Designing Edits for Electronic Economic Surveys and Censuses: Issues and Guidelines

Amy E. Anderson, Elizabeth D. Murphy, Elizabeth M. Nichols, Richard S. Sigman, Diane K. Willimack
U.S. Census Bureau, Washington, D.C. 20233

Key Words: Electronic reporting, establishment, usability testing, computerized self-administered questionnaires (CSAQs), data editing, electronic data collection.

1. Introduction

The purpose of this paper is to document and to reflect on the U.S. Census Bureau's experience with interactive data-editing strategies used in collecting data from business survey respondents. Such surveys are subject to many classes of unintentional human error, such as programming mistakes, omissions, miscalculations, keypunch errors, and interviewer misclassifications. The contribution of each class of inadvertent error to total survey error, however, may be controlled or at least somewhat mitigated by good survey design practice, forethought and planning, and by the advances of computer technology. In the survey context, all such efforts to detect and correct these errors fall under the purview of *data editing*.

Data editing may occur at almost any phase of data collection or analysis. Editing activities range from correcting typographical errors or out-of-range entries to elaborate statistical checks performed by computers identifying misshapen blocks of aggregate data. In longitudinal surveys, comparisons are made to previous results. In all cases, the goal is to identify and correct as much respondent error as possible.

In mail surveys of establishments, data editing has typically been performed during post-data-collection processing. Computer-assisted establishment surveys, however, may perform data editing interactively, during data collection. For example, in surveys that use computer-assisted telephone interviewing (CATI) or computer-assisted personal interviewing (CAPI), data editing rules, referred to as *edit checks*, can be incorporated into the CATI/CAPI instrument so that the interviewer is notified of a response that fails one or more edit checks. The interviewer can then probe the respondent for an alternative response or for verification that the flagged response is correct.

Edit checks are also incorporated into computerized self-administered questionnaires (CSAQs), which are delivered to the respondent by mailing disks or CD-ROMs (downloadable) or transmitted electronically over the Internet's World Wide Web. Browser-based CSAQs are also called *online Web-survey* or *Internet* questionnaires. When collecting data using CSAQs, respondents – not interviewers – are notified when a response fails one or more edit checks. The remainder of this paper focuses on interactive edit checks in downloadable and Internet questionnaires.

For comparison purposes, Section 2 briefly describes traditional automated, post-collection data-editing strategies. Section 3 describes the interactive editing approach currently incorporated into CSAQs collecting economic data from businesses. In Section 4, we offer preliminary guidelines for incorporating edit checks into CSAQs, based on several usability studies of electronic surveys of businesses or organizations. In Section 5, we discuss our findings and propose themes that emerge from the Census Bureau's approaches to interactive data editing; and we raise some issues for future research in Section 6.

2. Traditional Post-Collection Establishment Survey-Editing Practices

The Census Bureau has developed two generalized processing systems for editing economic data. One system is used for the economic censuses conducted every five years, and the second system is used for current economic surveys conducted more frequently. The editing system for the economic census, called "Plain Vanilla," provides basic editing capabilities, which can be augmented, if necessary, by trade-area-specific computer code. The editing system for current economic surveys is a module of the Standard Economic Processing System (StEPS), which performs a number of post-data-collection processing functions. Plain Vanilla is a highly automated system for the large data volumes associated with economic censuses, while StEPS is a very flexible and easily configured system for the large number and wide variety of current economic surveys.

Plain Vanilla has modules for performing several basic edit checks and a generalized module for validating discrete data items against item-specific reference lists. Subject-matter experts monitor the editing process and, if necessary, make adjustments to edit parameters or override edit actions. For more information about Plain Vanilla, see Wagner (2000).

Using the StEPS editing module for current economic surveys, subject-matter analysts can interactively define the following kinds of edits:

- Required-item edit: Verifies that a specified item has been reported

- Range edit: Verifies that item value lies in the range defined by a specified minimum and maximum value
- List-directed test: Verifies that a predefined list of values contains the value of the specified item
- Balance test: Verifies that a sum of specified detail items is equal to a specified total
- Survey-rule test: Validates complex inter-item relationships, such as inter-item ratios, ratios of data reported currently to data reported in a prior period, and other logical relationships between items reported by the same sample unit
- Negative test: Verifies that the value of the specified item is not negative

Interactive edits allow analysts to make corrections and immediately see whether the corrected data satisfy the edits. Executing edits in batch creates an edit reject file, which can be reviewed interactively by analysts or passed to the automated general imputation module which identifies items to be machine imputed. For more information about the StEPS editing module, see Sigman (2001).

3. Experiences with Edit Checks and Implementation Strategies for Electronic Economic Surveys and Censuses

CSAQ edit checks prompt respondents to clarify or resolve ambiguous or discrepant responses. Fields containing information that respondents can edit, including name, address and contact information, may be subject to intereactive edit checks. Thus, the kinds of edit checks incorporated into the Census Bureau's economic CSAQs cover a broader range of potential discrepancies than do those conducted during post-collection.

Economic CSAQs borrowed the following kinds of edit checks from post-collection edit routines:

- Balance edit: Verifies that the sum of detail data equals the appropriate total.
- Survey-rule test/Ratio edit: Verifies that the ratio of two data values lies in the range defined by a specified minimum and maximum value.
- Survey-rule test/Logical edit: Verifies that data reported for related items are logically valid.
- Required item or Missing value/Incomplete edit: Verifies that data have been reported.

The following kinds of edit checks tend to be administered only within CSAQs:

- Preventive edit: Blocks respondents from completing an action, occurring upon the first invalid keystroke.
- Alphanumeric edit: Verifies that the data meet the proper alphanumeric rules established for that field.
- Format edit: Verifies that the data have been entered in the expected form (e.g., date entered with dashes instead of slashes).
- Rounding Test: Checks to see if rounding has occurred in the data. (Some post-collection balance edit checks may use rounding tests.)

CSAQ designers have control over when and how the results of various edit checks are displayed to respondents. *Immediate edit checks* present edit messages instantly upon detection of discrepant data, and the system prompts the respondent for a correction or explanation. The result of an immediate edit check can be presented as either a pop-up window or an icon located near the questionable item. The results of *deferred edit checks* are presented to the respondent after the data have been entered and reviewed, usually in a list format. *Server-side edit checks* employ a type of deferred approach, since data have to be submitted to the server in order to generate the edit message.

Two questions are frequently asked in survey development: 1) How many edit checks are too many? 2) Should we allow submission of data with edit failures remaining? It is difficult to devise empirical rules to answer these two questions since each data collection situation is different. Instead we can speak to what we have done, what seems to work, and a general philosophy we have adopted.

Table 1 summarizes the editing features of six Census Bureau economic programs offering either downloadable or browser-based CSAQs – the Survey of Industrial Research and Development (R&D), the Manufacturer's, Shipments, Inventories, and Orders Survey (M3), the 2002 Economic Census, the Company Organization Survey (COS), the Annual Survey of Manufactures (ASM), and the Quarterly Financial Reports (QFR).

Table 1 shows that economic programs at the Census Bureau have embedded numerous edit checks into each electronic survey and, in two cases, that the number of edit checks exceeds the number of items collected (range = 0.31 to 2.43). Although the number of edits could be related to respondent burden, our experience indicates that respondents are generally receptive to edit checks. In situations with numerous edit checks, respondents could be bombarded with edit failures if they happen to trigger each edit rule. Fortunately this does not typically happen.

The question survey designers must address is the likelihood of a respondent triggering an edit check. If a respondent is highly likely to trigger a large number of edit checks, then perhaps the number of edits embedded into the CSAQ should be reduced. If it is likely that a large number of edit checks will be triggered, however, other questions to ask are whether the edit rules are too strict or whether there is a problem in the question phrasing or response field placement causing respondents' answers to trigger multiple edit checks. From the respondent's perspective, the purpose of edit checks is to help the establishment submit accurate, internally consistent data. Edit checks that do not clearly foster this result may be annoying to respondents.

Table 1 also shows that the Census Bureau's economic programs typically do not require respondents to resolve all edit failures before submission. For the situations where certain fields are required to pass the edit test before the survey can be submitted (also known as "hard" edits), failing to satisfy the edit test results in unit nonresponse. These items are considered so critical to survey response that missing or inaccurate responses make the submitted form unusable.

The philosophy of accepting data with unresolved edit failures stems from two principles: 1) Let the user be in control to the extent possible (a usability principle); and 2) Obtaining some data, even edit-failing data, is better than obtaining no data at all. Since the first principle is respondent-centered and the second is from the survey management perspective, CSAQ editing strategies become a balancing act. Generally, however, respondents want to provide accurate data; thus, if they have not changed data in a way to satisfy the edit failure, we should assume their response is correct in their business context. We also suspect that the more difficult it is to submit data electronically, the more likely is unit nonresponse.

4. Preliminary Guidelines from Usability Research on Organizational Surveys

4.1 Census Bureau Usability Research Methodology for Organizational Surveys

To learn about respondent interaction with various edit behaviors[1], the Census Bureau tests candidate editing approaches with respondents. We observe test respondents interacting with edit messages during usability testing. Do respondents recognize edit-failure notifications? Do respondents read the edit messages? If they read the messages, do they understand them? What kind of action do they take regarding the message? A response might consist of ignoring the edit check, modifying data values, or writing a remark to explain why data are accurate even though they failed the edit check. The latter task is particularly characteristic of business surveys since it is not uncommon for valid data to lie outside an expected range. Finally, how easy is it for respondents to interact with the CSAQ to respond to the edit check? For example, respondents may have to navigate to items referred to in the edit message.

Because it is virtually impossible to recruit business respondents to travel to the Census Bureau's usability lab, researchers travel with a video camera and laptop to the business locations to conduct usability testing. With the respondent's consent, video tape recording allows one or more researchers to analyze the session afterwards; the laptop is a necessary backup in case the CSAQ does not work properly on the respondent's workstation. Using a think-aloud protocol, a method often used in cognitive testing, we watch and listen as respondents interact with the CSAQ in their offices. Often we use internal staff members as supplemental subjects, since the usability literature typically recommends between five to 12 subjects (e.g., Dumas and Redish, 1999). As few as five subjects is considered sufficient to identify 80 percent of the usability issues (Virzi, 1992). Even with a small number of participants, usability methods are effective in uncovering issues in comprehension and navigation of the CSAQ.

Usability testing has its limits, however. Since usability testing uses a small number of subjects, generally from convenience samples, results cannot be tested for statistical significance. Statistical hypothesis testing is not appropriate in a usability-testing context because usability testing is not controlled experimentation. Usability testing is intended to identify problems that actual respondents might have with the CSAQ, not to find significant differences between groups.

Further, interview time devoted solely to edit checks is limited since usability testing focuses on the entire instrument. Edit behavior is often not fully functional when usability testing is conducted; such was the case during testing of the 2002 Economic Census prototype. Recognizing these disclaimers, the best practices we recommend for designing edit behavior arise from limited usability testing and should be subjected to additional research.

[1] "Behavior" here refers to the CSAQ's methods of communicating edit failures to the respondent: What is communicated, how is it communicated, and when?

4.2 Preliminary Guidelines

The following design guidelines summarize our interpretations of respondent reactions to interactive edits encountered during usability tests. After we state the guideline, we include evidence from our usability tests to support the guideline. In addition to the usability tests performed on surveys listed in Table 1, we also drew from usability reports for two institutional surveys: the Library Media Center Survey (LMC) and the Private School Survey (PSS). At the time of those tests, the LMC and PSS surveys were browser-based.

1. Good questionnaire design minimizes edit failures. Good questionnaire design includes communicating to respondents which data fields require answers and which are optional. For example, instructions should inform participants to click on "none" if the answer is zero or to enter a number when an entry is required. For dates or amounts, format can be built into fields automatically. Additionally, question text can include instructions on the correct format (Nichols et al., 2001).

2. Perform edit checks immediately unless checking for missing data or performing an inter-item edit. Defer activating those edit checks. Run them either immediately before the respondent finishes the form or after all the items in the inter-item edit have been entered. Study participants preferred immediate notification of edit failures, rather than receiving a list of edit messages at the end. Participants can learn to avoid similar mistakes if they are notified immediately. However, we caution against triggering edit rules too early. This happened during usability testing of the Quarterly Financial Report (QFR). A QFR edit checking the consistency between two items (an inter-item edit check) triggered the edit as soon as data were entered for the first of the two items. Participants thought the edit check was ill placed. This edit check should have been invoked on the second of the two items. We recommend activating the edit check when all the relevant fields have been manipulated, no matter the order of manipulation.

3. Research edit checks before implementing an edit that might be too strict. Some participants during ASM usability testing did not think the edit took in all the relevant factors when calculating ranges for a ratio edit. In usability testing of the survey of library media centers (LMC), many participants appeared frustrated when they correctly entered a number, which exceeded a range edit check. Based on the number of participants who triggered this edit check during usability testing, we determined this range edit was too strict.

4. Mixing editing with other functions, such as submitting, violates user expectations. Problems arose in the Private Schools Survey (PSS) because the server-side editing process was invoked when the respondent pressed the "Finished" button. At this point, respondents thought that the form would be submitted. When they saw that an edit check was run and edit messages appeared, they changed their understanding of the "Finished" button. They then believed that clicking on "Finished" again would iteratively check for edits until their forms were correct. This was not the case. Edit checks were only run the first time the "Finished" button was pressed. Although this design was most likely created to ensure respondents invoked the edit checks, the design violated respondents' understanding twice. Initially it violated their understanding of the word "Finished." It then violated their expectation of the ability to iteratively check for edits. During usability testing of the Company Organization Survey (COS), respondents were also surprised that the edit report was rerun when they tried to submit.

5. Allow edit failure reports to be run iteratively, as expected by respondents. In some CSAQs, edit failures for the entire CSAQ can be run as a batch (usually at the end of the questionnaire) and presented together as a list of failures. The batch process of presenting edit messages was not a problem in itself. The problem arose if the CSAQ did not allow this batch processing to be rerun and updated. For example, the Private School Survey (PSS) was designed for all the edit messages to appear together, at the top of the scrollable form, once the form was submitted. Designers likely expected respondents to make their way through the list, correcting each one in turn. During usability testing, however, some participants wanted to recheck their form after correcting only one edit, hoping that the list would reappear without the edit they had just corrected, but they were not permitted to do this. In the Annual Survey of Manufactures (ASM), we also observed respondents wanting to return to the review screen after correcting a failure. In this case, each time the review screen was invoked, the edit checks were rerun, generating

an updated list. This design met respondent expectations.

6. Allow for easy navigation back and forth between an edit failure list and the associated items. In both the COS and ASM, respondents easily navigated from the list of edit failures on the review screen to an item by clicking on the hyper-linked edit-failure text. Once at an item, however, returning to the review screen was confusing. Users of the private school CSAQ (PSS) wanted to be able to navigate easily back to the list of edit failures, once they were at an item. When they discovered the list was at the top of the form, they complained about having to scroll up to see the list. In the 2000 M3 CSAQ design, server-side edit checks were run, and the edit messages appeared on a separate page. Users had to exit that page to return to the form and correct their responses. They could not easily navigate between the two pages. Usability experts working on the M3 recommended placing the messages directly on the form, eliminating the need to navigate between two windows.

7. Clearly identify which edit failure goes with which item. In the PSS CSAQ, clicking on the edit-failure message at the top of the scrollable form caused the page to reset and display the line containing the edit-failing data-entry field at the top. However, the question text for this field was "above the fold" – that is, it was off-screen. To see the question text, respondents had to scroll up, thus they had to perform two tasks to find the item associated with the edit failure. In the CSAQ for library media centers (LMC), the pop-up edit messages were invoked when the respondent's cursor entered and gained focus in another data field. If respondents scrolled down this browser-based scrollable form, the item with the failed edit could be off the screen when the pop-up message appeared.

8. Edit messages should contain a location, a description of the problem, and an action to take. Respondents were always trying to decipher where the edit-failing item was located and what they needed to do to fix it. Every participant for the 2002 Economic Census prototype testing commented that many of the edit messages would have been easier to work with had the item number been available. Participants also wanted to know what action they needed to take to resolve an edit failure. The easiest messages to understand were those that said, "Please complete item *X*" (Nichols et al., 2001).

9. Edit messages should avoid jargon, be polite, use good grammar, be brief, use active voice, and use words that match the terminology used in the question. Problems arose when words used in the edit message did not match the terminology used in the item. Respondents were not sure they were on the right question (Nichols et al., 2001). Unclear edit messages were also a problem during usability testing for the LMC field test (Tedesco et al., 1999).

10. Avoid edit messages that offer only one solution when many are possible, and cognitively test them prior to implementation to assure interpretation as intended. Problems arose when solutions such as, "Check for a typing mistake" were contained in an edit message. Sometimes these solutions led respondents astray (Nichols et al., 2001). We noticed participants changing their answer to fit the upper bound of a range edit check in the LMC CSAQ when the range edit provided the bounds.

11. Do not automatically erase data that fail an edit check. In the LMC testing, we tested messages containing an "OK" and a "Cancel" button to close the pop-up edit message window. The edit messages warned participants that their data would be erased by clicking the "OK" button, while clicking "Cancel" retained respondent entries. Some participants did not understand the difference between the two buttons. We found that when participants' entries were erased by clicking the "OK" button, some were reluctant to re-enter data (Tedesco et al., 1999).

12. Inform respondents about the policy for submitting data with unresolved edit failures. CSAQ respondents in the testing of both the Annual Survey of Manufactures (ASM) and the Quarterly Financial Report (QFR) were not sure whether they could send data with errors remaining, although this was permissible.

13. Give the respondent as much control as possible. The user is not in control of pop-up edit failure messages, which appear automatically. Unsolicited pop-up windows containing edit messages caused problems for respondents in usability testing for the LMC Field Test. Several respondents did not read the message but automatically clicked a button in the window to make it disappear. When probed at the end of a testing session, one respondent didn't even

remember a pop-up window. Others thought it was a computer bug. The QFR also used pop-up windows to display the edit message, but the participant needed to click on an icon next to the edit-failing field to invoke the pop-up. Participants used this icon successfully and were in control of the timing of the edit message.

14. Use newly created icons with caution since they do not have universal meanings. Use standard icons only for expected purposes. The QFR and the ASM use icons to immediately notify respondents of edit failures. A red circle with a white "X" icon was used successfully by QFR respondents. When they clicked on the icon, an edit message displayed. The yellow "!" warning messages were rarely clicked on in the QFR testing, and a few ASM respondents were unaware of this property. Respondents thought the white bubble "i" icon, which was used only in the QFR, meant additional information, and they were surprised to find the message reminded them to make a remark. The standard use for an "i" bubble is to convey information, not to suggest a user action. Violation of the standard meaning of an icon confuses users.

5. Discussion and Emerging Themes

We summarize by discussing several themes that emerge from Census Bureau survey practices and research on incorporating editing into CSAQs and Web instruments:

The use of edit checks has increased for several reasons over the years. Historically, early CSAQs incorporated only a few basic edit checks because of a grave concern for additional respondent burden, which might result in unit nonresponse. In addition, early software could support only a few simple edit checks. Over time, more edit checks have been added to existing CSAQs and to newly created CSAQs. Indeed, the ratios of the number of edit checks to the number of questionnaire items presented in Table 1 seem high: A recent Web instrument developed by the Census Bureau, the M3, averages more than two edit checks per item on the questionnaire. This growth has occurred, in part, because of enhancements to the software, enabling the creation of edit checks that were not previously possible. Moreover, the number of edit checks has increased as survey staff experience and confidence have grown over multiple survey iterations.

A reasonable number of embedded CSAQ edits will not necessarily increase respondent burden or lead to unit nonresponse. Even though the number of edit checks has increased, it appears that embedded CSAQ edits do not necessarily lead to unit nonresponse. This is corroborated by usability research suggesting that respondents seem to appreciate edit checks, wanting to be informed of discrepancies in their data so they can make corrections. Thus, respondents do not necessarily consider edit checks to be "bad", and they do not appear to abandon the response task just because they received edit messages. In our experience, computer-literate respondents actually expect some automated checking of their data entries along with a capability to make interactive changes to their responses, and they are surprised if these features are not built into an electronic survey.

Only some post-collection edit checks can be embedded in CSAQs. The source of edit checks added to CSAQs is the set of edits typically applied during post-collection processing. For various reasons, however, not all post-collection edit checks can be moved into the CSAQ environment. Programming or technical issues may constrain development of embedded edit checks. For example, fewer edit checks were incorporated into the 2002 Economic Census because the system's design inhibited implementation of some edit checks. The utility of some edit checks is also limited by "one-size-fits-all" approaches to the design of electronic instruments. Because the correctness of many establishment survey data items depends on the industry, editing parameters vary by industry. CSAQs are not currently tailored by industry or size of business, limiting the value of certain kinds of edit checks. Other edit checks may be too complex to communicate to respondents and too cognitively challenging for respondents to interpret during the course of completing an electronic survey. Moreover, macro-level edits that look at summary results across all respondents can only be done post-collection, thus cannot be moved into the CSAQ.

Mission criticality, typical levels of data quality, and certain respondent characteristics guide the inclusion of edit checks in CSAQs. Because of various constraints, survey managers at the Census Bureau must prioritize edit checks incorporated into electronic surveys. Priorities are placed on items deemed mission-critical. Subject area knowledge of respondents' abilities to report particular data items and typical levels of response accuracy also guide the definition and selection of edit checks for CSAQs.

Respondent acceptance of edit checks depends on several factors, including perceived usefulness. Respondent reaction remains a valid concern.

Research shows that, to a great degree, instrument control needs to remain with respondents. Usability research suggests a number of guidelines for user-centered design and implementation of CSAQ edit checks to improve the usability of electronic surveys. Operational experience suggests that respondents easily accept edit checks ensuring that the data they enter meet required formats, and these kinds of edits are effective. In addition, different levels of edits– information, warning, and edit failure – provide respondents with information about severity and let respondents choose how to deal with edit messages.

Acceptance of electronic forms containing edit-failing data reflects a greater willingness to deal with measurement error rather than to absorb nonresponse error. Usability research suggests that the issue of respondent control over resolving edit failures is perhaps most critical at the data-submission stage. Many current Census Bureau CSAQs allow respondents to submit completed electronic survey forms with data that have failed the embedded edits. The main reason for this strategy is to avoid encouraging survey nonresponse due to unresolved edit failures. All survey programs prefer edit- failing data to no data (unit nonresponse), and they continue to rely on post-collection editing and imputation to cleanse reported data. Thus, it appears that survey managers are more willing to accept measurement error, than nonresponse error, in the collected data.

6. Future Directions and Research Issues

In general, the Census Bureau's incorporation of edit checks into electronic data collection for economic surveys embodies a conservative philosophy: At a minimum, the Census Bureau receives data from cooperative respondents. Those data may or may not pass basic edit checks. Research is needed to support a more ambitious philosophy, allowing the inclusion of additional post-collection edit checks in electronic instruments in order to reduce costs and increase data quality, while maintaining respondent cooperation.

Survey practitioners would very much like to have "generally accepted practices" or "rules of thumb" for resolving electronic survey-design issues, including the open issues in data editing. However, we expect this to be virtually impossible given the variety of surveys, administrations, and trade-offs related to data quality and response. Instead we think it would be more appropriate to develop a set of research-based guidelines to aid decisions related to editing. Derived from goals and principles, and supported by research, these guidelines should be revisited periodically to ensure their relevance as technology changes. Research is needed to determine whether a core set of best practices and heuristics could always be implemented.

Issues concerning data quality and resource allocation can arise when large mail surveys offer automated data collection. Large mail surveys have high variable costs (with respect to the number of respondents and the number of survey cycles) associated with data editing because clerks and subject matter experts review edit failures produced by post-data-collection edit checks. On the other hand, editing at the time of data collection, by respondents reviewing edit messages generated by automated edit checks, can have high fixed costs for programming and questionnaire testing; but the corresponding variable costs associated with data editing should be much lower than those for traditional post-data-collection editing. Such a paradigm shift would require modifications to survey organizational cultures, structures, and resource allocation.

Survey managers' preferences for receiving edit-failing data from respondents – as opposed to no data – raise the question of whether "submission with unresolved edit failures" is a satisfactory, cost-effective, "optimum" strategy in terms of data quality, which is affected by both nonresponse error and measurement error. Investigations into data quality suggest that the potential benefits of CSAQ edit checks are realized, resulting in fewer data items failing post-collection edits (Sweet and Ramos, 1995) and fewer items being changed during analyst review (Evans, 2003). Further research is needed to corroborate this encouraging conclusion, and to evaluate the trade-offs between measurement error and nonresponse error related to interactive edit checks in electronic data collection.

7. Acknowledgements

This report is released by the U.S. Census Bureau to inform interested parties of research and to encourage discussion of work in progress. The views expressed on methodological issues are those of the authors and not necessarily those of the Census Bureau. The authors wish to thank the following Census Bureau staff for helping us gather information for this paper: Patrick Kent, Joyce Kiessling, Yvette Moore, John Nogle, Yolando St. George, and Rita Williamson. Material in this paper was previously presented to the Federal Economic Statistics Advisory Committee, March, 2003, and to the Annual Conference of the American Association for Public Opinion Research (AAPOR), May, 2004.

8. References

Dumas, J. and Redish, J. 1999. *A Practical Guide to Usability Testing*. Portland, OR: Intellect.

Fellegi, I.P and Holt, D (1976). "A Systematic Approach to Automatic Edit and Imputation." *The Journal of the American Statistical Association*, 71, pp. 17-35.

Economic Electronic Style Guide Team. 2001. "Style Guide for the 2002 Economic Census Electronic Forms," U.S. Census Bureau, Economic Planning and Coordination Division.

Evans, I. 2003. "QFR CSAQ Evaluation." Internal Memorandum. U.S. Census Bureau, Company Statistics Division.

Nichols, E., Murphy, E, and Anderson, A. 2001. "Report from Cognitive and Usability Testing of Edit Messages for the 2002 Economic Census (First Round)," Internal Memorandum. U.S. Census Bureau, Statistical Research Division, Usability Lab, Human-Computer Interaction Memorandum Series #39.

Sigman, Richard (2001). "Editing and Imputation in a Standard Economic Processing System," *Proceedings of Statistics Canada Symposium 2001*, Ottowa, Ontario: Statistics Canada.

Sweet, E. and Ramos, M. 1995. "Evaluation Results from a Pilot Test of a Computerized Self-Administered Questionnaire (CSAQ) for the 1994 Industrial Research and Development (R&D) Survey," Internal Memorandum. U.S. Census Bureau, Economic Statistical Methods and Programming Division, #ESM-9503.

Tedesco, H., Zukerberg, A., and Nichols, E. 1999. "Designing Surveys for the Next Millennium: Web-based Questionnaire Design Issues," *Proceedings of the Third ASC International Conference*. The University of Edinburgh, Scotland, UK, September 1999, pp. 103-112.

Virzi, R. 1992. "Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough?," *Human Factors*, 34:457-468.

Wagner, Dennis L. (2000). "Economic Census General Editing – Plain Vanilla," *Proceedings of the Second International Conference on Establishment Surveys*, Alexandria, VA: American Statistical Association, pp. 561-570.

Table 1. Summary of Editing Features in U.S. Census Bureau Computerized Self-Administered Questionnaires (CSAQs)

						Resolution required to submit?
R&D	Downloadable	67/205 = 0.33	P, R, M	Immediate, Deferred	Review panel	N
M3	Browser-based	103/58 = 2.43	P, B, R, A, M	Deferred	Highlighted text	N
2002 Economic Census	Downloadable	66/95 = 0.69	B, P, L, M, F	Immediate, Deferred	Icon next to item, review panel	Y (for failure on one edit check)
COS	Downloadable	36/23 = 1.57	P, R, L, M, F, RT	Immediate, Deferred	Pop-up messages, review panel	Y (for a few key edit-check failures)
ASM	Downloadable	Unavailable/88	Unavailable	Immediate, Deferred	Pop-up messages, review panel	N
QFR	Downloadable	29/94 = 0.31	B, P, L, M	Immediate	Icon next to item	N

[1] R&D = Survey of Industrial Research and Development; M3 = Manufacturer's Shipments, Inventories, and Orders Survey; COS = Company Organization Survey; ASM = Annual Survey of Manufactures; QFR = Quarterly Financial Report.

[2] **B=Balance, P=Preventive, R=Ratio, L=Logical, A=Alphanumeric, M=Missing value/Incomplete, F=Format, RT=Rounding Test**