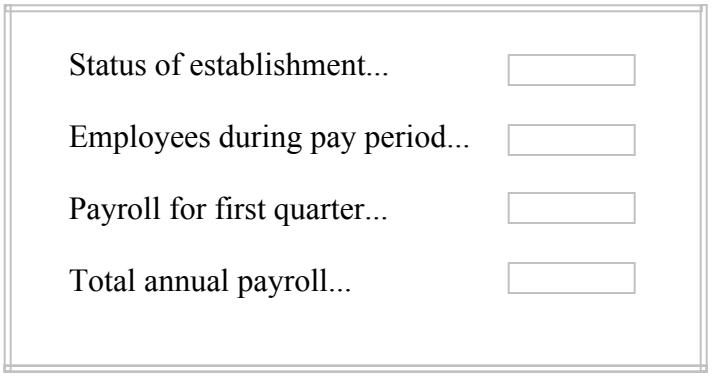

Figure 1. Example of a stacked format as an alternative to horizontal strings of items; recommended to fit all items within the viewable screen space.

Using a stacked format deviates from the horizontal layout used on the paper form. However, this format eliminates the need for the respondent to scroll to the right. Because respondents are likely to miss the subtle cues for horizontal scrolling, they are likely to omit items that are not completely visible within the computer screen's display area. Thus, the stacked format is preferable for the electronic form.

The error-handling mechanism implemented in the QFR is another design element adopted for the 2002 Economic Census Style Guide. In the QFR, an icon consisting of a red circle with a white X was used successfully to identify edit failures (i.e., answers that are out of predefined ranges or inconsistent with prior answers). As in QFR, this icon will appear to the right of an edit-failing data field in the electronic forms for the 2002 Economic Census. Clicking on the icon will open a dialog window containing the edit message, which will provide at least an explanation of the problem. The icon will disappear once the respondent deals with the edit failure.

In developing wording for edit messages, the team referred to wording that had been used for ASM, COS, and QFR edits as well as the 1997 Economic Census. In those CSAQs, some messages were preceded by the word "Warning" or "Error," depending on the severity of the discrepancy. Based on recommendations in the literature, an early team decision for the 2002 Style Guide was not to use the word "error" in any message because of its negative implications (e.g., Shneiderman, 1998). The general feeling was that a respondent's data can be correct even if certain entries technically fail an edit check.

Giving the respondent a Remarks section to explain anything that fails an edit check is expected to help the Census Bureau understand the respondent's rationale for not changing values in response to edit messages. In contrast to the paper form, where respondents often write explanations right beside an entry instead of going to the Remarks section, respondents in the electronic mode can enter explanations only in the Remarks section.

Cognitive and Usability Testing in the Laboratory

To resolve some user-interface design issues, the style-guide team needed further guidance beyond the available literature and best practices. In those cases, alternative user-interface designs were prototyped and submitted to cognitive and usability testing. Results helped to further define the contents of the style guide. For example, members of the team tested proposed wording of edit messages for the electronic forms.

Method: We conducted ten concurrent, cognitive, think-aloud laboratory interviews to test respondent comprehension, task difficulty, and sensitivity of proposed edits for the 2002 Economic Census electronic forms. Using paper prototypes, we also performed preliminary usability testing on the presentation of the edits. Test participants were members of the Census Bureau's economic staff.

We gave participants a partially completed paper questionnaire mimicking an on-line mortgage application. We asked participants to pretend that they were reviewing their own application. Participants read through the questionnaire

aloud.

In order to "fail the edits," some pre-filled answers did not make sense. For example, one question on the mortgage form asked respondents to indicate the percentage of their down payments and closing costs that would come from several types of funds (e.g., savings, investments). The pre-filled data totaled 110%, thus tripping the edit message, "Sum of down payment percentages is 110%. It should be 100%. If values are correct, explain discrepancy in the remarks section." We tested several types of edit messages.

Scattered through the questionnaire were three questions for which participants were to receive preventive edits. That is, the edit message for each of these questions was to be presented immediately instead of later, at the end of the session. Immediately after the participant read each of the three questions designated for preventive edits, the test administrator explained how those questions were to be answered, and the edits were presented on paper. Once participants completed reading the entire questionnaire, we gave them two sheets of paper with twelve edit messages relating to the responses they had just read. The participant's task was to find the item referenced by an edit message, and then to correct the answer to that item or go to the Remarks section and explain why the answer was correct.

In evaluating the edit messages, we needed to determine the importance of and placement of three attributes in the message: item number, item topic, and actual response. When the item topic was present, we varied its capitalization. For example, for several of the variable names, we used the following capitalization styles: TOTAL HOME COST, Homeowner's Fee, and found a new home. That is, variable names were given in all caps, initial caps, or no caps. We found that the following confusing edit message was the most confusing to participants: "Since you reported found a new home in item 15, please complete signed a sales contract in item 16," where "found a new home" and "signed a sales contract" were actually variable names inserted into the sentence. Deciphering this sentence was especially burdensome to respondents.

Results and Recommendations: We developed a set of general recommendations on the content and style of edit messages. These recommendations may apply in other contexts beyond the 2002 Economic Style Guide. The following examples are selected from the full set of recommendations (Nichols, Murphy, & Anderson, 2001):

- **Provide a location, a description of the problem, and an action to take**. Several error messages were written with key pieces of information missing, such as the location of the problem or what action was needed to resolve that particular error. Respondents reacted by spending extra time determining the location of the error and interpreting their course of action. Participants commented that working with the error messages would have been easier if these key pieces of information had been included. When the location of the problematic question was not readily available, respondents used other pieces of information within the edit to help give them "clues" as to where the problem was. The most readily used clues were subject-matter content that was inserted into the message, or the actual amounts reported in answer to the troublesome questions.

- **Beware of mismatches between language used in the messages and the questions**. The subject-matter content that was provided in the error messages was intended to match back to the content in the actual questions. For testing purposes, the subject-matter content that was provided in the error messages was slightly altered to see how participants would react. When respondents encountered this mismatch, they reported "thinking twice" about whether they were working with the correct question. This mismatch was another obstacle for the respondent.

- **Avoid jargon**. Avoiding the use of jargon is a standard usability recommendation when designing a computer application. This rule also extends to the wording that is used in edit messages. When presented with jargon, respondents noted that they were unfamiliar with its meaning.

- **Be polite**. Instead of phrasing the messages as commands, the word "please" was used to introduce instructions for respondents. It was intended to convey a pleasant tone and to reduce any negative perception of the Census Bureau as an imperative authority.

- **Be sensitive to the respondent**. Certain words that are synonymous with error messages can be very judgmental and imply that the user is wrong. Words typically associated with error messages that should be avoided are "error", "discrepancy", and "mistake."

- **Beware of offering only one of many possible solutions**. Several of the messages used in testing included reminders such as "please check for typing mistakes." These messages were meant to help the respondent figure out what could have gone wrong. Participants spent more time than necessary focusing on these reminders. The reminders seemed to be another hurdle that the respondents encountered in working with error messages. Reminders often led respondents down the garden path and away from thinking about the real problem.

- **Capitalize variable names.** Variable names referring to subject-matter content were used in many of the error messages. To the respondent, the variable name was a very important part of the error message, especially when it was the only piece of information given in the message. Participants noted that they preferred seeing the variable name presented in all uppercase because that style made this key piece of information stand out more.

- **Keep messages brief**. Participants noted that lengthy messages were more difficult to deal with than shorter messages that seemed to "get to the point" quicker. It is important to ensure that messages contain all of the necessary pieces of information in a succinct way.

- **Use good grammar**. Participants stumbled when awkward grammar was used in a few error messages. The difficulties resulting from insertion of variable names that did not "read" in context suggest that variable names encoded in the metadata need to make sense in any context in which they are inserted for the respondent.

- **Present edit messages as soon as possible**. Several respondents felt added burden when presented with their mistakes at the very end rather than at the point where the mistake had occurred. They would rather have dealt with any errors that occurred at the time and place where they occurred.

- **Format data-entry fields to prevent errors**. Participants noted that an edit could have been very easily avoided if the format that the application required for an answer was presented as part of the actual question. This convention is widely used in several computer applications and on web pages on the Internet.

Based on test participants' performance and verbal feedback, we also developed recommendations for wording the various kinds of edit messages. The following table provides examples of the wording that was tested and the generic re-wording recommended after testing (Nichols, Murphy, & Anderson, 2001):

Table 1. Examples of Tested Wording and Recommended Wording

Tested Wording	Recommended Wording
Homeowner's Fee is $1,200. Typical values are between $25 and $300. Please check for typing mistakes. If entries are correct, explain discrepancy in the remarks section.	In Item X, typical values are between A and B. You reported C. If this value is correct, please explain in the remarks section.
Since you reported found a new home in item 15, please complete signed a sales contract in item 16.	Please complete Item X, since you completed Item Y.
Sum of down payment percentages is 110%. It should be 100%. If values are correct, explain discrepancy in the remarks section.	In Item X, the values you reported add up to A%. The total should be 100%.

The recommended, generic wording for the various types of edit messages is documented in the style guide.

Review of Form Prototypes

As prototypes were developed for the paper versions of the economic forms, a programmer built corresponding prototypes of the electronic form. These electronic prototypes were revised according to style-guide team consensus. Decisions made during team discussions were documented for inclusion in the style guide. In some cases, situations arising from the prototypes called for topics to be added to the style guide (e.g., retention of column headers, indentation, presentation of running totals). This bottom-up approach complemented the top-down approach taken earlier in the evolution of the style guide.

As necessary, the team called in subject-matter experts from the various economic areas (e.g., Retail, Manufacturing) to assist in making decisions about necessary departures from the design of the paper forms. For example, to avoid horizontal scrolling, the grid design of the paper form for the manufacturing survey needed to be streamlined to fit the limited width of the computer screen. A representative from the manufacturing subject-matter area attended style-guide team meetings and participated in developing several options that departed from the corresponding paper form. One option was to use the stacked design that has been used successfully by the COS (Figure 1). Other manufacturing subject-matter experts agreed to this design change, and their representative brought this decision back to the style-guide team. Rather than risk loss of data, the electronic form was re-designed to present all the items within the horizontal display area of the screen.

Cognitive and Usability Testing in the Field

In the second round of user testing, a partially functioning electronic prototype of the retail form was administered to participants at several companies in the Washington, DC Metropolitan Area. Accompanied by one or two other team members, a survey researcher visited business locations and administered the prototype questionnaire to participants. The prototype CSAQ was administered to participants using a laptop from the Census Bureau, i.e., it was not loaded onto the participants' own personal computers (PCs) at their worksites. During each test session, the laptop's screen was videotaped to record the participant's actual navigational path through the prototype CSAQ.

Researchers wanted to determine whether the electronic format would support the respondent in comprehending what data to provide and where to enter data. Of interest was the respondent's use of navigational features. As shown in Figure 2, three navigational features were embedded in the prototype: a set of menu options in the left navigation bar, “next” and “previous” navigational buttons, and hyperlinks for skip instructions and data verification.

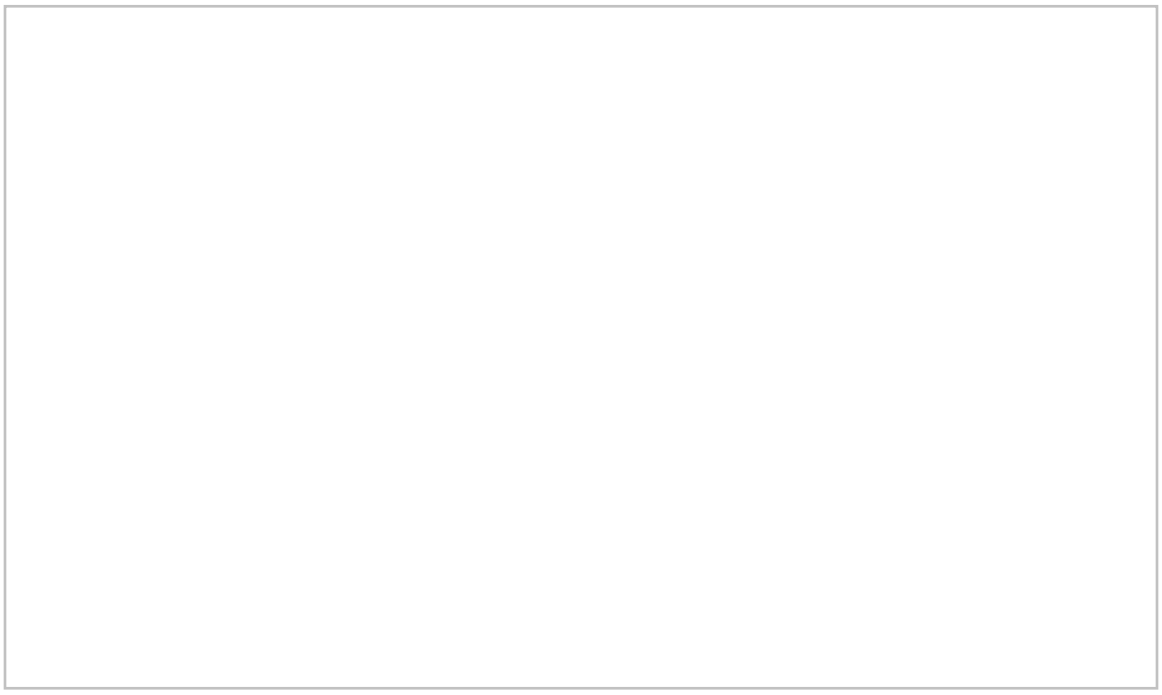

Figure 2. Prototype of electronic form for the 2002 Economic Census showing a hyperlinked skip instruction (...go to 29), hyperlinked menu options in the left navigation bar, and navigational buttons at the lower right.

Participants in the testing used all three of these navigational features appropriately. As expected, the hyperlinked skip instructions were the least likely to be used, and the next and previous buttons were most likely to be used.

Another issue of interest in testing was the method provided for respondents to select answers. Experience in QFR usability testing had indicated that combining radio-button functionality with tabbing can have serious, undesired consequences. For example, tabbing from one question to the next can cause the respondent's selected radio button to jump from "Yes" to "No" without the respondent noticing this change. Thus, for "select one" choices, the style-guide team is considering checkboxes instead of radio buttons.

The difficulty with checkboxes is that multiple responses can be selected. In test situations, participants have quickly discovered that they can select more than one option. As of this writing, all test participants have noticed the instruction to select only one response and have assumed initially that the computer will allow them to check only one. To change a response, however, participants have found that they need to perform the additional step of unchecking a box to proceed. One participant assumed that this "feature" would be corrected in the final form. The team has documented the desired behavior of checkboxes and other user-interface elements in the style guide.

Summary

Development of the 2000 Economic Census Style Guide drew upon several different but related methods. Consulting usability principles and user-interface design guidelines got the team started in a top-down direction, but it was necessary to test alternatives and use bottom-up approaches in formulating design rules at the required level of specificity. Past results (what worked and what did not work for participants) from usability testing of other CSAQ's was valuable in informing team decisions. Cognitive and usability testing were useful in providing human performance data and observations on which to base decisions. Once the style guide is complete, for designers will use it in conjunction with forms-design software to design the CSAQs for the 2002 Economic Census. Revision of the Style Guide for future use can occur as necessary now that the foundation has been laid.

References

Compuware, Inc. (1998). *GUI Guidelines* (Help template). Madison, WI: Author.

Microsoft, Inc. (1995). *The Windows Interface Guidelines for Software Design*. Redmond, WA: Author.

Nichols, E., Murphy, B., and Anderson, A. E. (2001). *Results from Cognitive and Usability Testing of Edit Messages for the 2002 Economic Census* (Human-Computer Interaction Memorandum Series #39). Washington, DC: U. S. Census Bureau.

Nielsen, J. (1993). *Usability Engineering*. New York: AP Professional/Academic Press.

Shneiderman, B. (1998). *Designing the User Interface* (3rd ed.). Reading, MA: Addison-Wesley.

---

[1] This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The authors are grateful to Kent Marquis, Erica Olmsted, and Diane Willimack for their constructive reviews.

[2] The economic census is mandated by law under Title 13 of the United States Code (sections 131, 191, and 224).

[3] http://www.census.gov/epcd/www/econ97.html

[4] Officially known as the Report of Organization Survey.