USABILITY METHODS FOR DESIGNING A COMPUTER-ASSISTED DATA COLLECTION INSTRUMENT FOR THE CPI[1]

Jean E. Fox
Bureau of Labor Statistics

Abstract

The Bureau of Labor Statistics (BLS) is developing a Computer-Assisted Data Collection (CADC) instrument for the Commodities and Services (C&S) survey for the Consumer Price Index (CPI). The instrument must be flexible enough to handle a variety of procedures and situations, but it must also be logical and useful for the data collectors and provide tools to help to ensure data quality. Therefore, it was critical to incorporate usability into the design process. This paper describes the usability techniques we used, including a description of how we tailored some of them to the situations we faced. The paper also discusses the benefits we found by including usability engineering in the development process. So far, the feedback from the field has been predominantly positive, suggesting that our usability methods are leading to a useful and usable instrument.

Keywords: Usability, User-Centered Design, Computer-Assisted Data Collection, Development Lifecycle

Introduction

The Bureau of Labor Statistics (BLS) produces the Consumer Price Index (CPI), one of the primary economic indicators for the United States. The CPI reflects "the average change over time in the prices paid by urban consumers for a market basket of consumer goods and services." (BLS Website). Data for the CPI are collected in two surveys: (1) the Housing Survey, which collects information on housing costs, and (2) the Commodities and Services (C & S) Survey, which collects information on the cost of goods (such as food and fuel) and services (such as medical care and education).

This paper will discuss the usability engineering methods used in the development of a Computer-Assisted Data Collection (CADC) Instrument for the C & S Survey. It will also discuss how we adapted the methods to accommodate our unique situations. The paper will conclude with some of the benefits of our user-centered design approach.

The Commodities and Services Survey

For the C & S Survey, field staff members visit 25,000 outlets each month to collect price information on a total of approximately 80,000 items. Data collectors report prices for items and services in all types of establishments, such as supermarkets, department stores, and doctors' offices. Currently, this survey is collected on paper. The C&S survey is a complicated survey with many procedures, but no consistent order for data collection. In fact, different types of items have different reporting requirements (e.g., size must be collected for cans of tuna, but not for medical services). Further, the procedures can be altered depending on the policies of individual outlets (e.g., a university may allow data collection only twice a year, when they set tuition). Analysts in Washington set the requirements based on their needs for analysis.

---

[1] The views expressed in this paper are solely those of the author and do not reflect the official opinion of the Bureau of Labor Statistics.

Commodities and Services Computer-Assisted Data Collection (CADC)

BLS began developing the C & S CADC Instrument in early 1999. We plan to deploy the instrument by the end of 2002. The instrument will run on a pen tablet computer (11" x 8.5" x 1.1" and 3.2 pounds) with a color display. Because of the variety in procedures and requirements, we could not use standard survey software such as Blaise. Therefore, we developed the instrument in Visual Basic to give data collectors and analysts the flexibility required.

The Design Process

From the start, the development team employed a "user-centered design approach," meaning we placed strong emphasis on meeting the needs of the users by involving them throughout design and development. We have incorporated a variety of usability engineering methods into our design process to identify and address the users' needs.

To ensure that usability remained a key component of the C&S CADC development process, a user interface designer was assigned to the development team. The designer coordinated the usability effort with the rest of the development process and ensured that the development team addressed usability issues as they arose. The team is building the instrument iteratively, and we are using a variety of usability methods in each stage, culminating with user testing at the end of each stage. The iterative process has enabled us to resolve usability problems and other bugs when they arise rather than waiting until just before deployment, when they become more complicated and expensive to fix.

Each iteration has followed approximately the same progression. We started out by acquiring appropriate background information. In the first iteration, we conducted user and task analyses; in later iterations, we reviewed the user feedback from the previous stage. Next, we designed and developed a prototype, conducting heuristic evaluations throughout the process. Once we completed development for each stage, we conducted user testing remotely with data collectors from all over the country. We then started the next stage by addressing the problems uncovered in the user testing.

The Usability Methods

The methods we used were based on traditional usability methods. However, as can happen in some situations, we often had to tailor the methods to fit our unique needs. The remainder of this paper will describe the usability methods in more detail, including the adaptations we had to make and the types of information we uncovered. Finally, the paper will list the benefits we gained by incorporating usability methods into the design process.

User and Task Analyses

The team used these tools primarily at the beginning of the project. We needed to understand who the users were, what their skills and experience were, the tasks they needed to perform, and the environment in which they worked. We included all types of users, including the field data collectors, their managers, the analysts in Washington, and even the respondents (who could be especially burdened if the instrument hindered the data collectors).

First we conducted a user analysis to understand who the users were. We analyzed each user population and determined what characteristics would affect design. For example, we found that two characteristics of the data collectors would have a significant impact on design. First, many are middle aged or older, so we have to be sure they can read the display. Second, many have very little computer experience. Therefore, we have to design simple interfaces with large fonts and sufficient contrast so data collectors can read and use the displays.

Next, we conducted a task analysis. The goal of a task analysis is to identify all the tasks that users perform. In a task analysis, a user interface designer generally first works with users to identify high-level user goals (Kirwan and Ainsworth, 1992). Then, the goals are further analyzed to specify individual activities required to meet each goal. The analysis identifies the tools and information required for each activity. A task analysis helps designers understand what users need to do for each task, so that the instrument can be adapted to meet the users' needs.

Some of the task analyses we conducted were more formal than others. The data collection manual already described the procedures established for paper collection. The manual provided much of the task information we needed, and we frequently consulted it to understand the data collectors' tasks. However, there was no manual for the analysts' job, so we met with representative analysts to learn how they use data that would be collected with the instrument.

Evaluating the users' work environments was fairly straightforward. The data collectors work primarily inside businesses. Generally, the lighting and environmental conditions are good. However, occasionally they work outside - either at an outdoor stand or in their car. When they are outside, lighting (especially glare) can be a problem.

Thus, the user and task analyses provided important background information needed to make informed design decisions. In some cases, the information was already available. In other cases, we had to do some research to find the answers.

Heuristic Evaluation

Armed with the knowledge gained from the user and task analyses, we began to design the instrument. First we provided the programmers with ideas and drawings for each screen. They created draft screens in the "look and feel" of the instrument. Then, based on our previous knowledge and experience regarding C & S and usability, the team worked together to revise the screens to better meet the users' needs.

We followed the general procedures for heuristic evaluations as outlined by Nielsen and Mack (1994) to evaluate the draft screens. The authors provide a list of 10 "heuristics" (usability principles) that an application should follow. These heuristics include rules such as "Speak the users' language" and "Maintain consistency." The goal of an heuristic evaluation is to identify aspects of an application that are not in accordance with these usability principles. According to Nielsen and Mack, three to five usability specialists should conduct the evaluations, in order to identify the majority of usability problems. Further, the results of the evaluation should be a written report, describing the problems and the heuristics they violate.

We tailored the method somewhat for our development process. We used a less formal method so we could respond to draft screens faster. Below are some of the adaptations we made to better incorporate heuristic evaluations into our development process:

- Rather than bringing in several usability specialists for this activity, the user interface designer conducted the analyses with several members of the Office of Field Operations (OFO) at BLS. We all served on the development team, so we were already familiar with the survey. We did not have the benefit of an objective observer, but this adaptation allowed us to do the evaluations quickly.
- At this stage, we reviewed mostly paper screen shots. In a few cases, developers had new features working for us to try.
- In a sense, we combined heuristic evaluations with the method of cognitive walkthroughs (where designers "walk through" the steps a user would follow to perform specific tasks). Cognitive walkthroughs can identify where the application does not logically support the users' tasks (see Wharton, Rieman, Lewis, and Polson, 1994). We took tasks the users would have to complete, and stepped through how they would do it in the CADC instrument. We revised the interface when we felt the steps were illogical, confusing, or incomplete.
- We tried to describe the problems we found, but we did not categorize them by the heuristic(s) they violated.
- We did not write a formal report detailing our findings. Instead, we presented our findings in person at meetings with developers, using drawings and brief task descriptions when necessary to support our opinions.

The heuristic evaluations helped us fix many usability problems early in design, before there was much coding to support the instrument. The heuristic evaluations also facilitated discussion between the user interface designers and the developers. This increased communication let each party know about the needs and limitations of the other.

Remote User Testing

After each stage of development was completed, we conducted user testing with approximately 10 users. We used an instrument that included all the features from the previous stages plus all the new features we designed and developed. These users worked from geographically diverse locations (often called "remote" users). Usability testing with remote users has become a significant challenge for user interface designers, because you cannot easily observe the users' behavior directly.

The challenge for us was how to solicit feedback from the users. Conventional wisdom in the usability field is that you must observe the users to truly know where the problems are. Literature describes methods such as the following for observing users:

- Providing a button for users to click (to store video clips) whenever they experience a critical incident as they use a software prototype (Hartson, Castillo, Kelso, and Neale, 1996).
- Using a shared windowing tool and the telephone to conduct sessions remotely (Hammontree, Weiler, and Nayak, 1994).

Options such as these require fairly sophisticated tools we did not have. Another option was to have all participants travel to Washington, D.C., or to have the user interface designer travel to the participants. Unfortunately, we did not have the resources for travel, either.

We needed an inexpensive alternative. We settled on the use of questionnaires to solicit feedback from the users. We provided users with a pen computer loaded with the instrument. We gave them instructions for using the instrument, conducting the test, and completing the questionnaire. They completed the entire test from their home or local office.

We carefully constructed the questionnaire to solicit user feedback. We made the questions very specific, so users would report on the issues we would have focused on during observations (see Fox, 2000 for more information). Each screen in the instrument had a corresponding page in the questionnaire. At the top of the page, we put a picture of the screen. Below that, we put several questions asking the users to rate different aspects of the screen. At the bottom of the page, we included several open-ended questions, which allowed users to add any additional thoughts they had.

The questionnaire was quite successful at uncovering usability problems. The ratings identified general areas of problems, and the open-ended questions allowed users to express specific concerns or ideas. Several factors may have influenced the success of the questionnaire:

- The participants were extremely motivated to respond. They knew that they would have to use the instrument on an almost daily basis, and they wanted to be sure it would work for them.
- The specific questions and the screen shots helped participants focus on the issues.
- After the first iteration, users could see that we really listened to them, so they continued to provide input.

The questionnaire took the participants some time to complete, but it provided us with extremely useful feedback. We would have preferred to use a more traditional method involving direct observation, but we did not have the resources. Our users, like the data collectors for many other large-scale federal surveys, are geographically dispersed, so we could not observe them cost effectively.

Usability Testing and Focus Groups

Although we were not able to conduct traditional "usability tests" for each iteration of development, we did get a chance to observe data collectors using a version of the instrument. The field staff came to our Washington office for several days of “refresher training.” As part of their activities, the data collectors spent some time working with the current version of the instrument and providing feedback. We had too many people to conduct a real one-on-one usability test. Therefore, we combined some of the techniques of usability testing, where typical users try to complete typical tasks with the application under development (see Dumas and Redish, 1999), with techniques for focus groups, where 10-12 participants talk about a particular subject in a discussion led by a skilled moderator (see Morgan, Krueger, and King, 1998).

First, we gave the data collectors a demonstration of the instrument. Then, the data collectors spent approximately an hour working through sample schedules (one schedule contains all the forms for one outlet). We urged them to try to enter unusual situations they have faced to see how well the instrument worked. During this activity, several members of the development team wandered through the room, observing how the participants used the instrument. We also answered their questions and asked them our own questions about what they would like to see in future iterations.

After the participants finished with the instrument, we conducted a focus group discussion with the entire group. At this point, we had an idea of what problems they faced. We presented possible solutions and asked the whole group for input. Focus groups allow participants to bounce ideas off each other, sometimes creating solutions they might not have come up with on their own. We were also able to get a feel for how the group felt as a whole about certain issues.

We found this stage of the testing to be particularly useful. We were able to identify serious problems when many participants experienced the same difficulties. We got insightful feedback on how they thought the instrument should work, and we were able to suggest possible modifications and get the participants’ reactions. Further, in our evaluation, the users worked in pairs, so we also learned a lot by observing their interaction. This feedback helped us focus on design changes for the next iteration.

Benefits of User Centered Design

The user-centered design approach has led to a better product than we would have had otherwise. We knew that the users might not be comfortable with computers, but we wanted them to be at ease with the instrument. We wanted them to feel that the instrument truly helps them do their work. So far, the feedback has been optimistic. The data collectors seem pleased with the instrument and happy that we worked so hard to incorporate their suggestions, which has encouraged them to continue to provide feedback. The usability of the system should ease the transition during deployment, because the field staff are eager to learn and use this tool. The instrument enforces procedures when possible and checks the data to be sure it is complete, accurate (e.g., within appropriate ranges), and consistent before it is sent into Washington. Therefore, we expect to get more accurate data than from the paper forms, leading to more accurate calculations for the CPI.

Summary

This paper described four user-centered design methods that we used in the development of the C & S CADC Instrument. Building an instrument to facilitate the data collectors’ tasks will aid data collection and is expected to have a positive impact on data quality. The users were pleased to be included them in the design process; transitioning to the new instrument will be easier because the field is eager for the technology.

The four usability methods we used often required adaptations from traditional textbook practices to accommodate our working environment. Many of these situations, including diverse, remote users with limited availability, are no doubt common to other federal statistical programs. The adaptations we used were quite successful in our environment and should transfer to other agencies. Adopting a user-centered design philosophy has already provided a variety of benefits, and we expect the benefits to continue into deployment. Therefore, survey programs should be encouraged to consider how user centered design methods, either with traditional procedures or with adaptations, can improve other computer-assisted surveys as well.

References

BLS Website. *CPI Frequently Asked Questions*. http://stats.bls.gov/cpifaq.htm

Dumas, J.S. and Redish, J.C. (1999). *A Practical Guide to Usability Testing*. Norwood, NJ: Ablex

Fox, J.E. (2000). Collecting data from data collectors: Using questionnaires to gather feedback from remote users. Companion paper to poster presented at *The Usability Professionals' Association Conference*, Asheville, NC.

Hammontree, M., Weiler, P., and Nayak, N. (1994). Methods and tools: Remote usability testing. *Interactions*, 1(3) 21-25.

Hartson, H.R., Castillo, J.C., Kelso, J., and Neale, W.C. (1996). Remote evaluation: The network as an extension of the usability laboratory. *Proceedings of ACM CHI 96 Conference on Human Factors in Computing Systems*, v.1, 228-235.

Kirwan, B. and Ainsworth, L.K. (Eds.). (1992). *A Guide to Task Analysis*. Washington, DC: Taylor & Francis

Morgan, D. L., Krueger, R. A. and King, J.A. (1998). *The Focus Group Kit, Volumes 1-6*. Thousand Oaks, CA: Sage Publications

Nielsen, J. and Mack, R.L. (1994). *Usability Inspection Methods*. New York: John Wiley & Sons, Inc.

Wharton, C., Rieman, J., Lewis, C., and Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In Nielsen, J. and Mack, R.L. (Eds) *Usability Inspection Methods*. New York: John Wiley & Sons, Inc.