

The Census Bureau plans to conduct additional research under the generic clearance for questionnaire pretesting research (OMB number 0607-0725). We will be conducting usability research using 20 questions from the April 2011 Internet Test version of the American Community Survey (ACS) and the 2010 ACS Content Test. This study is a continuation of two prior studies (which were sent to OMB on 6/6/2011 and 10/28/2011), which sought to find a way to identify confused Internet respondents. The first two studies identified specific mouse cursor movements that are associated with difficulty or confusion answering a question on an Internet survey. We then attempted to generate a model, using the different movements, that can predict when a respondent is having difficulty with a question in real-time. The principal application of the prior research is the ability to provide respondents with tailored real-time assistance while they are taking an Internet survey. The goal is to implement this feature in a future ACS Internet test.

Past research has suggested that while real-time assistance leads to more accurate responses, respondents do not seem to like unsolicited assistance. Therefore, this study will experiment with different methods of providing real-time assistance with the goal of maximizing respondent satisfaction and accuracy. To ensure respondents have trouble answering questions and need help, this study will deliberately manipulate respondent difficulty answering a question using scenarios. Each question will have two versions: one complicated and one straightforward. The level of difficulty will be manipulated either by using two different scenarios for one question, or two questions for one scenario. Each respondent will receive 10 complicated questions and 10 straightforward questions. Each question will have Help text associated with it. Participants can click on a Help link or, if a time threshold is surpassed (median response time taken from the prior study, which used the same questions and scenarios), the model will initiate the Help. Three different treatments for providing model-initiated Help are tested in this study: text, audio, and chat. Each respondent will be randomly assigned to one of the three treatments. A copy of the questions, scenarios, and Help text is enclosed. In addition to manipulating difficulty, using scenarios also allows us to know the true answer to each question, and thereby assess accuracy. After participants complete the survey, they will be asked to complete a satisfaction questionnaire on the help they received (enclosed). These questions ask about overall satisfaction with the survey, as well as with the Help they received.

The three methods for providing Help will be compared based on accuracy and the satisfaction ratings from the follow-up survey. Since each question/scenario combination has a single correct response, we can compare accuracy across treatment. Specifically, we will calculate the proportion of correct responses when Help was provided by the model and compare it to the proportion of correct responses when Help was requested by the respondent (for each treatment) and when no Help was received. We will statistically test the differences between these proportions to determine whether one of the treatments stands out as yielding a higher proportion of correct responses across the different comparisons. We will also create an index for participant satisfaction with the Help they received. This index will be created using factor analysis. The initial factors will be collected from the questions on the satisfaction questionnaire. Questions measuring the same underlying factor will be combined into a single composite satisfaction score. This satisfaction index will be predominantly based on the initiation treatment, but will also include satisfaction with the Help content. Again, we will compare the average score across treatments to determine which treatment participants prefer. Respondents report preferring more human-like interactions, so we expect the chat treatment to yield the highest satisfaction score. This research will help future researchers provide respondents with assistance they need without deterring them from completing the survey.

In September and October 2012 we will conduct approximately 150 laboratory interviews (50 participants will be assigned to each treatment). This sample size will allow us to detect differences

of .25 with a power of 0.80 at a 0.05 significance (this calculation is based on differences seen in Conrad and his colleagues' 2007 paper that looked at accuracy and satisfaction after providing model-initiated help to respondents). Recruiting will take place using Craigslist advertisements and flyers posted on the University of Maryland campus. The interviews will be conducted at the Joint Program in Survey Methodology at the University of Maryland, in a lab equipped with Tobii eye- and mouse-tracking capabilities. Participants of varying ages, education levels, and levels of computer experience will be recruited using a screening questionnaire. A copy of the screening questionnaire is enclosed. Eligible participants will answer 20 ACS questions on a computer using a condensed version of the ACS Internet instrument. Following each question, the participant will also be asked to rate the previous question on its level of difficulty using a five-point scale. Additionally, respondents will answer a series of demographic questions using a paper form (a copy of this questionnaire is attached). Finally, they will be debriefed on their experience with the survey. A copy of the research protocol is enclosed.

Interviews will be recorded and the participant's eye movements will be tracked, with respondent permission, to facilitate analysis of the results. Participants will be informed that their response is voluntary and that the information they provide is confidential and will be seen only by employees involved in the research project. Participants will receive \$30 for their participation.

Recruiting and screening will take approximately 10 minutes per participant and interviews will take approximately 45 minutes each. Thus, the total estimated burden for this research is 140 hours.

The contact person for questions regarding data collection and statistical aspects of the design of this research is listed below:

Rachel Horwitz
Decennial Statistical Studies Division
U. S. Census Bureau
4600 Silver Hill Road
Washington, DC 20233
301-763-2834
rachel.t.horwitz@census.gov