STUDY SERIES
(Survey Methodology #2009-01)

Final Report for First-Round Usability Testing of Data-Reliability Indicator Prototypes

Kathleen T. Ashenfelter
Jenna Beck
Elizabeth D. Murphy

Statistical Research Division
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: February 12, 2009

Date: Thursday, February 12, 2009

To: American Community Survey Data-Reliability Indicator Team: Jennifer Tancreto and Michael Springer, Decennial Statistical Studies Division (DSSD)

From: Usability Team: Kathleen T. Ashenfelter, Jenna Beck, and Elizabeth D. Murphy, Statistical Research Division (SRD)

Subject: Final Report for First-Round Usability Testing of Data-Reliability Indicator Prototypes

---

Table of Contents

List of Figures

List of Tables

Executive Summary

All American Community Survey (ACS) estimates are currently released with their associated margin of error (MOE). However, many users do not think that the MOE provides enough information about an estimate's reliability. We conducted usability testing to examine whether the addition of reliability indicators to new prototypes of ACS data tables helped people use the tables with greater accuracy, efficiency, and satisfaction. All three prototypical reliability indicators were based on the coefficient of variation (CV), which is defined as the standard error divided by the mean of the estimate.

Each prototypical table with a reliability indicator tested in this study implemented a color-coded "Reliability" column and a legend explaining the meaning of the color codes to provide users with guidance as to whether the proportion of error to the estimate itself might be considered unacceptably high.

Each prototype table was defined by the number of "levels" that its Data Reliability Indicator used to label the estimates. Specifically, the two-level indicator had the levels "blank" (CV<=0.30) and "use caution" (yellow; CV>0.30); the three-level indicator included the levels "good" (green; CV<=0.30), "fair" (yellow; 0.30<CV<=0.61), and "poor" (red; CV>0.61); and the four-level indicator had the levels "excellent" (green; CV<=0.10), "good" (yellow; 0.10<CV<=0.30), "fair" (orange; 0.30<CV<=0.61), and "poor" (red; CV>0.61). Nine participants completed usability tasks using the prototypes, and three participants completed the tasks using baseline versions of the current ACS data tables without the indicators. Full versions of the prototype tables can be found in Appendix A.

The results showed that the four-level indicator was associated with the highest accuracy and satisfaction scores. In fact, eight of the twelve participants indicated that they preferred the four-level indicator overall. Users were able to complete the tasks more efficiently using the prototypes than the baseline tables, and they also expressed a strong preference for all of the prototypes over the baseline tables.

1.0 Introduction

This usability research project aims to address an issue that arises with the American Community Survey (ACS) data tables because the estimates in these tables have varying levels of reliability.* Some of the data, especially some single-year estimates, have an unacceptably high standard error (typically measured by the coefficient of variation (CV), which is defined as the standard error of an estimate divided by the mean of that estimate). Some users may use these estimates without taking their accuracy into account (Whitford & Weinberg, 2008). The goal of this project is to provide some context (e.g., a data reliability indicator) to help data users more easily detect when there are potential issues with interpretability due to large values for the CV. This data reliability indicator is intended to provide some information about the size of the standard error relative to the estimate, but is not intended as a decision-making tool for whether or not to use the estimate. This decision, which should be based on the context of the estimate's use, is ultimately the responsibility of the data user.

The current method for addressing the issue of estimate reliability was to color-code each estimate with the level of confidence, as measured by the CV. The CV was chosen as the estimate of sampling error for these data reliability indicators because it is a relative measure of the size of the standard error in relation to the estimate. Also, the CV can be compared across estimates and across different tables. The existing data quality standard, Quality Requirements for Releasing Data Products, states: "If the estimated coefficients of variation (CV) for key statistics are larger than 30 percent, the data product will be released under the requirements for category 2 or category 3 data" (Cahoon, Donnalley, Gore, Kostanich, Runyan, Detlefsen, and Stern (2007), Page 7 section 1.B). As a starting point, a four-level categorization (a stoplight color indicator) based on this documented Census CV<=0.30 standard was proposed by the sponsor.

The first round of usability testing on the three prototypes of the proposed Data-reliability indicator, which can be found in Appendix A, took place from September 29 to October 2, 2008. Testing with twelve internal participants occurred in the Usability Suite at Census Bureau Headquarters (Room 5K509 and associated test rooms). Three participants performed all of the tasks for three pairs of current tables (e.g., no indicators) and nine participants completed all of the tasks using the proposed table prototypes with the indicators. ACS team members observed a few of the sessions. In this report, each prototype is defined by the number of "levels" that its data reliability indicator uses. Specifically, the two-level indicator has the levels "blank" and "use caution" (See Figure 1).

*Note: If a table generally has extremely unreliable estimates, it will be filtered from viewing on the ACS Web site. Users will not be able to see these tables.

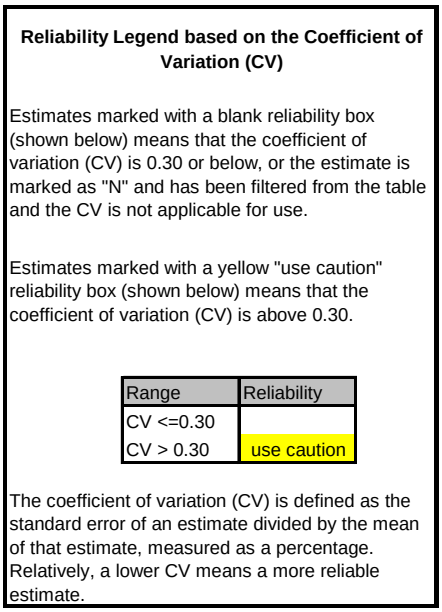


Figure 1: Legend from two-level indicator

The two-level indicator legend box was the only version to include an explanation of what an “N” inside an estimate box means. Several participants expressed confusion over the meaning of these “N”s. The explanation was added to this version of the indicator box following a dry-run of the usability test where the participant expressed confusion about the meaning of these “N”s in the two-level condition.

The three-level indicator includes the levels “good,” “fair,” and “poor”(Figure 2).

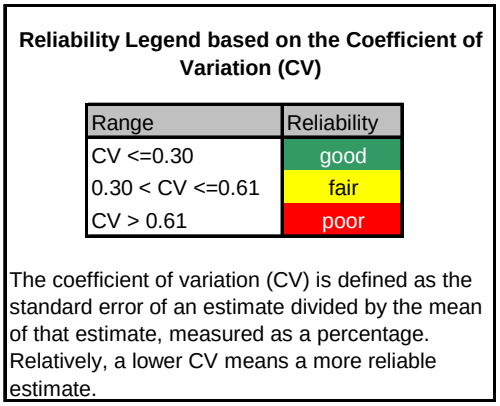


Figure 2: Legend from three-level indicator

This version of the reliability indicator is the closest of the three to the concept of a traffic “stoplight”, which was the conceptual framework around which these prototypes were developed. The vertical order of these indicators (green on top, yellow in middle, red on bottom), is the reverse of an actual traffic stoplight, and a few participants noted this

difference, although this reverse order did not cause any noticeable difficulties for the participants.

Finally, the four-level indicator has the levels “excellent,” “good,” “fair,” and “poor” (Figure 3).

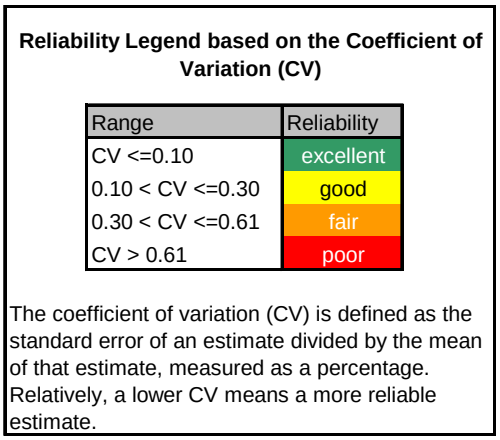


Figure 3: Legend from four-level indicator

This four-level indicator was the most popular of the three versions among participants in the current round of testing and was associated with the highest accuracy and user satisfaction scores.

Screen captures of the entire two-level, three-level, and four-level prototype tables are provided in Appendix A.

1.1 Purpose

The purpose of the testing was to examine how well the data-reliability indicator worked for users (especially as compared to the current ACS data tables without the indicator) and to identify any problems that actual users might have with the data tables. The data reliability indicator was based on the coefficient of variation (CV), which is defined as the standard error of an estimate divided by the mean of that estimate.

Another purpose of this testing was to examine whether users would notice and use the margin of error (MOE) when answering questions about the estimates from the table. This second testing goal was based on the observation that although the MOE is currently provided with each estimate, the MOE is routinely ignored by ACS data users (Whitford & Weinberg, 2008). This low-fidelity testing is the first round in a series of planned tests that are part of a larger research project about Data-Confidence Indicators (Whitford & Weinberg, 2008).

The usability testing performed for this stage of the project involved providing participants with one version of the ACS data tables at a time and having them perform a series of “tasks” intended to mirror the type of research questions for which novice users

of ACS data (such as small-town reporters, students, or marketing researchers) might use the ACS data tables. For instance, one task asked the participants to find the answer to the following vignette, "*You are researching a marketing research report for work and need to find the number of people of Swiss descent in Yakima, WA. What do you report in the paper based on your findings in the tables?*" Each of these tasks, in some way, was intended to gauge whether or not users would report a measure of standard error (MOE, CV, or data reliability indicator), and/or whether this measure of standard error was used correctly. The tasks used in the subsequent rounds of testing will be worded as closely to these tasks as possible in order to support a comparison of results across iterative rounds. During a meeting of the Census Bureau Math Stat Council (01/08/09), the recommendation was made to remove the word "reliability" from the tasks because it could potentially bias the user toward looking for a reliability column or indicator (and therefore make it harder to compare the baseline and prototype tables). The tasks will be reworded to avoid this bias in future rounds of testing.

1.2 Scope

A specific set of user interactions with the Tables (as portrayed in the low-fidelity prototypes provided by the sponsor) was within the scope of the usability evaluation. The user interface was not tested for compliance with the Section 508 regulations, although the SRD/DSSD team did meet with Laura Yax of SSD to discuss potential Section 508 issues while developing the reliability indicator prototypes before the testing occurred. Since these tables are available through a government Web site (www.Census.Gov), they must comply with Section 508 regulations before the Web site becomes available, unless a waiver is granted.

1.3 Usability Goals

The usability goals for this study were defined in three categories: user accuracy, efficiency, and satisfaction.

Goal 1-Accuracy: To achieve a high level of accuracy in completing the given tasks using the ACS data reliability indicators. The predetermined goal set by the sponsor in collaboration with SRD was that the user was expected to be able to successfully complete 70 % of the tasks given.

Goal 2-Efficiency: To achieve a high level of efficiency in using the ACS data reliability indicators site. The test participants should be able to complete the tasks in an efficient manner taking no longer than 3 minutes for a harder task and 1 minute for an easier task. Overall, the goal for this testing was to achieve an average task completion time of less than 2 minutes.

Goal 3-Satisfaction: For the users to experience a moderate to high level of satisfaction from their experience with the ACS data reliability indicators. A Questionnaire for User Interaction Satisfaction (QUIS) was implemented. The overall mean of the QUIS ratings

was expected to be well above the mean (at 5 or above on a seven-point scale, where 1 is the lowest rating and 7 is the highest rating). The same was expectation was true for the individual QUIS items.

1.4 Assumptions

- Participants had at least one year of prior Internet and computer experience
- Participants had prior knowledge of how to navigate a Web site.
- Participants had not had extensive prior experience in using the previous ACS data tables.
- Participants had no known disabilities.

Note: Since all of the participants for this round of testing were internal Census employees, they did have some knowledge of Census terminology. However, no participants were directly involved with current ACS operations.

1.5 Method and Experimental Design

The ACS data tables used for this round of testing were first downloaded from the Web site in Excel format in order to calculate the CVs and code the estimates. The data tables were then formatted to look as close as possible to their appearance on the Web site. The tables were then converted into HTML Web pages and posted to an internal Census server (http://develop.ssd.census.gov) so that they could be used with the Tobii eyetracking system.

Because this test was designed to compare the prototypes to each other and to the current format of the tables, three participants completed all of the tasks using the current tables as a baseline condition. Nine participants completed the tasks using each of three prototypes in a randomized order to control for practice effects. During the debriefing portion of the testing session, the baseline participants were shown the prototypes and the prototype participants were shown the baseline tables and asked to state their preference.

For this testing, the independent variable was the number of categories/levels included in indicator and the dependent variables were: 1) User satisfaction based on a tailored version of the Questionnaire for User Interaction Satisfaction (QUIS), 2) efficiency (time to complete each task), and 3) accuracy (whether or not the participant found the correct answer). Some other basic descriptive statistics and qualitative eye-tracking results are provided in the results section.

As mentioned above, the number of “levels” of the indicator was used as a label to distinguish among the indicators. Each different type of indicator was associated with a different pair of geographical locations, which the participants used to answer the tasks. Each pair of locations included one area with an extremely large population and reliable estimates and one area with a small population with less reliable estimates for

comparison. The smaller geographical areas had a population of about 65,000 or above, since this is the minimum for one-year estimates for the ACS. The indicators were associated with the geographical areas listed in Table 1.

Table 1: Indicators and Associated Geographical Areas

	Large Area	**Small Area**
Two-Level	New York, NY	Longmont, CO
Three-Level	Cook County, IL	Yakima, WA
Four-Level	California	Wilmington, DE

The baseline participants saw the tables in the following order:

BLP1 NewYork-Longmont /CookCounty-Yakima/California-Wilmington
BLP2 CookCounty-Yakima / California-Wilmington / NewYork-Longmont
BLP3 California-Wilmington / NewYork-Longmont / CookCounty-Yakima

To control for order effects, each pair of baseline tables was seen once as the first table presented to a participant, once as the second table presented, and once as the third table presented. The only main difference between the tables in the baseline and prototype conditions was that all of the prototypes included a data reliability indicator, which included a color-coded "Reliability" column and a legend explaining that coding (see Appendix A). The geographical area pairs were the same in both the baseline and prototype conditions.

Here is the specific breakdown for the test design for the prototype:

Nine participants saw all three of the prototypes with the indicators.

- Three participants saw the two-level first, where at least one participant saw the three-level second, and at least one saw the four-level second.

- Three participants saw the three-level first, where at least one participant saw the two-level second, and at least one saw the four-level second.

- Three participants saw the four-level first, where at least one participant saw the two-level second, and at least one saw the three-level second.

The prototype participants saw the tables in the following order:

P1 three-level/four-level/two-level
P2 two-level/four-level/three-level
P3 two-level/four-level/three-level
P4 two-level/three-level/four-level
P5 three-level/two-level/four-level

P6 four-level/two-level/three-level
P7 three-level/four-level/two-level
P8 four-level/three-level/two-level
P9 four-level/two-level/three-level

1.5.1 Participants and Observers

Participants in the September testing were recruited internally from the Census Bureau. These internal participants were sampled to give the team a feel for which features of the low-fidelity prototypes worked for users and which did not. However, we did not try to stratify this sample by age, gender, level of sophistication with statistics, etc. This level of stratification may be performed for future rounds of testing. Each test participant had at least one year of prior experience in navigating different Web sites. Participants varied in their levels of familiarity with ACS and data. Observers from the ACS Data Reliability Indicator team were invited to watch the usability tests on television screens in a separate room from the test participant and test administrator. At the end of each sponsor-observed test session, the test administrator and observer(s) discussed the preliminary findings from that session.

1.5.2 Facilities and Equipment

1.5.2.1 Testing Facilities
The test participant sat in a small room, facing a one-way glass and a wall camera, in front of an LCD monitor equipped with an eye-tracking machine that was placed on a table at standard desktop height. The test participant and test administrator were in the same room.

1.5.2.2 Computing Environment
The tester's workstation consisted of a Dell OptiPlex GX150 personal computer with a Pentium IV processor and 1 GB of RAM, a Tobii LCD monitor equipped with cameras for eye-tracking, a standard 101/102 key quiet keyboard, and a PS2 IntelliMouse with a wheel. The operating system was Windows XP for all test participants.

1.5.2.3 Audio and Video Recording
Video of the application on the test participant's monitor was fed through a PC Video Hyperconverter Gold Scan Converter, mixed in a picture-in-picture format with the camera video, and recorded via a Sony DSR-20 digital Videocassette Recorder on 124-minute, Sony PDV metal-evaporated digital videocassette tape. Audio for the videotape was picked-up from one desk and one ceiling microphone near the test participant. The audio sources were mixed in a Shure audio system, eliminating feedback, and fed to the videocassette recorder.

1.5.3 Eye-Tracking
Tracking of the participants' eye movements was recorded during the usability test. Using the ClearView 2.0 software program, the Tobii eye-tracking device monitored the

participants' eye movements and recorded eye-gaze data. This type of eye-tracking requires the calibration of each eye. Data collected from the eye-tracking device included eye-gaze position, timing for each data point, eye position, and a heat map showing where and how often a participant looked at a certain region of the computer screen. Red areas on the heat maps are "hotspots" that were fixated on more frequently.

1.5.4 Materials

Testing materials included the items described in the following subsections. These items are provided in the appendices noted.

1.5.4.1 General Introduction

The test administrator read some background material and explained several key points about the session. The general introduction emphasizes that the participant's skills and abilities are not being tested. The participant is helping in an evaluation of the site's look and feel. A draft is provided as Appendix B.

1.5.4.2 Consent Form

Prior to beginning the usability test, the test participants completed a general consent form (supplied in Appendix C). The consent form documents the participant's agreement to permit videotaping of the testing session.

1.5.4.3 Questionnaire on Computer Use and Internet Experience

After completing the consent form, the test participant completed a satisfaction questionnaire, which gathers self-reported information on the participant's computer use and Internet experience (Appendix D). This information helps the usability analysts understand a participant's performance in the context of their previous experience. Occasionally, a participant inflates his or her previous experience, but that is generally evident in the kinds of problems they have with the tasks (e.g., basic navigational issues).

1.5.4.4 Tasks and Prototypes

Working collaboratively, members of the ACS data reliability indicators team and members of the Census Bureau's Usability Lab created the tasks. The tasks were designed to capture the participant's interaction with, and reactions to, the design and functionality of the ACS data confidence indicators. Appendix E provides the version of the tasks that were used in the first round of testing, and which we plan to use again for future rounds of testing. The design includes a baseline condition in which a small set of separate participants completed the same tasks using the current format of the ACS tables.

Most of the participants completed the tasks using the new ACS Data Reliability Indicator prototypes in a pseudorandom order. Specifically, the first table that each participant saw was randomly assigned using the SAS Proc Plan function, but the second and third table order was pseudorandomly assigned. This assignment was done such that

three participants saw the two-level first, where at least one participant saw the three-level second, and at least one saw the four-level second; three participants saw the three-level first, where at least one participant saw the two-level second, and at least one saw the four-level second; and three participants saw the four-level first, where at least one participant saw the two-level second, and at least one saw the three-level second.

The inclusion of both baseline and experimental conditions allowed us to evaluate whether users were consistently not looking at or using the Margin of Error (MOE) column to begin with, and whether the indicator and/or legend box in the new ACS table would help if added.

1.5.4.5 Questionnaire for User Interaction Satisfaction (QUIS)

The original version of the QUIS includes dozens of items related to user satisfaction with a user interface (Chin, Diehl, and Norman, 1988). In a usability test at the Census Bureau, SRD typically uses 10 to 12 items that the usability team has tailored to the particular user interface being evaluated. In this study, we used a modified version that includes items worded for the ACS data reliability indicators context (Appendix F). In all, participants completed three QUIS evaluations (one for each prototype presented). Participants completed the QUIS immediately after finishing the tasks for each specific prototype.

1.5.4.6 Debriefing Questions

After completing the tasks, the test participants answered debriefing questions about their overall experience using the prototype ACS Data Confidence Indicator (Appendix G).

1.5.5 Procedure

The test participants reported to the usability lab at the U.S. Census Bureau headquarters building in Suitland, Maryland. Upon arriving at the lab, the test participant was seated in the testing room. The test administrator greeted the participant and read the general introduction. Next, the participant read and signed the consent form. After signing the consent form, the test participant completed the questionnaire on computer use and Internet experience. During this time, the test administrator started the audio-video recording.

In order to make sure that the participants understood what was expected by thinking aloud, they engaged in practice think-aloud tasks during which they walked through their thought process while performing a task using a commonly accessed Web page (the instructions for this practice can be found at the end of Appendix B).

Since this test used the eye-tracking device, the participant's eyes were calibrated at this time. Calibration was completed in about fifteen to twenty seconds, by having the

participant look at a dot moving across the computer screen. After calibration the test administrator operated the eye-tracker.

Following calibration, the participant completed the tasks using the ACS data reliability indicators prototype. At the start of each task, the participant read the task aloud. While completing the task, the participants were encouraged to think aloud and share what they were thinking about the task. This interaction was not intended to be a conversation. The participant was instructed to keep up a running commentary akin to a stream of consciousness narrative. If at anytime the participant became quiet, the test administrator reminded the participant to think aloud. The content of the "think-aloud" protocol allows us to gain a greater understanding on how the participant dealt with each task and to identify issues with the tables.

At the conclusion of each task, the participant stated a "final answer" to the task. During the task, the test administrator noted any indicators of confusion, frustration, or other behaviors. After the participant completed all tasks, the eye-tracking device was stopped, but the video recording continued. The test participant then completed the modified satisfaction questionnaire silently.

Each test participant then answered some debriefing questions. This was an opportunity for a conversational back-and-forth exchange, but the test administrator was trained to remain neutral. At the conclusion of the usability evaluation, the video recording was stopped. In general, each usability session lasted about 60 minutes.

2.0 Results

In this section, statements such as "X table's scores were more accurate than Y table's scores" refer to general trends and overall averages in the data and do not imply statistical significance. For this round of testing, the sample size was too small for inferential testing, but may be increased for this purpose for future rounds of testing.

Usability Goals: Accuracy, Efficiency and Satisfaction

2.1 Goal 1: Accuracy

Accuracy was calculated by evaluating whether the participant found the correct answer in the time allotted for the task. If the participant failed to report the correct answer in the allowed amount of time (based on the difficulty of the task), this was counted as a failure. Since many of the tasks included multiple questions, each sub-question was individually coded, and the score for that task was computed by taking the average of the scores for the sub-questions. This means that each score is a potential fraction instead of just a 0 or 1. So, if there were two sub-questions and a participant got one sub-question correct and one incorrect, the accuracy score for that task would be 0.5. Also, the response was marked as incorrect if it took longer than 1 minute to complete for an easy task or over 3

minutes for a difficult task (although this rarely happened). Information about task difficulty can be found with the task list in Appendix E.

The breakdown for the scoring of the tasks for accuracy is as follows:

California/Wilmington, DE

Task 1: What is the first thing you notice about these tables?
N/A- this is a qualitative question that cannot be coded for accuracy

Task 2: Your supervisor asks you to find the total number of women ages 15 to 50 who gave birth in the past 12 months for your hometown of Wilmington, DE [Answer A: 975]. He wants you to also provide the same number for California as a point of comparison [Answer B: 515,991]. What would you tell your boss based on the tables?
Average taken of Answers A and B.

Task 3. You are researching a paper and need to find the number of people of West Indian descent in Wilmington, DE [Answer A: 0]. What do you report in the paper based on your findings in the tables? [Answer B: Don't use estimate]
Average taken of Answers A and B.

Task 4. Find the total number of people with German ancestry [Answer A: 4,206] and the total number of people with Slovak ancestry [Answer B: 70] for Wilmington, DE. Which category of ancestry do you think is a better estimate in terms of data quality? [Answer C: German]. Please explain why you think this is a better estimate of data quality.
Average taken of Answers A, B, and C.

Task 5. a. For both California and Wilmington, DE, please find any three estimates that have a high reliability. [Answer A: any 3 "excellent" estimates] b. Please find any three estimates that are low in reliability. [Answer B: any 3 "poor" estimates]
Task 6: You are asked to report to state leaders the number of people of Italian descent living in California. What answer would you give them? [Answer A: 1,560,870] Would you recommend using this number? [Answer B: Yes] Why or why not?
Average taken of Answers A and B.

Task 7: Look at the table for Wilmington, DE and table for California. In general, which area has more reliable estimates? [Answer: California]. Explain why you gave this answer.

Cook County, IL/ Yakima, WA

Task 1: What is the first thing you notice about these tables?
N/A- this is a qualitative question that cannot be coded for accuracy

Task 2: Your friend asks you to find the total number of males who are separated for her hometown of Yakima, WA [Answer A: 779]. She wants to know the same number for Cook County, IL for comparison [Answer B: 941,158]. What would you tell your friend based on the tables? Probe if indicator not mentioned: Would you tell your friend about the differences in reliability or quality between the two estimates?
Average taken of Answers A and B.

Task 3: Find the total number of people with English [Answer A: 9,178] ancestry and the total number of people with Welsh ancestry [Answer B: 60] for Yakima, WA. Which category of ancestry do you think is a better estimate in terms of data quality? [Answer C: English]. Please explain why you think this is a better estimate of data quality.
Average taken of Answers A, B, and C.

Task 4: a. For both Cook County, IL and Yakima, WA, please find any three estimates that have a high reliability. [Answer A: any 3 "good" estimates (all had to be correct for a score of "1" on this sub-question)] b. Please find any three estimates that are low in reliability. [Answer B: any 3 "poor" estimates]
Average taken of Answers A and B.

Task 5: You are researching a marketing research report for work and need to find the number of people of Swiss descent in Yakima, WA. [Answer A: 0] What do you report in the paper based on your findings in the tables [Answer B: don't use number]?
Average taken of Answers A and B.

Task 6: You are asked to report to state leaders the number of people of Polish descent living in Cook County, IL. What answer would you give them? [Answer A: 507,498] Would you recommend using this number? [Answer B: Yes] Why or why not?
Average taken of Answers A and B.
Task 7: Look at the table for Yakima, WA and the table for Cook County, IL. In general, which area has more reliable estimates? [Answer: Cook County, IL]. Explain why you gave this answer.

New York, NY/ Longmont, CO

Task 1: What is the first thing you notice about these tables?
N/A- this is a qualitative question that cannot be coded for accuracy

Task 2: You are asked to report to state leaders the number of people of Portuguese descent living in New York, NY. What answer would you give them? [Answer A: 15,905] Would you recommend using this number? [Answer B: Yes] Why or why not?
Average taken of Answers A and B.

Task 3. You are researching a marketing research report for work and need to find the number of people of Dutch descent in Longmont, CO [Answer A: 2,365]. What do you report in the paper based on your findings in the tables? [Answer B: Yes, use the number]
Average taken of Answers A and B.

Task 4. Find the total number of people with Subsaharan African ancestry [Answer A: 0] and the total number of people with Norwegian [Answer B: 2,608] ancestry for Longmont, CO. Which category of ancestry do you think is a better estimate in terms of data quality? [Answer C: Norwegian]. Please explain why you think this is a better estimate of data quality.
Average taken of Answers A, B, and C.

Task 5. a. For both New York, NY and Longmont, CO, please find any three estimates that have a high reliability. [Answer A: any 3 unmarked estimates] b. Please find any three estimates that are low in reliability. [Answer B: any 3 "use caution" estimates]
Average taken of Answers A and B.

Task 6. Your coworker asks you to find the total number of people age 5-15 with a disability for his hometown of Longmont, CO [Answer A: 387]. He wants to know the same number for New York, NY for comparison [Answer B: 62,075]. What would you tell your coworker based on the tables? Probe if indicator not mentioned: Would you tell your coworker about the differences in reliability or quality between the two estimates?
Average taken of Answers A and B.

Task 7. Look at the table for Longmont, CO and the table for New York, NY. In general, which area has more reliable estimates? [Answer: New York, NY]. Explain why you gave this answer.

Tables 2 and 3 give summaries of the Accuracy scores for this round of testing for the Baseline (n=3; BLP1-BLP3) and Prototype (n=9; PTP1-PTP9) participants, respectively. In the interest of space conservation in the following tables, the specific versions of the tables are referred to by their large geographical area (e.g., California, Cook County, or New York). The full accuracy results are available in Appendix H.

Table 2: Baseline-Condition Accuracy Scores by Table and Participant (Collapsed across all tasks)

Table	BLP1	BLP2	BLP3	Average Success by Table
California	0.92	0.83	0.53	76%
Cook County	0.75	0.67	0.69	70%
New York	0.83	0.69	0.60	71%
Average Success by Participant	83%	73%	61%	Overall Average 72% (n=3)

Table 3: Prototype-Condition Accuracy Scores by Table and Participant (Collapsed across all tasks)

Table	PTP1	PTP2	PTP3	PTP4	PTP5	PTP6	PTP7	PTP8	PTP9	Average Success by Table
California	1.00	1.00	0.92	0.83	0.92	0.81	0.83	0.92	0.83	90%
Cook County	0.58	0.92	1.00	0.83	0.92	1.00	0.83	1.00	0.92	80%
New York	1.00	1.00	0.92	1.00	0.83	0.92	0.83	0.53	0.89	79%
Average Success by Participant	86%	97%	94%	89%	89%	91%	83%	81%	88%	Overall Average 83% (n=9)

Goal 1: Average Accuracy > 70 %
Baseline: Average Accuracy=72%; PASS
Prototype: Average Accuracy=83%; PASS

Both the baseline and the prototype tables passed the basic usability accuracy goal of 70 percent. Remarkably, the four-level tables had a 90 percent accuracy rate overall, followed by the three-level indicator at 80 percent and the two-level indicator at 79 percent. All of the prototype tables had higher accuracy scores than the average baseline accuracy of 72 percent.

2.2 Goal 2: Efficiency

Efficiency scores are simply the total time it took a participant to complete a particular task from the time s/he finished reading the task aloud to when s/he finished reporting his or her final answer. Although responses were marked as incorrect if the participant took

longer than one minute for an easy task or three minutes for a harder task for accuracy, s/he was not stopped at this time limit but was allowed to attempt to complete the task (due to the exploratory nature of this round of testing). So, participants could have an efficiency score of over three minutes for a particular task (although this happened in-frequently). Tables 4 and 5 provide a summary of the efficiency scores for the baseline and prototype tables for this round of testing.

Table 4: Baseline-Condition Efficiency Scores in Minutes and Seconds by Table and Participant (Collapsed across all tasks)

Table	BLP1	BLP2	BLP3	Average Time by Table
California	0:58	1:24	NA*	1:11
Cook County	1:08	1:17	NA*	1:12
New York	1:52	1:21	NA*	1:36
Average Time by Participant	1:19	1:21	NA*	Overall Average Time 1:20

***The video recording was not codable for time stamps for this participant**

Table 5: Prototype-Condition Efficiency Scores in Minutes and Seconds by Table and Participant (Collapsed across all tasks)

Table	PTP1	PTP2	PTP3	PTP4	PTP5	PTP6	PTP7	PTP8	PTP9	Average Time by Table
California	0:51	1:07	0:45	1:45	0:33	1:35	1:13	3:17	1:17	1:14
Cook County	1:31	0:44	0:45	1:08	1:03	1:20	1:36	1:25	1:09	1:04
New York	1:00	1:34	2:26	1:30	1:37	1:39	1:06	0:29	1:19	1:16
Average Time by Participant	1:07	1:08	1:18	1:28	1:04	1:31	1:19	1:44	1:15	Overall Average Time 1:11

On average, the prototype participants completed the tasks more quickly than the baseline participants did. However, the individual prototype tables do not show a clear advantage with respect to shorter time-on-task as compared to the baseline tables. This finding could be due to the prototype participants' requiring more time to interpret the data reliability indicators and thinking about standard error/data reliability. On the other hand, participants in the baseline condition expressed more uncertainty than the prototype participants did when looking for answers. So, baseline participants take longer to decide with certainty on an answer, while prototype participants may be using the data reliability

indicators to take standard error and the reliability of the estimates into account before reporting an answer. Further testing is recommended to examine this issue.

Goal 2: Average Time-on-Task > 2 minutes
Baseline: Average Time-on-Task=1 minute, 20 seconds; PASS
Prototype: Average Time-on-Task=1 minute, 11 seconds; PASS

Full efficiency score results can be found in Appendix I.

2.3 Goal 3: Satisfaction-Questionnaire for User Interaction Satisfaction Score

Overall, each of the prototypes received a higher average satisfaction score than the baseline tables did. The four-level indicator received the highest satisfaction score with an average rating of 7.36, but the three-level indicator was close behind with an average rating of 7.25. Table 6 shows a summary of these scores.

Table 6: User Satisfaction Score Results

Table	Average QUIS Score
Baseline	6.79
two-level	7.04
three-level	7.25
four-level	7.36

Goal 3: QUIS score >= 5
Baseline: Average QUIS=6.79; PASS
Prototype: Average QUIS=7.22; PASS

2.4 Additional Descriptive Statistics

In addition to coding whether or not the participant found the right target answer, some additional descriptive measures were also coded for the following tasks. The baseline and prototypes had unequal sample sizes, so the percentages reflect the number of participants who provided the specific response out of a total of 3 for the baseline condition and out of a total of 9 for the prototype condition. The following percentages are included to provide a general idea of user behavior, but are not intended as inferential statistical tests.

California/Wilmington, DE

Task 1: Did the participant mention anything about the reliability of the estimate without being probed (MOE, CV, color-coded reliability indicator)?
Baseline: 0 (0%)
Prototype: 4 (44%)

Task 2: Did the participant mention reliability, MOE, or CV without being probed?
Baseline: 3 (100 %)
Prototype: 7 (78%)

Task 3: Did the participant mention the reliability, the MOE, or the CV without being probed?
Baseline: 1 (33%)
Prototype: 7 (78%)

Task 4: Did the participant mention reliability, MOE, or CV? [No probing was done for this issue.]
Baseline: 2 (67%)
Prototype: 6 (67%)

Task 5: Did the participant mention the reliability, the MOE, or the CV without being probed?
Baseline: 2 (67%)
Prototype: 9 (100%)

Task 7: Did the participant use the reliability color indicator in their response?
Baseline: NA
Prototype: 8 (89%)

Cook County, IL/Yakima, WA

Task 1: Did the participant mention anything about the reliability of the estimate without being probed (MOE, CV, color-coded reliability indicator)?
Baseline: 0 (0%)
Prototype: 5 (56%)

Task 2: Did the participant mention reliability, MOE, or CV without being probed?
Baseline: 2 (67%)
Prototype: 4 (44%)

Task 3: a) Did the participant mention reliability? [No probing was done for this issue.]
Baseline: 1 (33%)
Prototype: 4 (44%)

AND b) Did the participant mention sample size? [No probing was done for this issue.]
Baseline: 0 (0%)
Prototype: 1 (11%)

Task 4: Did the participant mention reliability, MOE, or CV without being probed?
Baseline: 3 (100%)
Prototype: 9 (100%)

Task 5: Did the participant mention reliability, MOE, or CV without being probed?
Baseline: 2 (67%)
Prototype: 8 (89%)

Task 6: Did the participant mention reliability, MOE, or CV without being probed?
Baseline: 1 (33%)
Prototype: 7 (78%)

Task 7: a) Did the participant use the reliability color indicator in their response?
Baseline: NA
Prototype: 7 (78%)

AND b) Did the participant mention MOE or CV?
Baseline: 1 (33%)
Prototype: 3 (33%)

AND c) Did the participant mention sample size?
Baseline: 0 (0%)
Prototype: 1 (11%)

New York/Longmont, CO

Task 1: Did the participant mention anything about the reliability of the estimate without being probed (MOE, CV, color-coded reliability indicator)?
Baseline: 0 (0%)
Prototype: 8 (89%)

Task 2: a) Did the participant mention the lack of an indicator label in the reliability column?
AND b) Did the participant mention MOE or CV? [No probing was done for this issue.]
Baseline: NA
Prototype: 7 (78%)

Task 3: a) Did the participant mention reliability, MOE, or CV without being probed?
Baseline: 2 (67%)
Prototype: 9 (100%)

AND b) Did the participant mention checking another data source (previous release of ACS, etc.)?
Baseline: 0 (0%)
Prototype: 2 (22%)

Task 4: a) Did the participant mention reliability? [No probing was done for this issue.]
Baseline: 1 (33%)
Prototype: 4 (44%)

AND b) Did the participant mention sample size? [No probing was done for this issue.]
Baseline: 0 (0%)
Prototype: 2 (22%)

Task 5: Did the participant mention reliability, MOE, or CV without being probed?
Baseline: 3 (100%)
Prototype: 9 (100%)

Task 6: Did the participant mention reliability, MOE, or CV without being probed?
Baseline: 2 (67%)
Prototype: 7 (78%)

Task 7: a) Did the participant use the reliability color indicator in their response?
Baseline: NA
Prototype: 9 (100%)

AND b) Did the participant mention MOE or CV? [No probing was done for this issue.]
Baseline: 1 (33%)
Prototype: 2 (22%)

AND c) Did the participant mention sample size? [No probing was done for this issue.]
Baseline: 1 (33%)
Prototype: 5 (56%)

2.5 Expressed Preference During Debriefing

After having seen all three prototypes and the current tables, each participant in both the baseline and prototype conditions, was asked if he or she had a preference for any table (or tables), i.e., which table(s) they liked best. All participants readily indicated that they had a favorite table. No participants preferred the current tables. The following is a breakdown of the self-reported preference for the baseline (BL) and prototype (PT) condition participants:

BLP1: two-level
BLP2: four-level
BLP3: four-level
PTP1: two-level
PTP2: three-level
PTP3: four-level
PTP4: three-level
PTP5: four-level
PTP6: four-level
PTP7: four-level
PTP8: four-level
PTP9: four-level

Two participants said that they preferred the two-level prototype, two participants preferred the three-level prototype, and eight preferred the four-level prototype. This finding is consistent with the results of the satisfaction survey, where participants gave the four-level indicator the highest overall satisfaction rating.

Task 1: What is the First That You Noticed About These Tables?

This section provides a qualitative summary of the table features that each participant reported as the ones that they "first noticed." The results below are reported only for the first table that each participant saw.

BLP1 (New York/Longmont, CO): Household information (bold header)
BLP2 (Cook County/Yakima, WA): There is a lot of information
BLP3 (California/Wilmington, DE): "The bold headers pop out at you"
PTP1 (Cook County/Yakima, WA): Noticed the color-coding (reliability column)
PTP2 (New York/Longmont, CO): Noticed the reliability column and was intrigued. "I am guessing it means use caution if you are going to make a big deal about that particular estimate." "I work with reporters a lot… and this seems like an easy way to tell them to watch out for them."
PTP3 (New York/Longmont, CO): Blank Reliability Column: "Blank for reliability; I kind of hate that, I don't know why it's there then if it's blank."
PTP4 (New York/Longmont, CO): "I noticed the box to the right... and then to the title…and then there is a note." [*Note: This is the only participant that mentioned the box (reliability indicator) for the first table that he/she saw.*]
PTP5 (Cook County/Yakima, WA): Bold Headers, Color-Coded Reliability Column: "The MOE is the first thing that stands out, then I guess the reliability… the color makes it stand out."
PTP6 (California/Wilmington, DE): Bold Headers: Mentioned title of the table then the column headers.
PTP7 (Cook County/Yakima, WA): Color-coded reliability column; also, she thought the legend makes the purpose of the CV clear. She mentioned that she would not know what the CV was without reading the legend.
PTP8 (California/Wilmington, DE): The participant took some time and read over the different items on the table. Liked the bolded header and total. Said, "The first thing that I notice is the green."
PTP9 (California/Wilmington, DE): The participant thought that the reliability column is for standard error at first: "The green column, I have never seen that before, I am not sure what that means." Also, the participant said, "I think it is a little odd having these colors." Participant suggested having the estimate as coded as green instead of having the reliability column. The usability staff acknowledges that color-coding the estimate column itself is not possible because of Section 508 accessibility problems with the PDF version of the tables (e.g., cannot alt tag these automatically or easily as with Web Pages).

Participants frequently mentioned that they were drawn first to the bold columns and color-coded reliability column. The eye-tracking data support the finding that participants

looked at the colorful reliability column before they looked at the reliability indicator box. For the next round of testing, the plan is to move the legend closer to the table or place it above the reliability column.

2.6 Eye-tracking Results

Because of some issues with the computer used to run the Tobii eye tracker in the usability lab (most probably insufficient Random Access Memory (RAM)), not all participant sessions produced usable eye-tracking results. However, those sessions that were successfully recorded led to the following findings:
1) The participants who saw the baseline and two-level prototype table first looked at the bold column headers, especially in the "Selected Social Characteristics in the United States: 2006" column.
2) For participants who saw the three- and four-level tables first, they looked first at the colors in the data reliability indicator column, which means that they noticed and looked at the column before the indicator's legend box at the far right of the screen at the top of the page.

For the current round of testing, qualitative plots of gaze path (also called a "scan path" in the literature) and heatmaps of fixations were available as results from the eye-tracking analysis. Fixations are areas where a person focused visual attention on a specific area of the screen for approximately 200 – 300 milliseconds (Pan et al., 2004).

With the gaze-path plots, the fixations are numbered in the temporal order in which they are made. Fixations with a larger circumference indicate longer fixations. For the heatmaps, red areas indicate areas where participants fixated most frequently, while green areas indicate areas that were fixated upon, but not as frequently as the red areas. Unmarked areas are regions where the participants did not fixate their eyes. The red Xs indicate where participants clicked using the mouse. More quantitative measures for analysis of eye-tracking data are currently being developed by the lead author of this report.

Baseline and Two-Level Prototype: Bold Headings

When asked what they noticed first about the tables (Task 1), several of the participants in the baseline condition as well as participants who saw the two-level indicator first reported that their eyes were first drawn to the bold headings in the tables. The eye-tracking results are consistent with their self-report results for Task 1. Figure 4 shows the gaze pattern (e.g., series of locations where the participant fixated) for a participant in the baseline condition. The first fixation recorded, fixation 6, is on the geographical title of Yakima, WA (subsequent fixations are numbered with increasing numerals following this fixation). Then the participant looks at the "Households by type," then at or near the "Total households" header, then at the estimate column, then at the "NOTE" above the table, then again at "HOUSEHOLDS BY TYPE," and then at the contents of the first few rows of the estimate and margin of error columns. The larger the circle representing a

fixation, the longer the fixation duration was. Although the participant fixated on the MOE column contents here, they did not look at the title of the column first and only fixated on this column briefly.

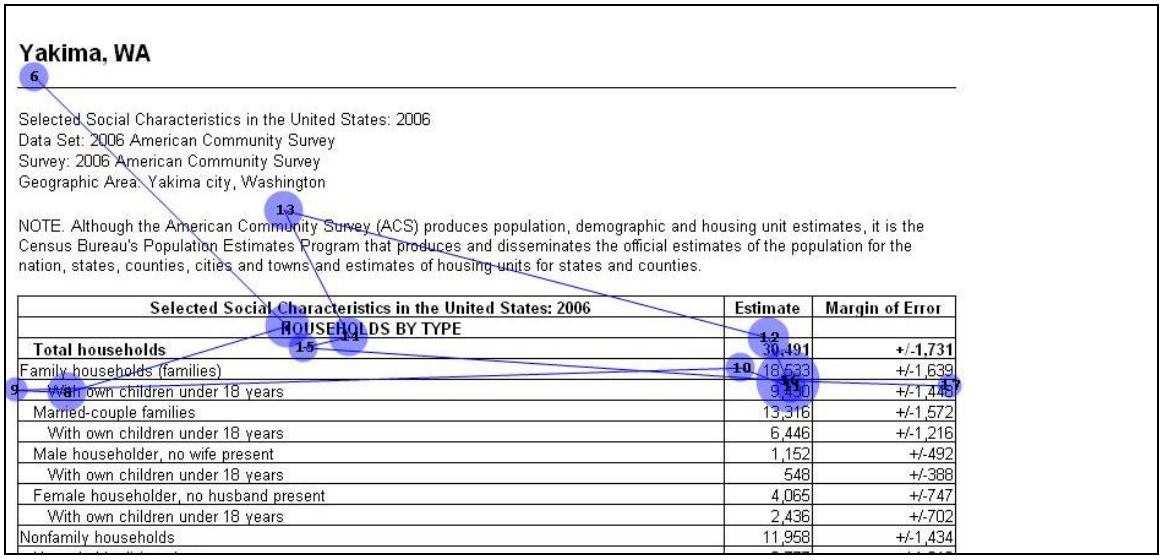
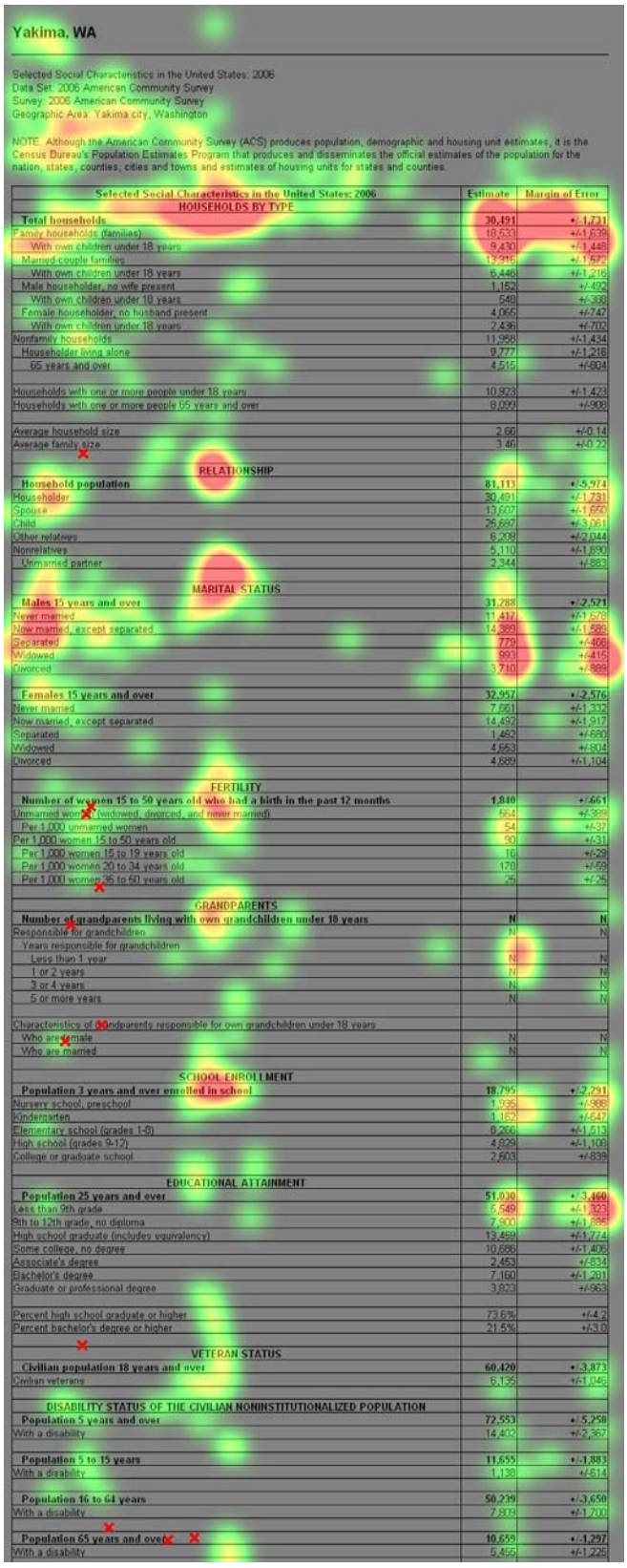


Figure 4: Gaze path of fixations of the first ten seconds of Task 1 for Baseline Participant 2, who saw this table first. This participant looked mainly at the bold headings.

From Figure 4, it is evident that this participant was drawn to the bold headers and top content of the table. Current users will expect that important information will be located in this area. The baseline participants did not look in the blank region at the upper right corner of the table where the data reliability indicator legend was placed for the prototype tables.

Figure 5 shows the heatmap for the composite sessions of Baseline Participants 2 and 3 for this table. This figure includes the heat map of the entire session of tasks for both participants with this table instead of just the first ten seconds. The red areas indicate regions where the participants looked (e.g., fixated) most often.



Yakima, WA

Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Yakima city, Washington

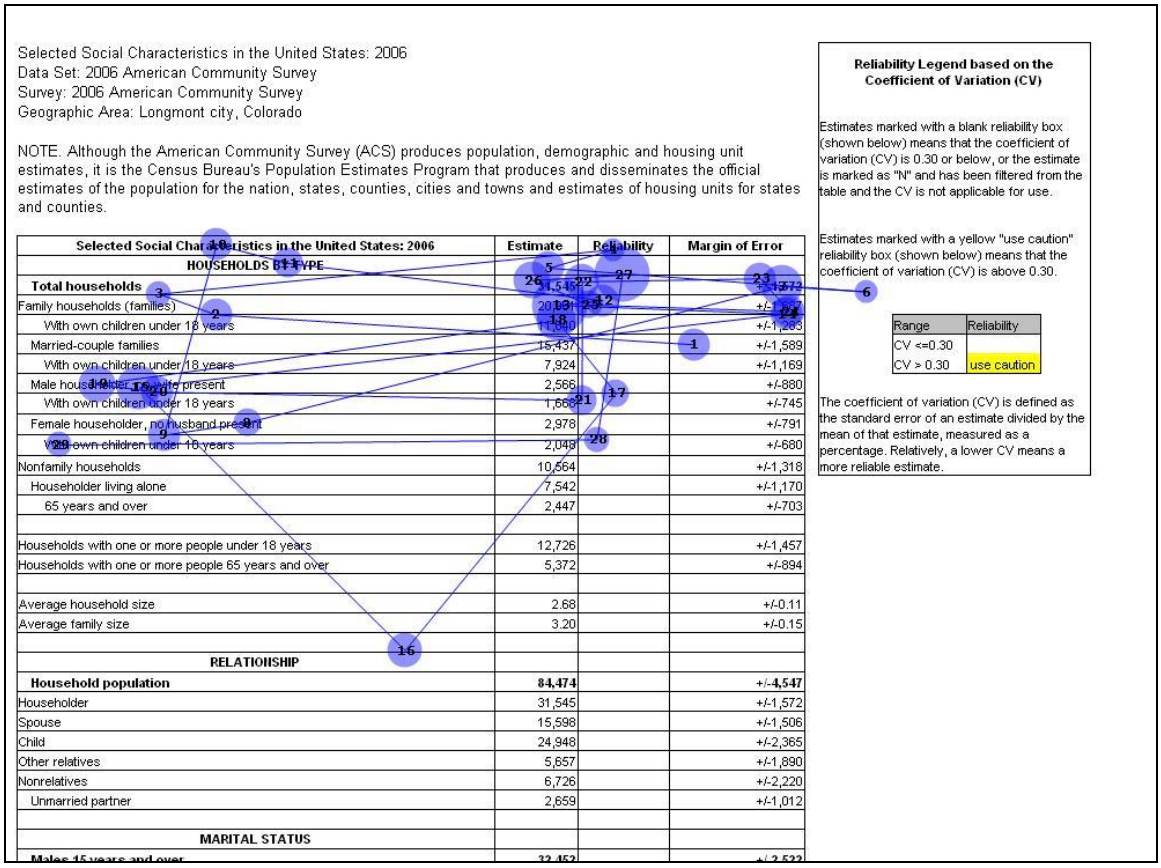
NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Selected Social Characteristics in the United States: 2006	Estimate	Margin of Error
HOUSEHOLDS BY TYPE		
Total households	**30,491**	**+/-1,731**
Family households (families)	18,533	+/-1,639
With own children under 18 years	9,430	+/-1,448
Married-couple families	13,316	+/-1,579
With own children under 18 years	6,446	+/-1,216
Male householder, no wife present	1,152	+/-492
With own children under 18 years	548	+/-388
Female householder, no husband present	4,065	+/-747
With own children under 18 years	2,436	+/-702
Nonfamily households	11,958	+/-1,434
Householder living alone	9,777	+/-1,216
65 years and over	4,515	+/-804
Households with one or more people under 18 years	10,923	+/-1,423
Households with one or more people 65 years and over	8,099	+/-908
Average household size	2.66	+/-0.14
Average family size	3.46	+/-0.22
RELATIONSHIP		
Household population	**81,113**	**+/-5,974**
Householder	30,491	+/-1,731
Spouse	13,607	+/-1,660
Child	25,687	+/-3,061
Other relatives	6,209	+/-2,044
Nonrelatives	5,110	+/-1,690
Unmarried partner	2,344	+/-883
MARITAL STATUS		
Males 15 years and over	**31,288**	**+/-2,521**
Never married	11,412	+/-1,678
Now married, except separated	14,389	+/-1,589
Separated	779	+/-406
Widowed	993	+/-415
Divorced	3,710	+/-888
Females 15 years and over	**32,957**	**+/-2,576**
Never married	7,661	+/-1,332
Now married, except separated	14,492	+/-1,917
Separated	1,462	+/-680
Widowed	4,653	+/-804
Divorced	4,689	+/-1,104
FERTILITY		
Number of women 15 to 50 years old who had a birth in the past 12 months	**1,840**	**+/-661**
Unmarried women (widowed, divorced, and never married)	664	+/-369
Per 1,000 unmarried women	54	+/-37
Per 1,000 women 15 to 50 years old	90	+/-31
Per 1,000 women 15 to 19 years old	16	+/-29
Per 1,000 women 20 to 34 years old	170	+/-69
Per 1,000 women 35 to 50 years old	25	+/-26
GRANDPARENTS		
Number of grandparents living with own grandchildren under 18 years	**N**	**N**
Responsible for grandchildren	N	N
Years responsible for grandchildren		
Less than 1 year	N	N
1 or 2 years	N	N
3 or 4 years	N	N
5 or more years	N	N
Characteristics of grandparents responsible for own grandchildren under 18 years		
Who are female	N	N
Who are married	N	N
SCHOOL ENROLLMENT		
Population 3 years and over enrolled in school	**18,795**	**+/-2,791**
Nursery school, preschool	1,035	+/-988
Kindergarten	1,162	+/-647
Elementary school (grades 1-8)	8,266	+/-1,513
High school (grades 9-12)	4,829	+/-1,108
College or graduate school	2,603	+/-838
EDUCATIONAL ATTAINMENT		
Population 25 years and over	**51,030**	**+/-3,160**
Less than 9th grade	5,549	+/-1,323
9th to 12th grade, no diploma	7,900	+/-1,895
High school graduate (includes equivalency)	13,459	+/-1,774
Some college, no degree	10,686	+/-1,406
Associate's degree	2,453	+/-834
Bachelor's degree	7,160	+/-1,291
Graduate or professional degree	3,823	+/-963
Percent high school graduate or higher	73.6%	+/-4.2
Percent bachelor's degree or higher	21.5%	+/-3.0
VETERAN STATUS		
Civilian population 18 years and over	**60,420**	**+/-3,873**
Civilian veterans	6,135	+/-1,046
DISABILITY STATUS OF THE CIVILIAN NONINSTITUTIONALIZED POPULATION		
Population 5 years and over	**72,553**	**+/-5,258**
With a disability	14,402	+/-2,367
Population 5 to 15 years	**11,655**	**+/-1,883**
With a disability	1,138	+/-614
Population 16 to 64 years	**50,239**	**+/-3,650**
With a disability	7,809	+/-1,700
Population 65 years and over	**10,659**	**+/-1,297**
With a disability	5,455	+/-1,225

Figure 5: Heatmap of top portion of baseline table for Yakima, WA (Baseline Participants 2 & 3)

Consistent with the results of Task 1 for the two-level indicator where participants did not immediately notice the color-coding, they also did not look at the "reliability" column or the indicator legend as they did for the three- and four-level indicators. Instead, they too looked at the bold columns. The eye-tracking results of participants initially looking at the two-level indicator are similar to the results for the baseline tables in that the bold headings captured their attention first.

Figure 6 shows the first ten seconds of fixations from Participant 3, who saw this two-level indicator table from Longmont, CO first. This participant mainly fixates on the headers and column content of the table, and fixates only once briefly inside of the indicator legend box (fixation number 6). However, this fixation is not near the colored "use caution" portion of the indicator, and the participant did not appear to read any of the text during this time. Instead, the participant's eyes return immediately to the "Margin of Error" column content (fixation 7), then to the "Households By Type" column content, then back to the bold "Selected Social Characteristics in the United States: 2006" column, then proceed to look at more headers at the top of the table and also the bold "Relationship" header further down the page.



Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Longmont city, Colorado

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households			
Family households (families)	20,193		+/-1,373
With own children under 18 years	11,080		+/-1,283
Married-couple families	15,437		+/-1,589
With own children under 18 years	7,924		+/-1,169
Male householder, no wife present	2,566		+/-880
With own children under 18 years	1,568		+/-745
Female householder, no husband present	2,978		+/-791
With own children under 18 years	2,048		+/-680
Nonfamily households	10,564		+/-1,318
Householder living alone	7,542		+/-1,170
65 years and over	2,447		+/-703
Households with one or more people under 18 years	12,726		+/-1,457
Households with one or more people 65 years and over	5,372		+/-894
Average household size	2.68		+/-0.11
Average family size	3.20		+/-0.15
RELATIONSHIP			
Household population	**84,474**		**+/-4,547**
Householder	31,545		+/-1,572
Spouse	15,598		+/-1,506
Child	24,948		+/-2,365
Other relatives	5,657		+/-1,890
Nonrelatives	6,726		+/-2,220
Unmarried partner	2,659		+/-1,012
MARITAL STATUS			
Males 15 years and over	32,452		+/-2,522

Reliability Legend based on the Coefficient of Variation (CV)

Estimates marked with a blank reliability box (shown below) means that the coefficient of variation (CV) is 0.30 or below, or the estimate is marked as "N" and has been filtered from the table and the CV is not applicable for use.

Estimates marked with a yellow "use caution" reliability box (shown below) means that the coefficient of variation (CV) is above 0.30.

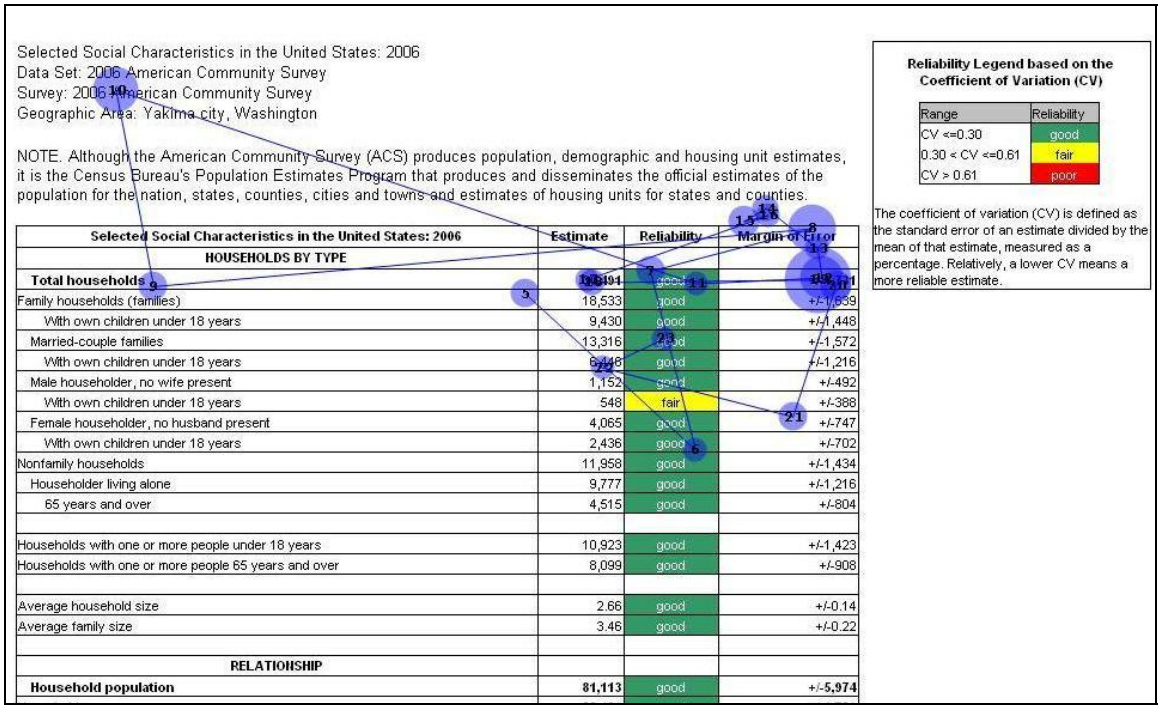
Range	Reliability
CV <=0.30	
CV > 0.30	use caution

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Figure 6: Gaze plot for Longmont, CO two-level indicator (Participant 3, who saw this table first)

Prototype Tables: Three- and Four-level Indicators

Consistent with the self-report results from Task 1 for these tables, participants who saw the three- or four-level indicator first tended to still look at the bold headings, but also at the color-coded "Reliability" column with the first few fixations, often quite some time before they looked at the data reliability indicator legend box for an interpretation of these colors. Figure 7 shows the first ten seconds from Participant 5, who saw the three-level indicator Yakima, WA table first.



Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Yakima city, Washington

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Reliability Legend based on the Coefficient of Variation (CV)

Range	Reliability
CV <=0.30	good
0.30 < CV <=0.61	fair
CV > 0.61	poor

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	19,191	good	+/-1,171
Family households (families)	18,533	good	+/-1,639
With own children under 18 years	9,430	good	+/-1,448
Married-couple families	13,316	good	+/-1,572
With own children under 18 years	5,446	good	+/-1,216
Male householder, no wife present	1,152	good	+/-492
With own children under 18 years	548	fair	+/-388
Female householder, no husband present	4,065	good	+/-747
With own children under 18 years	2,436	good	+/-702
Nonfamily households	11,958	good	+/-1,434
Householder living alone	9,777	good	+/-1,216
65 years and over	4,515	good	+/-804
Households with one or more people under 18 years	10,923	good	+/-1,423
Households with one or more people 65 years and over	8,099	good	+/-908
Average household size	2.66	good	+/-0.14
Average family size	3.46	good	+/-0.22
RELATIONSHIP			
Household population	81,113	good	+/-5,974

Figure 7: First ten seconds of gaze fixations for Participant 5 (who saw this three-level indicator table first)

This participant looked at the color-coded "Reliability" column, the bold headings at the top of the table, the "Margin of Error" column header and content and the year of the data (2006) within the first 10 seconds of seeing the table. However, this participant did not look at the Data Reliability Indicator legend box within this initial short period of time. This pattern is consistent with the general trend of participants' noticing the color-coded column some time before they notice and/or use the legend box.

Figure 8 shows the first ten seconds of gaze fixations from Participant 6, who saw this four-level indicator first.

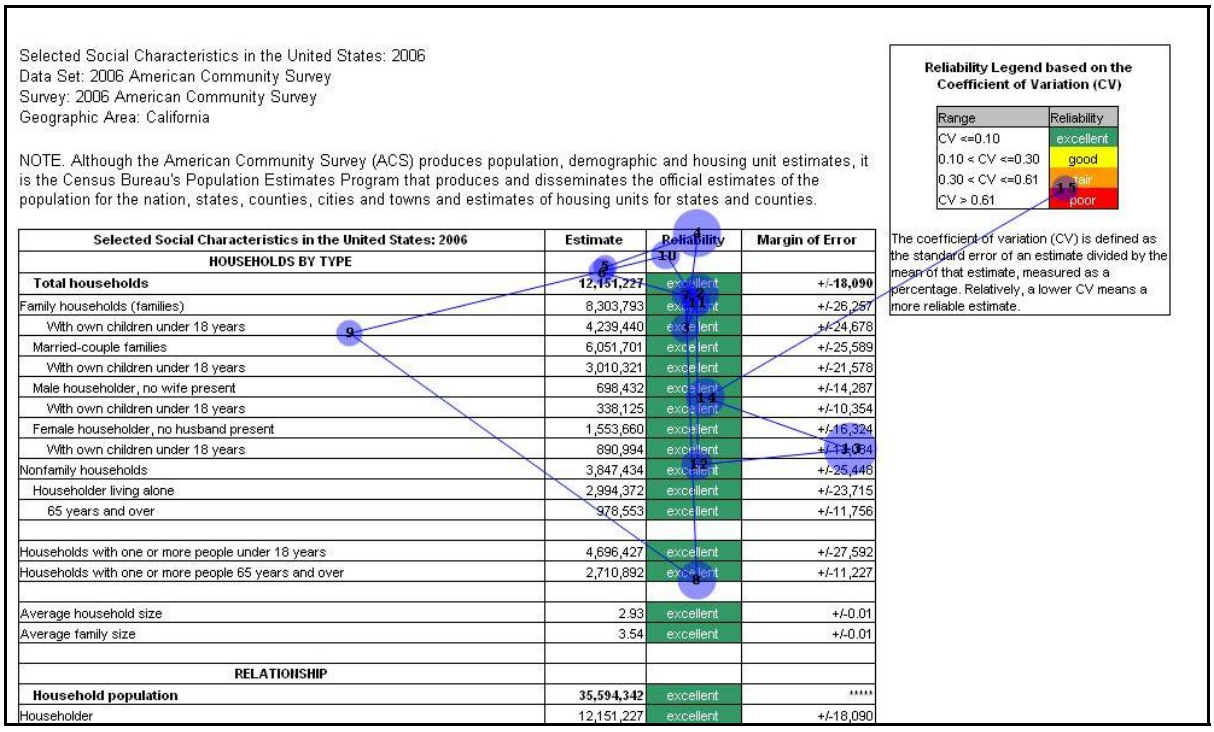


Figure 8: First ten seconds of gaze fixations for Participant 6 (who saw this four-level indicator table first)

Figure 8 shows that Participant 6, who saw this four-level indicator table first, made an initial fixation (number 4) on the bold "Reliability" header; then the next two fixations are on the bold estimate at the top of the "Estimate" column for "Total Households" (i.e., 12,551,227, fixations 5 and 6). Then, the participant fixates on the colored content of the "Reliability" column (fixations 7 and 8). After this, the person's gaze stops briefly inside the white area of the "Selected Social Characteristics in the United States: 2006" column (fixation 9) en route to another fixation on the bold "reliability" header (fixation 10). The participant then fixates again on the colored content of the "Reliability" column (fixations 12), then the "Margin of Error" column (fixation 13), then the Data Reliability Indicator legend box. Again, the bold headers and colored column seem to attract the user's attention before the indicator legend box does.

Composite heatmaps of the three-and four-level indicator tables show similar patterns. Figures 9-12 show the overall composite heatmaps with 6 participants apiece for the three- and four- level indicators.

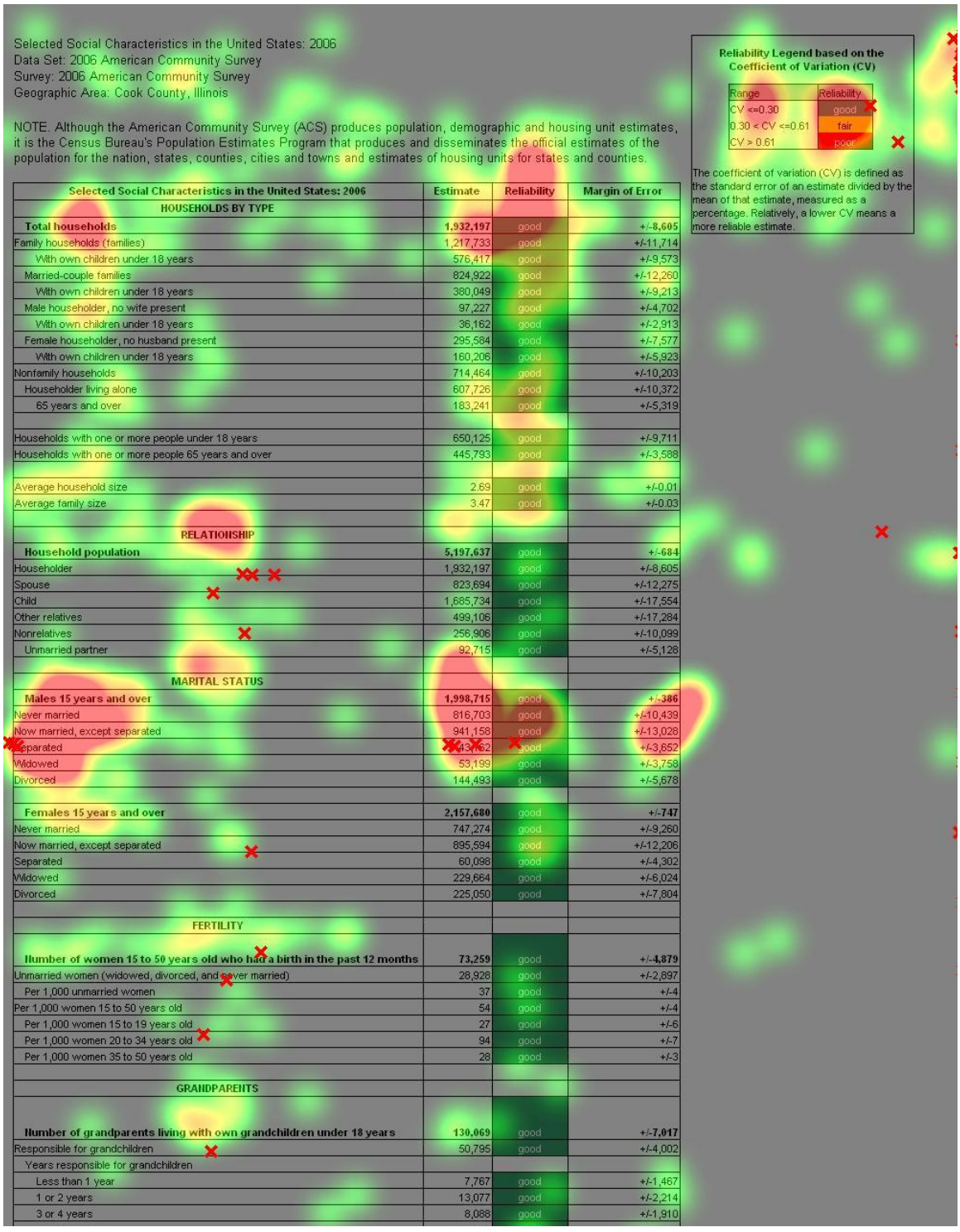
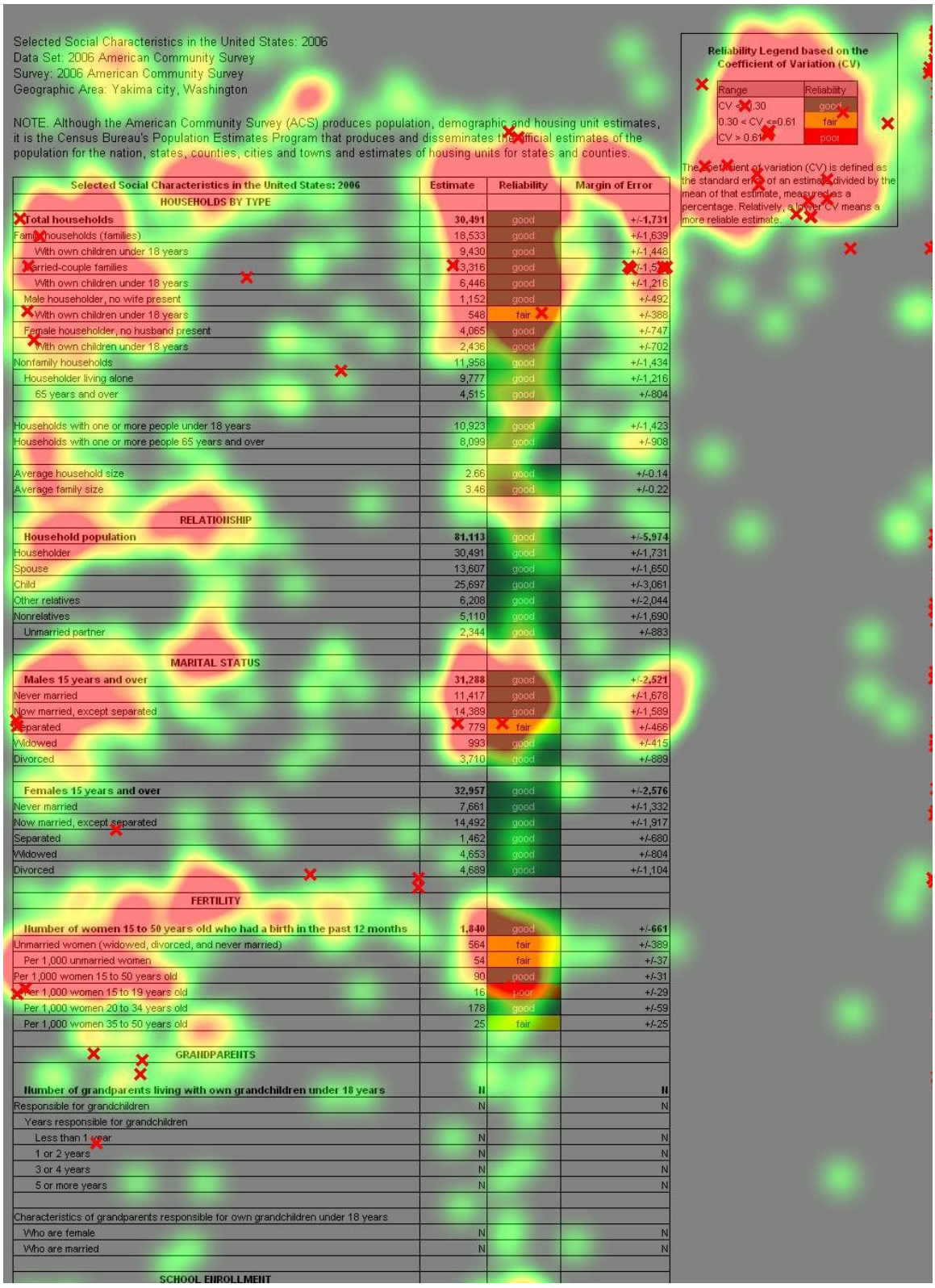


Figure 9: Top section of composite heatmap from six participants' eye-tracking data on the three-level Cook County table (over entire set of tasks for this table).

Figure 9 shows that all six people who had usable eye-tracking data for the three-level Cook County, IL table (not just those who saw this table first) looked at the bold

headings, “Reliability” column, and the Data Reliability Indicator legend box. In comparison to the baseline composite heatmap plot in Figure 5, it is apparent that participants look less at the “Margin of Error” column and instead look at the “Reliability” column for information (although they still do look at this column intermittently). This is consistent with High-Priority Usability Finding 3 (pg. 33), which states that participants stopped using the MOE as they got used to using the Data Reliability Indicator for the prototypes. All participants were probed about reliability and/or MOE if they did not mention either indicator in their response to Task 1 for the first table they saw. So, they were given information that would lead them to look at MOE and reliability over the course of the test.

Figure 10, which shows the composite heatmap for 6 participants on the Yakima, WA three-level indicator table, shows a similar pattern.



Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Yakima city, Washington

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Reliability Legend based on the Coefficient of Variation (CV)

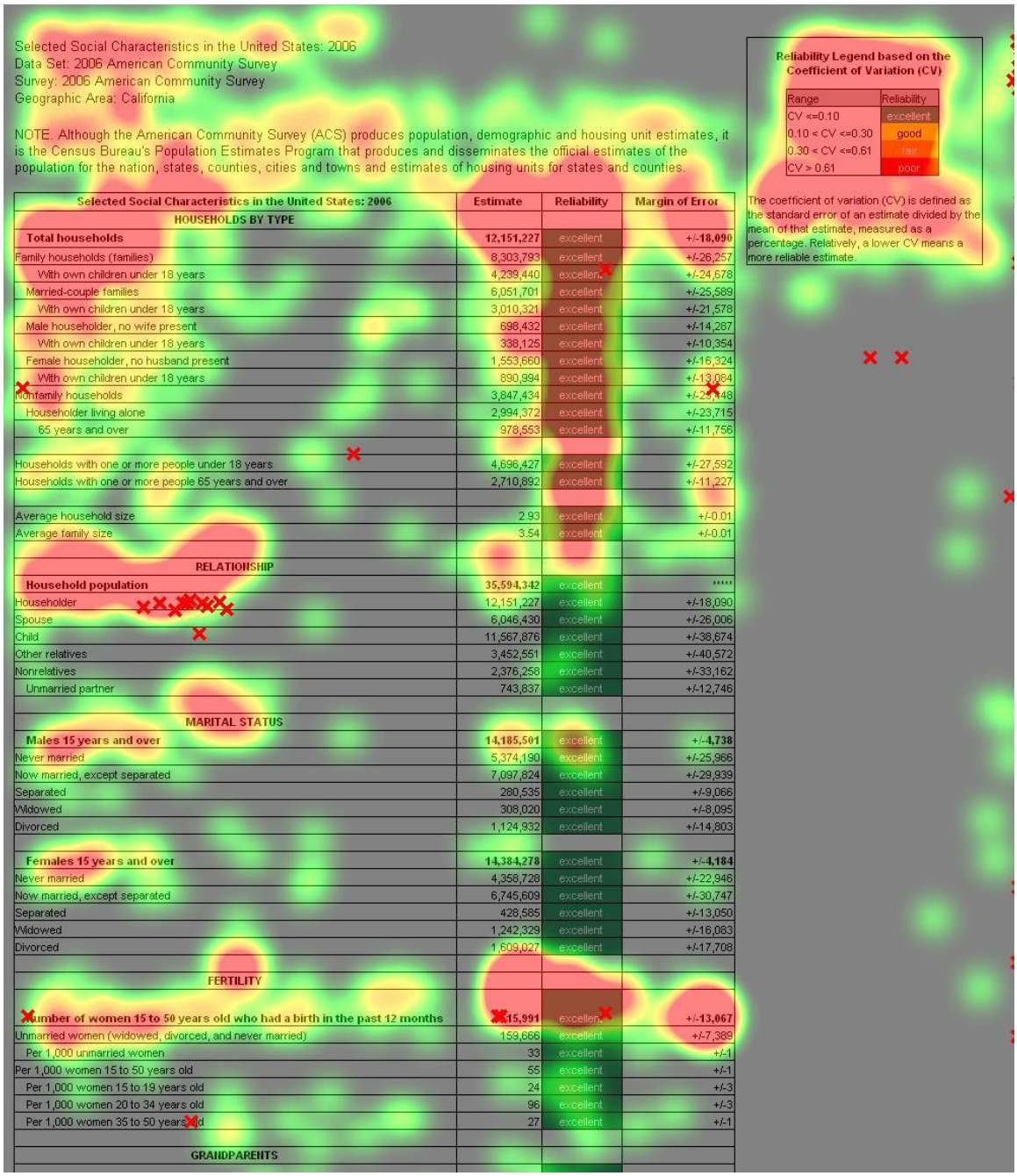
Range	Reliability
CV <= 0.30	good
0.30 < CV <=0.61	fair
CV > 0.61	poor

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**30,491**	good	**+/-1,731**
Family households (families)	18,533	good	+/-1,639
With own children under 18 years	9,430	good	+/-1,448
Married-couple families	13,316	good	+/-1,5
With own children under 18 years	6,446	good	+/-1,216
Male householder, no wife present	1,152	good	+/-492
With own children under 18 years	548	fair	+/-388
Female householder, no husband present	4,065	good	+/-747
With own children under 18 years	2,436	good	+/-702
Nonfamily households	11,958	good	+/-1,434
Householder living alone	9,777	good	+/-1,216
65 years and over	4,515	good	+/-804
Households with one or more people under 18 years	10,923	good	+/-1,423
Households with one or more people 65 years and over	8,099	good	+/-908
Average household size	2.66	good	+/-0.14
Average family size	3.46	good	+/-0.22
RELATIONSHIP			
Household population	**81,113**	good	**+/-5,974**
Householder	30,491	good	+/-1,731
Spouse	13,607	good	+/-1,650
Child	25,697	good	+/-3,061
Other relatives	6,208	good	+/-2,044
Nonrelatives	5,110	good	+/-1,690
Unmarried partner	2,344	good	+/-883
MARITAL STATUS			
Males 15 years and over	**31,288**	good	**+/-2,521**
Never married	11,417	good	+/-1,678
Now married, except separated	14,389	good	+/-1,589
Separated	779	fair	+/-466
Widowed	993	good	+/-415
Divorced	3,710	good	+/-889
Females 15 years and over	**32,957**	good	**+/-2,576**
Never married	7,661	good	+/-1,332
Now married, except separated	14,492	good	+/-1,917
Separated	1,462	good	+/-680
Widowed	4,653	good	+/-804
Divorced	4,689	good	+/-1,104
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**1,840**	good	**+/-661**
Unmarried women (widowed, divorced, and never married)	564	fair	+/-389
Per 1,000 unmarried women	54	fair	+/-37
Per 1,000 women 15 to 50 years old	90	good	+/-31
Per 1,000 women 15 to 19 years old	16	poor	+/-29
Per 1,000 women 20 to 34 years old	178	good	+/-59
Per 1,000 women 35 to 50 years old	25	fair	+/-25
GRANDPARENTS			
Number of grandparents living with own grandchildren under 18 years	**N**		**N**
Responsible for grandchildren	N		N
Years responsible for grandchildren			
Less than 1 year	N		N
1 or 2 years	N		N
3 or 4 years	N		N
5 or more years	N		N
Characteristics of grandparents responsible for own grandchildren under 18 years			
Who are female	N		N
Who are married	N		N
SCHOOL ENROLLMENT			

Figure 10: Top section of composite heatmap from six participants' eye-tracking data on the three-level Yakima, WA table (over entire set of tasks for this table).

Figures 11 and 12 show the eye-tracking results for four-level California and Wilmington, DE prototype tables.



Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: California

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

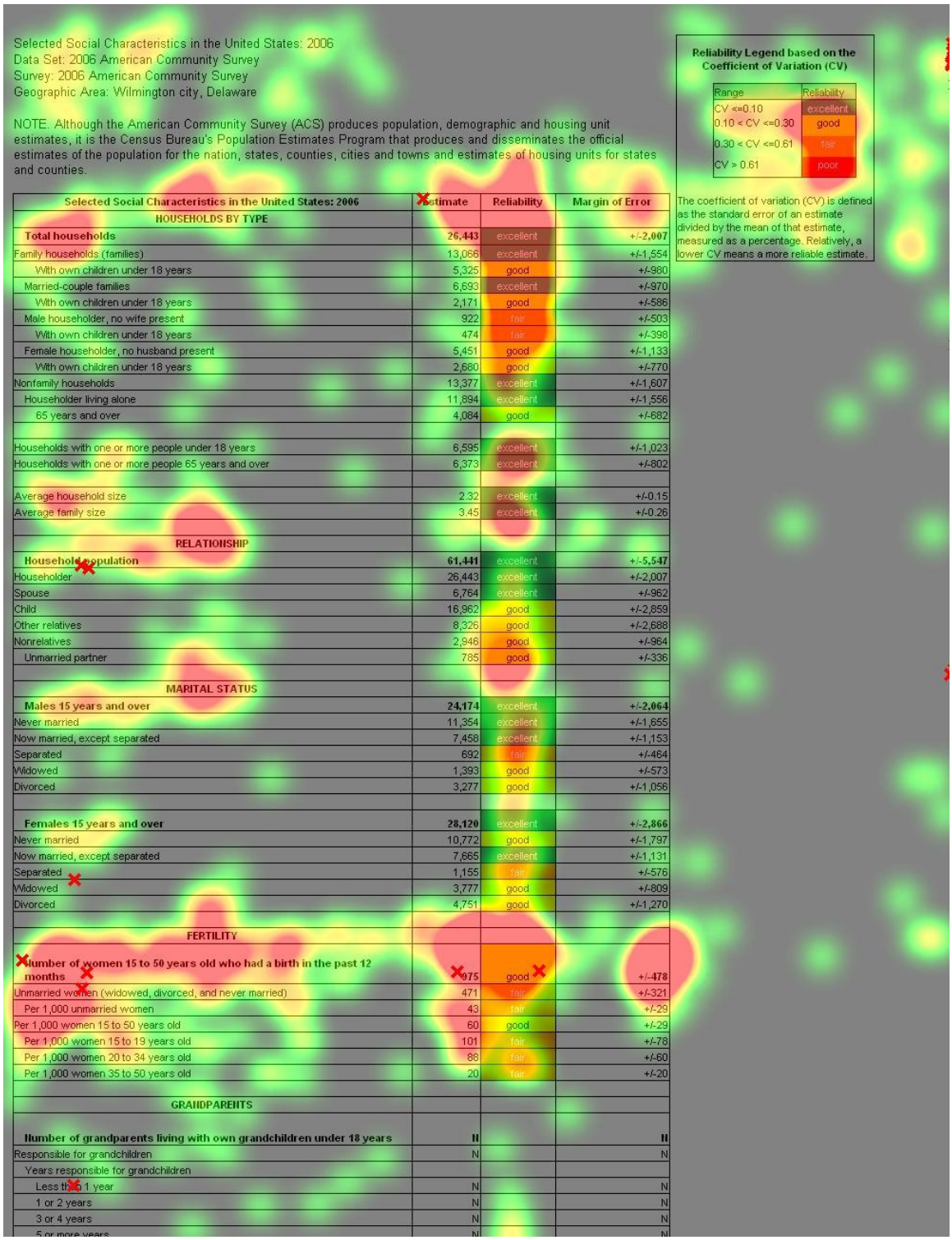
Reliability Legend based on the Coefficient of Variation (CV)

Range	Reliability
CV <=0.10	excellent
0.10 < CV <=0.30	good
0.30 < CV <=0.61	fair
CV > 0.61	poor

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**12,151,227**	excellent	**+/-18,090**
Family households (families)	8,303,793	excellent	+/-26,257
With own children under 18 years	4,239,440	excellent	+/-24,678
Married-couple families	6,051,701	excellent	+/-25,589
With own children under 18 years	3,010,321	excellent	+/-21,578
Male householder, no wife present	698,432	excellent	+/-14,287
With own children under 18 years	338,125	excellent	+/-10,354
Female householder, no husband present	1,553,660	excellent	+/-16,324
With own children under 18 years	890,994	excellent	+/-13,084
Nonfamily households	3,847,434	excellent	+/-23,348
Householder living alone	2,994,372	excellent	+/-23,715
65 years and over	978,553	excellent	+/-11,756
Households with one or more people under 18 years	4,696,427	excellent	+/-27,592
Households with one or more people 65 years and over	2,710,892	excellent	+/-11,227
Average household size	2.93	excellent	+/-0.01
Average family size	3.54	excellent	+/-0.01
RELATIONSHIP			
Household population	**35,594,342**	excellent	*****
Householder	12,151,227	excellent	+/-18,090
Spouse	6,046,430	excellent	+/-26,006
Child	11,567,876	excellent	+/-38,674
Other relatives	3,452,551	excellent	+/-40,572
Nonrelatives	2,376,258	excellent	+/-33,162
Unmarried partner	743,837	excellent	+/-12,746
MARITAL STATUS			
Males 15 years and over	**14,185,501**	excellent	**+/-4,738**
Never married	5,374,190	excellent	+/-25,966
Now married, except separated	7,097,824	excellent	+/-29,939
Separated	280,535	excellent	+/-9,066
Widowed	308,020	excellent	+/-8,095
Divorced	1,124,932	excellent	+/-14,803
Females 15 years and over	**14,384,278**	excellent	**+/-4,184**
Never married	4,358,728	excellent	+/-22,946
Now married, except separated	6,745,609	excellent	+/-30,747
Separated	428,585	excellent	+/-13,050
Widowed	1,242,329	excellent	+/-16,083
Divorced	1,609,027	excellent	+/-17,708
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**15,991**	excellent	**+/-13,067**
Unmarried women (widowed, divorced, and never married)	159,666	excellent	+/-7,389
Per 1,000 unmarried women	33	excellent	+/-1
Per 1,000 women 15 to 50 years old	55	excellent	+/-1
Per 1,000 women 15 to 19 years old	24	excellent	+/-3
Per 1,000 women 20 to 34 years old	96	excellent	+/-3
Per 1,000 women 35 to 50 years old	27	excellent	+/-1
GRANDPARENTS			

Figure 11: Top Section of composite heatmap from six participants' eye-tracking data on the four-level California table (over entire set of tasks for this table).



Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Wilmington city, Delaware

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Reliability Legend based on the Coefficient of Variation (CV)

Range	Reliability
CV <=0.10	excellent
0.10 < CV <=0.30	good
0.30 < CV <=0.61	fair
CV > 0.61	poor

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**26,443**	excellent	**+/-2,007**
Family households (families)	13,066	excellent	+/-1,554
With own children under 18 years	5,325	good	+/-980
Married-couple families	6,693	excellent	+/-970
With own children under 18 years	2,171	good	+/-586
Male householder, no wife present	922	fair	+/-503
With own children under 18 years	474	fair	+/-398
Female householder, no husband present	5,451	good	+/-1,133
With own children under 18 years	2,680	good	+/-770
Nonfamily households	13,377	excellent	+/-1,607
Householder living alone	11,894	excellent	+/-1,556
65 years and over	4,084	good	+/-682
Households with one or more people under 18 years	6,595	excellent	+/-1,023
Households with one or more people 65 years and over	6,373	excellent	+/-802
Average household size	2.32	excellent	+/-0.15
Average family size	3.45	excellent	+/-0.26
RELATIONSHIP			
Household population	**61,441**	excellent	**+/-5,547**
Householder	26,443	excellent	+/-2,007
Spouse	6,764	excellent	+/-962
Child	16,962	good	+/-2,859
Other relatives	8,326	good	+/-2,688
Nonrelatives	2,946	good	+/-964
Unmarried partner	785	good	+/-336
MARITAL STATUS			
Males 15 years and over	**24,174**	excellent	**+/-2,064**
Never married	11,354	excellent	+/-1,655
Now married, except separated	7,458	excellent	+/-1,153
Separated	692	fair	+/-464
Widowed	1,393	good	+/-573
Divorced	3,277	good	+/-1,056
Females 15 years and over	**28,120**	excellent	**+/-2,866**
Never married	10,772	good	+/-1,797
Now married, except separated	7,665	excellent	+/-1,131
Separated	1,155	fair	+/-576
Widowed	3,777	good	+/-809
Divorced	4,751	good	+/-1,270
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**975**	good	**+/-478**
Unmarried women (widowed, divorced, and never married)	471	fair	+/-321
Per 1,000 unmarried women	43	fair	+/-29
Per 1,000 women 15 to 50 years old	60	good	+/-29
Per 1,000 women 15 to 19 years old	101	fair	+/-78
Per 1,000 women 20 to 34 years old	88	fair	+/-60
Per 1,000 women 35 to 50 years old	20	fair	+/-20
GRANDPARENTS			
Number of grandparents living with own grandchildren under 18 years	**N**		**N**
Responsible for grandchildren	N		N
Years responsible for grandchildren			
Less than 1 year	N		N
1 or 2 years	N		N
3 or 4 years	N		N
5 or more years	N		N

Figure 12: Top section of composite heatmap from six participants' eye-tracking data on the four-level Wilmington, DE table (over entire set of tasks for this table).

Overall, it is apparent that users do notice and use the color-coded "Reliability" column and Data Reliability Indicator legend box, although they do not tend to read the legend box as one of their first steps when using the tables.

Recommendations Based on Eye-tracking Data

The Indicator box might be better located above the color-coded "reliability" column because users of the current tables tend to favor the area at the top of the table and just above it. In an ideal setting, a legend box that followed the user's scrolling would be preferable, but it may also be helpful for future testing to include the legend at the top and bottom of the page. Also, the heatmaps show that the area in between the descriptive label for the estimate on the left and the estimate for the right is wasted space because participants did not fixate there. This is consistent with several participants' suggestion of adding a shaded bar or color scheme for easier tracking across the table because it is difficult to match the estimate with its corresponding description because of the large horizontal blank space between them. Again, we recommend that the longer titles (such as "DISABILITY STATUS OF THE NONINSTITUTIONALIZED CIVILIAN POPULATION") be wrapped to two lines so that this column width can be lessened. Also, shortening titles like this for easier reading and usability (e.g., writing for the Web (Redish, 2007)) is recommended for future versions of these tables.

3.0 Qualitative Results: Potential Usability Issues

Ranking System

The issues identified in this document have been grouped according to the following ranking system:

- High Priority - These items are likely to prevent users from accomplishing their tasks. These are also items that have been specifically mentioned as areas to focus on in the review of the Web site.
- Medium Priority - These items are likely to cause navigation errors or to cause performance issues resulting in additional time required for users to complete tasks.
- Low Priority - These items are not absolutely connected to user performance, but are important points to be considered in further development of the site.

3.1 High-Priority Findings and Recommendations:

Findings are based on observations made during usability testing within the context of known best practices for online user-interface design.

1. WIDTH OF ESTIMATE DESCRIPTION COLUMN

Observations: Several participants in both the prototype and baseline conditions mentioned that it was difficult to keep track of which number went with each description (e.g.., column labeled "Selected Social Characteristics in the United States: 2006")

because there was so much horizontal distance and white space separating them. Indeed, the width of the column depends on the length of the longest category name both in the current ACS tables and the prototypes. A few participants in the prototype condition mentioned that they wished the color went all the way across the column for the smaller geographies because it would help them keep track of the categories (although it should be noted that this type of shading may be impossible because of issues with Section 508 accessibility compliance). One participant suggested adding gray or colored bands every five lines or so to aid in the horizontal scanning. Another participant suggested that the longer descriptions be wrapped to the second line so that the column width would be reduced.

As a note, the width of this descriptive column (see Appendix I) was reduced slightly for the prototypes in order to convert the Excel file version of the table into an HTML Web site that could be used with the Usability Lab's Tobii eyetracker and to allow the legend box to appear on the screen without the need for horizontal scrolling. However, this slight reduction did not resolve the horizontal scrolling problem. Several participants physically moved the mouse across the screen from left to right when locating the estimate for a particular category.

Usability Issue: It is difficult for users to keep track of which estimate is associated with each description. This could result in novice users reporting the wrong estimate from a row above or below the intended estimate.

Recommendations: We recommend reducing the width of the description column substantially. This will allow users to scan the tables more quickly and may reduce instances of users misreporting estimates because they are looking at the incorrect row in the table. We recommend that a version of the tables with a much narrower description column be included in future rounds of testing to allow for a comparison of efficiency scores for the tasks. The usability team acknowledges that changing the width of the ACS data tables is out of scope for the current project and that the sponsor has mentioned this issue to the staff in charge of these tables for possible future revision.

2. BASELINE TABLES: USE OF MOE

Observations: The baseline participants mainly noted the geographic location associated with the table and then scrolled down and mentioned the bolded headers for the different sections of categories. When asked in Task 1 to report the first things they noticed about the tables, many did not mention the MOE until prompted. The prototype participants were able to complete the tasks more quickly than the baseline participants, and they also more frequently mentioned the indicator, margin of error or coefficient of variation, in their responses. During debriefing, the baseline participants who were presented with the three- and four-level indicators first almost always mentioned that their eye was drawn to the colors first. The use of a reliability indicator does seem to encourage users to use and/or report some measure of error along with the estimate they choose. This finding is consistent with previous eye-tracking research that showed that while users scanned

yellow page advertisements, those with color were perceived more quickly, more often, and longer than black-and-white text ads without color (Lohse, 1997).

Usability Issues: The baseline participants took longer overall to complete the tasks than did the prototype participants. The baseline participants did not always report a measure of reliability with their response.

Recommendations: We recommend the inclusion of a reliability indicator for future releases of the ACS data tables. However, the use of CV as the scale of reliability can be confusing when MOE is the measure of error listed in the table itself. In future testing, it might be of interest to have a column of the actual CV instead of calculating it behind the scenes. Additionally, concerns have been expressed about the use of CV for this type of estimate (Weidman, 2008). Further research and testing is recommended to examine this issue.

3. PROTOTYPE TABLES: USE OF MOE

Observations: One side effect of the testing design was that, as participants became accustomed to using the prototypes, many of them stopped reporting the MOE and instead relied only on the indicator column.

Usability Issues: Many participants stopped using the MOE and relied completely on the reliability column. While this finding suggests that participants found the indicator helpful, it also suggests that people may use the CV indicator and still fail to report the MOE even if both are provided. This can be problematic if the user has misinterpreted the meaning of the indicator (see High Priority Issue number 4, below).

Recommendations: Explain somewhere on the table, in brief text, the relationship between CV and MOE. Further testing of this issue is recommended.

4. TWO-LEVEL INDICATOR

Observations: Participants overall tended to take longer to understand the coding scheme for the two-level indicator prototype than they did for the other two prototypes. Several participants commented that they did not like the blank column entries for estimates with acceptable reliability. One participant commented, “Blank for reliability. I kind of hate that; I don’t know why it's there then if it’s blank.” Another participant had a low efficiency score for all of the tasks for the two-level indicator because he misinterpreted the meaning of the “use caution” message and did not read the legend. A few participants noted that the white-to-yellow contrast was not pleasing to the eye or that there was not enough contrast between the colors.

Usability Issues: One participant interpreted the two-level indicator as having the opposite meaning to what was intended. That is, he thought that the “use caution” estimates were more reliable than the unmarked estimates because they were marked with yellow. Several participants initially thought that the blank entries in the reliability

column indicated that there was missing information or that the table had not been completed by the Census Bureau. Also, the explanation of why some cells are blank in the legend for this prototype makes its text lengthier than the legend for the other two prototypes. As mentioned above, a few of the participants commented that "there was no way" that they were going to read all of the content in the legend for this prototype. There is also a potential section 508 accessibility issue with labeling some reliability cells and not others, since there is no accessible text information about the reliability of the acceptable estimates (i.e., there is something available to read for "use caution," but not for all of the other cells.)

Recommendation: Overall, the lack of a contrasting message to "use caution" leads to a broad range of interpretation for the blank cells. We recommend considering a label such as "acceptable" instead of leaving the cells blank. One participant suggested using red instead of yellow for the "use caution" cells because it would be easier to see. This suggestion, however, conflicts with the standard use of yellow for caution and red for danger in Western cultures like the United States. Because the primary users of these ACS data tables are Americans, yellow is the recommended color to use when caution is indicated.

5. THREE-LEVEL INDICATOR

Observations: Several participants spontaneously mentioned the "stoplight" analogy that motivated this study. One participant mentioned that she felt more confident in her responses because she felt the stoplight analogy gave her more definitive information. One color-blind participant used the stoplight analogy to aid in interpreting the meaning of the coding because she could not tell for certain what the colors were, saying, "I think green, yellow, and red make sense because they are from stoplight colors. Green means you can go forward you can go forward with using them. Isn't green usually on the bottom of the stop light?" During debriefings, a few participants mentioned that the "good" message in the green cells of the three-level tables was not as helpful as the "excellent" message from the four-level tables (i.e., "good" was too weak a word for the best possible reliability). One participant wrote on her satisfaction form, "This one added excellent so it gave me more confidence than the good of the other one."

Usability Issues: Although the participants who saw the three- or four-level indicator tables first mentioned the color of the reliability column, a few did not notice or look at the indicator legend itself. The fact that the vertical order of the colors is the opposite pattern of an actual stoplight may cause some confusion (traffic stoplight has red on top, but prototype has green on top). This reverse order of the stoplight colors is somewhat counterintuitive from a usability perspective.

Recommendations: Consider placing the legend in a position above the table so that it is closer to the reliability column. In terms of putting green on top of the stoplight instead of red, further testing and research on possible solutions to this issue is recommended. Finally, we recommend using "excellent" for the high-reliability estimates for future

rounds of usability testing. This might also require the widening of the “fair” range because CVs on the high end of “good” might not be considered “excellent.”

6. FOUR-LEVEL INDICATOR

Observations: One color-blind participant mentioned that she was glad the text was used in the reliability column and not just the color, saying, “"I think its good that the word is there and not just the color… it would be more difficult for me to use without the word embedded in color”. As mentioned above in number 5, several users said that they preferred this prototype over the three-level prototype because the “excellent” and “good” labels provided more information about the more reliable estimates. Still other participants commented that they liked the fact that the four-level indicator just gave them more information overall.

Usability Issues: A few participants mentioned that the use of this color scheme made the smaller geography tables appear overly “busy”. One participant said, “"Whoa, that’s a lot of color," and another said, “It seems like it’s starting to get busy with the colors." However, participants tended to find the answers to tasks that asked for high-quality estimates more quickly with the four-level indicator than they did with either of the other prototypes or with the baseline tables.

Recommendations: The overall “busy-ness” may deter some users from using the reliability column. However, this issue should be further examined in future rounds of testing.

7. INDICATOR BOX

Observation: For all of the prototypes, several participants mentioned that they did not like having to scroll all the way back to the top of the table in order to read the indicator box again. This comment occurred most often for the four-level indicator, which contained more categories of information as compared to the other indicators.

Usability Issues: The frustration that users expressed at having to scroll to the top of the page to look at the legend when reporting estimates from the bottom of the table (e.g., Ancestry section) could potentially lead the user to stop looking at the legend. There may be other issues with the indicator box being displayed correctly on the screen due to the excessive width of the table (see High Priority Issue number 1, above).

Recommendations: Some participants recommended repeating the legend at the bottom of the page, and several suggested having the legend move down the side of the page as the user scrolls down the table. However, we acknowledge that the technology to support a moving table legend most likely is out of scope for the Adobe PDF and Excel versions of these tables. We recommend repeating the indicator box at the bottom of the page so that the user will only have to scroll vertically for half of the length of the table to refer to the indicator legend.

8. USER UNFAMILIARITY WITH CV

Observations: Although several baseline and prototype participants explicitly mentioned that they did not know what the coefficient of variation was, they went on to explain their answers with reference to the size of the estimate in proportion to the margin of error. A few used the word "proportion" of the error to the estimate (i.e.., the relative size of the error to the estimate itself) and said they would not use estimates that were higher than a certain proportion (e.g., one baseline participant repeatedly used "fifty percent" as her standard for deciding that an estimate's MOE was too high to use). Several participants mentioned that they wished that the actual CV value was listed instead of a derivative interpretation of it, especially for the two-level indicator where most of the reliability cells were blank.

Usability Issues: Although the CV is commonly used by the Census Bureau, it is a less familiar measure of error to the general population. Many novice users of the ACS data tables may not understand what the indicator is really measuring. The fact that several participants said that they did not know that the CV was, but still reported the proportion of MOE to the estimate, indicates that they understand the concept behind the use of CV as the measure for the indicator. However, since the participants did not recognize that the color-coding was based on the CV, which represents the proportion of error to the estimate, there is a problem with the way the legend is being interpreted.

Recommendations: Although the only documented Census Bureau standard for data reliability is a coefficient of variation of .3 or less, which can be found in *Quality Requirements for Releasing Data Products*, page 7, section 1.B. (Cahoon et al., 2007), this may not be an intuitive standard to novice users. We recommend revising the text in the legend to explain that the CV does measure the proportion of the standard error to the estimate. Below, we offer more specific recommendations for making this change.

Here is the current text from the four-level indicator legend box:

"The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate."

Participants tended not to relate their own concept of the relationship between the estimate and its associated error to the "percentage" mentioned in this text. One possible explanation for this is that the indicator legend uses the format ".30" and not "30%." We recommend testing the use of the term "proportion" instead of "percentage" in the explanatory test in future rounds of testing.

9. ABSENCE OF GUIDANCE ON WHETHER OR NOT TO USE ESTIMATES LABELED "FAIR"

Observation: Several participants explicitly stated that they were not sure whether they should use the estimates that were coded with "fair." A few said that they wished there

were some guidance on whether they should use those estimates (i.e., whether they should "report" the estimates).

Recommendation: The labels used in the data reliability indicator may need further explanation in the legend. However, the length of the text in the two-level indicator prevented some participants from reading it. One concern with providing information about whether to use the estimates is that this decision is ultimately at the discretion of the user, and there are no set guidelines for how and when to use estimates. Essentially, we cannot make the users' decisions for them, and they will ultimately have to perform some evaluation of the data quality on their own. Further research into this issue is recommended, but we also recommended the future usability testing of a cautionary note about users needing to make a decision whether to use an estimate based on the situation (see also Weidman (2008)).

3.2 Medium Priority Findings and Recommendations:

1. NOTE ABOUT POPULATION ESTIMATES ABOVE TABLE

Observation: Several participants mentioned that the note above these tables about using ACS population estimates versus the official Population estimates was one of the first things that they noticed on the first table that they saw. This is the content of that note: "NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties."

Many participants also said that they thought this note was going to tell them about the data in the table, which it did not. More specifically, a few said that they thought this note would tell them whether the data were given in thousands, or represented a total/count.

Usability Issues: Participants expected that this space would be used for explanation of how to use the tables instead of a note that does not directly relate to using the data within the table itself. The purpose and meaning of the note are likely to be lost on some users.

Recommendation: The usability staff realizes that there is not much flexibility in the location of this footnote, but it might be a good idea to put this note in a less prominent location and instead add some short guidelines for using the table in the current location. The note itself needs to clarify why users need to know this information. It is unclear what decisions, if any, users are expected to make based on this note. In a future version, it may be a good idea to add a note indicating that the reliability indicator is just a guideline for helping users to determine whether to use the estimate, but leaving that decision to their own discretion.

2. CELLS CONTAINING “N”

Observation: Several tables contained an “N” instead of a valid estimate. Only the two-level indicator legend explained that this “N” means that the sample was too small to produce an estimate.

Usability Issues: Several participants remarked that they did not know what the “N” meant, and a few thought that these characters were produced due to a calculation error by the Census Bureau. Although an explanation for this “N” value was included in the legend box for the two-level indicator, many participants either did not initially read the box or never read it. Some participants mentioned that they were looking for this particular explanation of the “N” values for the two-level indicator, but they did not find it in the legend box.

Recommendation: Briefly define what the “N” values stand for in plain language displayed prominently on the page. This information would be appropriate in the space where the Note, mentioned in Medium Priority Usability Issue number 1, is currently located.

4.0 Summary

Overall, the prototype participants completed the tasks more accurately and more quickly than did the baseline participants, and the prototype participants more frequently reported a measure of error along with the estimate in their response. Every participant indicated a preference for one of the prototype tables over the current version of the tables. Some of the most important issues to resolve are the misuse or misunderstanding of blank reliability cells in the two-level indicator, participants either missing or not reading the meaning of “N” cells, and participants not making the connection that the indicator is based on CV and, therefore, represents the proportional relationship between the error and the estimate. Further usability testing is recommended to explore the issues identified above.

5.0 References

Cahoon, L., Donnalley, G., Gore, E., Kostanich, D., Runyan, R., Detlefsen, R., and Stern, S. (2007). *Quality Requirements for Releasing Data Products*. U.S. Census Bureau official documentation, 03/22/07.

Lohse, G.L. (1997) *Consumer Eye Movement Patterns on Yellow Pages Advertising*', Journal of *Advertising* 26(1), 61–73.

Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., & Newman, J. K. (2004). *The Determinants of Web Page Viewing Behavior: An Eye-Tracking Study*. New York: Association for Computing Machinery.

Redish, J. (2007). *Letting go of the words: Writing Web content that works*. San Francisco: Elsevier/Morgan Kaufmann.

Weidman, L. (2008). *Issues concerning the ACS Research and Evaluation project proposal "ACS Data Confidence Indicator (Stoplight)"*, White Paper, 07/01/08.

Whitford, D. and Weinberg, D. (2008). *Proposal to Highlight American Community Survey Data with a Data Confidence Indicator*, 05/01/08.

Appendix A: Prototypes of Data Reliability Indicator Displays

Four-Level Indicator Tables

Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: California

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Reliability Legend based on the Coefficient of Variation (CV)

Range	Reliability
CV <=0.10	excellent
0.10 < CV <=0.30	good
0.30 < CV <=0.61	fair
CV > 0.61	poor

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**12,151,227**	excellent	**+/-18,090**
Family households (families)	8,303,793	excellent	+/-26,257
With own children under 18 years	4,239,440	excellent	+/-24,678
Married-couple families	6,051,701	excellent	+/-25,589
With own children under 18 years	3,010,321	excellent	+/-21,578
Male householder, no wife present	698,432	excellent	+/-14,287
With own children under 18 years	338,125	excellent	+/-10,354
Female householder, no husband present	1,553,660	excellent	+/-16,324
With own children under 18 years	890,994	excellent	+/-13,084
Nonfamily households	3,847,434	excellent	+/-25,448
Householder living alone	2,994,372	excellent	+/-23,715
65 years and over	978,553	excellent	+/-11,756
Households with one or more people under 18 years	4,696,427	excellent	+/-27,592
Households with one or more people 65 years and over	2,710,892	excellent	+/-11,227
Average household size	2.93	excellent	+/-0.01
Average family size	3.54	excellent	+/-0.01
RELATIONSHIP			
Household population	**35,594,342**	excellent	*****
Householder	12,151,227	excellent	+/-18,090
Spouse	6,046,430	excellent	+/-26,006
Child	11,567,876	excellent	+/-38,674
Other relatives	3,452,551	excellent	+/-40,572
Nonrelatives	2,376,258	excellent	+/-33,162
Unmarried partner	743,837	excellent	+/-12,746
MARITAL STATUS			
Males 15 years and over	**14,185,501**	excellent	**+/-4,738**
Never married	5,374,190	excellent	+/-25,966
Now married, except separated	7,097,824	excellent	+/-29,939
Separated	280,535	excellent	+/-9,066
Widowed	308,020	excellent	+/-8,095
Divorced	1,124,932	excellent	+/-14,803
Females 15 years and over	**14,384,278**	excellent	**+/-4,184**
Never married	4,358,728	excellent	+/-22,946
Now married, except separated	6,745,609	excellent	+/-30,747
Separated	428,585	excellent	+/-13,050
Widowed	1,242,329	excellent	+/-16,083
Divorced	1,609,027	excellent	+/-17,708
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**515,991**	excellent	**+/-13,067**
Unmarried women (widowed, divorced, and never married)	159,666	excellent	+/-7,389
Per 1,000 unmarried women	33	excellent	+/-1
Per 1,000 women 15 to 50 years old	55	excellent	+/-1
Per 1,000 women 15 to 19 years old	24	excellent	+/-3
Per 1,000 women 20 to 34 years old	96	excellent	+/-3
Per 1,000 women 35 to 50 years old	27	excellent	+/-1
GRANDPARENTS			
Number of grandparents living with own grandchildren under 18 years	**956,156**	excellent	**+/-16,048**
Responsible for grandchildren	273,070	excellent	+/-9,822
Years responsible for grandchildren			
Less than 1 year	57,416	excellent	+/-5,078
1 or 2 years	66,327	excellent	+/-5,218
3 or 4 years	46,167	excellent	+/-3,729
5 or more years	103,160	excellent	+/-6,855
Characteristics of grandparents responsible for own grandchildren under 18 years			
Who are female	62.7%	excellent	+/-0.9
Who are married	71.5%	excellent	+/-1.3
SCHOOL ENROLLMENT			
Population 3 years and over enrolled in school	**10,385,391**	excellent	**+/-31,813**
Nursery school, preschool	605,484	excellent	+/-12,202
Kindergarten	501,274	excellent	+/-10,650
Elementary school (grades 1-8)	4,179,120	excellent	+/-14,597
High school (grades 9-12)	2,313,813	excellent	+/-15,819
College or graduate school	2,785,700	excellent	+/-25,952
EDUCATIONAL ATTAINMENT			
Population 25 years and over	**23,133,174**	excellent	**+/-5,870**
Less than 9th grade	2,462,134	excellent	+/-27,635

9th to 12th grade, no diploma	2,143,220	excellent	+/-23,165
High school graduate (includes equivalency)	5,368,962	excellent	+/-38,936
Some college, no degree	4,684,571	excellent	+/-32,448
Associate's degree	1,767,989	excellent	+/-21,487
Bachelor's degree	4,311,428	excellent	+/-32,369
Graduate or professional degree	2,394,870	excellent	+/-23,306
Percent high school graduate or higher	80.1%	excellent	+/-0.1
Percent bachelor's degree or higher	29.0%	excellent	+/-0.1
VETERAN STATUS			
Civilian population 18 years and over	**26,789,221**	excellent	**+/-5,076**
Civilian veterans	2,142,367	excellent	+/-22,070
DISABILITY STATUS OF THE CIVILIAN NONINSTITUTIONALIZED POPULATION			
Population 5 years and over	**33,211,121**	excellent	**+/-5,156**
With a disability	4,283,468	excellent	+/-32,065
Population 5 to 15 years	**5,765,810**	excellent	**+/-10,458**
With a disability	274,922	excellent	+/-7,791
Population 16 to 64 years	**23,637,212**	excellent	**+/-11,405**
With a disability	2,443,463	excellent	+/-28,342
Population 65 years and over	**3,808,099**	excellent	**+/-3,894**
With a disability	1,565,083	excellent	+/-15,094
RESIDENCE 1 YEAR AGO			
Population 1 year and over	**35,932,693**	excellent	**+/-12,161**
Same house	29,915,828	excellent	+/-65,533
Different house in the U.S.	5,696,055	excellent	+/-64,003
Same county	3,959,241	excellent	+/-55,479
Different county	1,736,814	excellent	+/-31,401
Same state	1,214,709	excellent	+/-29,479
Different state	522,105	excellent	+/-17,045
Abroad	320,810	excellent	+/-15,675
PLACE OF BIRTH			
Total population	**36,457,549**	excellent	*****
Native	26,555,482	excellent	+/-50,752
Born in United States	26,159,970	excellent	+/-51,833
State of residence	19,089,635	excellent	+/-57,197
Different state	7,070,335	excellent	+/-39,658
Born in Puerto Rico, U.S. Island areas, or born abroad to American parent(s)	395,512	excellent	+/-10,569
Foreign born	9,902,067	excellent	+/-50,752
U.S. CITIZENSHIP STATUS			
Foreign-born population	**9,902,067**	excellent	**+/-50,752**
Naturalized U.S. citizen	4,264,806	excellent	+/-35,816
Not a U.S. citizen	5,637,261	excellent	+/-46,698
YEAR OF ENTRY			
Population born outside the United States	**10,297,579**	excellent	**+/-51,833**
Native	395,512	excellent	+/-10,569
Entered 2000 or later	51,937	excellent	+/-4,971
Entered before 2000	343,575	excellent	+/-10,211
Foreign born	9,902,067	excellent	+/-50,752
Entered 2000 or later	1,954,316	excellent	+/-33,181
Entered before 2000	7,947,751	excellent	+/-40,904
WORLD REGION OF BIRTH OF FOREIGN BORN			
Foreign-born population, excluding population born at sea	**9,901,970**	excellent	**+/-50,736**
Europe	682,224	excellent	+/-15,636
Asia	3,356,805	excellent	+/-26,736
Africa	143,882	excellent	+/-9,892
Oceania	66,273	excellent	+/-5,187
Latin America	5,517,741	excellent	+/-42,576
Northern America	135,045	excellent	+/-6,824
LANGUAGE SPOKEN AT HOME			
Population 5 years and over	**33,784,883**	excellent	**+/-2,490**
English only	19,414,977	excellent	+/-61,768
Language other than English	14,369,906	excellent	+/-61,691
Speak English less than "very well"	6,778,944	excellent	+/-50,426
Spanish	9,588,622	excellent	+/-46,828
Speak English less than "very well"	4,700,922	excellent	+/-44,612
Other Indo-European languages	1,459,471	excellent	+/-31,551
Speak English less than "very well"	474,782	excellent	+/-18,180
Asian and Pacific Islander languages	3,040,704	excellent	+/-26,199
Speak English less than "very well"	1,515,494	excellent	+/-22,205
Other languages	281,109	excellent	+/-15,424
Speak English less than "very well"	87,746	excellent	+/-6,620
ANCESTRY			
Total population	**36,457,549**	excellent	*****
American	1,081,295	excellent	+/-28,738

Arab	238,846	excellent	+/-14,455
Czech	93,229	excellent	+/-5,272
Danish	207,023	excellent	+/-9,367
Dutch	439,105	excellent	+/-13,070
English	2,592,100	excellent	+/-25,666
French (except Basque)	807,575	excellent	+/-17,106
French Canadian	125,694	excellent	+/-7,132
German	3,640,022	excellent	+/-37,155
Greek	148,356	excellent	+/-8,630
Hungarian	129,159	excellent	+/-8,283
Irish	2,830,975	excellent	+/-31,114
Italian	1,560,870	excellent	+/-24,224
Lithuanian	50,506	excellent	+/-5,290
Norwegian	410,755	excellent	+/-12,644
Polish	518,536	excellent	+/-14,491
Portuguese	391,649	excellent	+/-14,544
Russian	444,146	excellent	+/-14,060
Scotch-Irish	471,320	excellent	+/-12,813
Scottish	590,878	excellent	+/-14,759
Slovak	27,015	excellent	+/-3,629
Subsaharan African	232,617	excellent	+/-13,078
Swedish	479,942	excellent	+/-14,260
Swiss	122,593	excellent	+/-6,569
Ukrainian	93,697	excellent	+/-7,881
Welsh	181,864	excellent	+/-8,294
West Indian (excluding Hispanic origin groups)	73,004	excellent	+/-7,375

Source: U.S. Census Bureau, 2006 American Community Survey

Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Wilmington city, Delaware

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Reliability Legend based on the Coefficient of Variation (CV)

Range	Reliability
CV <=0.10	excellent
0.10 < CV <=0.30	good
0.30 < CV <=0.61	fair
CV > 0.61	poor

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**26,443**	excellent	**+/-2,007**
Family households (families)	13,066	excellent	+/-1,554
With own children under 18 years	5,325	good	+/-980
Married-couple families	6,693	excellent	+/-970
With own children under 18 years	2,171	good	+/-586
Male householder, no wife present	922	fair	+/-503
With own children under 18 years	474	fair	+/-398
Female householder, no husband present	5,451	good	+/-1,133
With own children under 18 years	2,680	good	+/-770
Nonfamily households	13,377	excellent	+/-1,607
Householder living alone	11,894	excellent	+/-1,556
65 years and over	4,084	good	+/-682
Households with one or more people under 18 years	6,595	excellent	+/-1,023
Households with one or more people 65 years and over	6,373	excellent	+/-802
Average household size	2.32	excellent	+/-0.15
Average family size	3.45	excellent	+/-0.26
RELATIONSHIP			
Household population	**61,441**	excellent	**+/-5,547**
Householder	26,443	excellent	+/-2,007
Spouse	6,764	excellent	+/-962
Child	16,962	good	+/-2,859
Other relatives	8,326	good	+/-2,688
Nonrelatives	2,946	good	+/-964
Unmarried partner	785	good	+/-336
MARITAL STATUS			
Males 15 years and over	**24,174**	excellent	**+/-2,064**
Never married	11,354	excellent	+/-1,655
Now married, except separated	7,458	excellent	+/-1,153
Separated	692	fair	+/-464
Widowed	1,393	good	+/-573
Divorced	3,277	good	+/-1,056
Females 15 years and over	**28,120**	excellent	**+/-2,866**
Never married	10,772	good	+/-1,797
Now married, except separated	7,665	excellent	+/-1,131
Separated	1,155	fair	+/-576
Widowed	3,777	good	+/-809
Divorced	4,751	good	+/-1,270
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**975**	good	**+/-478**
Unmarried women (widowed, divorced, and never married)	471	fair	+/-321
Per 1,000 unmarried women	43	fair	+/-29
Per 1,000 women 15 to 50 years old	60	good	+/-29
Per 1,000 women 15 to 19 years old	101	fair	+/-78
Per 1,000 women 20 to 34 years old	88	fair	+/-60
Per 1,000 women 35 to 50 years old	20	fair	+/-20
GRANDPARENTS			
Number of grandparents living with own grandchildren under 18 years	**N**		**N**
Responsible for grandchildren	N		N
Years responsible for grandchildren			
Less than 1 year	N		N
1 or 2 years	N		N
3 or 4 years	N		N
5 or more years	N		N
Characteristics of grandparents responsible for own grandchildren under 18 years			
Who are female	76.7%	good	+/-19.1
Who are married	45.6%	fair	+/-36.6
SCHOOL ENROLLMENT			
Population 3 years and over enrolled in school	**14,597**	good	**+/-2,829**
Nursery school, preschool	854	fair	+/-510
Kindergarten	1,022	fair	+/-631
Elementary school (grades 1-8)	5,838	good	+/-1,505
High school (grades 9-12)	3,471	good	+/-1,279
College or graduate school	3,412	good	+/-1,087
EDUCATIONAL ATTAINMENT			
Population 25 years and over	**44,522**	excellent	**+/-3,155**
Less than 9th grade	3,949	good	+/-1,054

9th to 12th grade, no diploma	5,724	good	+/-1,374
High school graduate (includes equivalency)	13,514	good	+/-2,320
Some college, no degree	7,940	good	+/-1,493
Associate's degree	3,320	good	+/-1,153
Bachelor's degree	5,496	good	+/-1,058
Graduate or professional degree	4,579	good	+/-1,244
Percent high school graduate or higher	78.3%	excellent	+/-3.7
Percent bachelor's degree or higher	22.6%	excellent	+/-3.7
VETERAN STATUS			
Civilian population 18 years and over	**49,276**	excellent	**+/-3,711**
Civilian veterans	N		N
DISABILITY STATUS OF THE CIVILIAN NONINSTITUTIONALIZED POPULATION			
Population 5 years and over	**58,765**	excellent	**+/-4,922**
With a disability	12,909	excellent	+/-2,009
Population 5 to 15 years	**7,992**	good	**+/-1,702**
With a disability	936	fair	+/-836
Population 16 to 64 years	**42,773**	excellent	**+/-3,954**
With a disability	8,027	good	+/-1,477
Population 65 years and over	**8,000**	excellent	**+/-1,067**
With a disability	3,946	good	+/-792
RESIDENCE 1 YEAR AGO			
Population 1 year and over	**63,187**	excellent	**+/-5,502**
Same house	52,978	excellent	+/-4,657
Different house in the U.S.	9,858	good	+/-2,979
Same county	8,307	good	+/-2,727
Different county	1,551	good	+/-626
Same state	441	fair	+/-352
Different state	1,110	good	+/-526
Abroad	351	fair	+/-337
PLACE OF BIRTH			
Total population	**63,752**	excellent	**+/-5,536**
Native	58,885	excellent	+/-5,394
Born in United States	57,963	excellent	+/-5,387
State of residence	36,369	excellent	+/-4,133
Different state	21,594	excellent	+/-2,861
Born in Puerto Rico, U.S. Island areas, or born abroad to American parent(s)	922	fair	+/-560
Foreign born	4,867	good	+/-1,679
U.S. CITIZENSHIP STATUS			
Foreign-born population	**4,867**	good	**+/-1,679**
Naturalized U.S. citizen	1,344	good	+/-629
Not a U.S. citizen	3,523	good	+/-1,666
YEAR OF ENTRY			
Population born outside the United States	**5,789**	good	**+/-1,731**
Native	922	fair	+/-560
Entered 2000 or later	0	poor	+/-261
Entered before 2000	922	fair	+/-560
Foreign born	4,867	good	+/-1,679
Entered 2000 or later	1,938	fair	+/-1,032
Entered before 2000	2,929	good	+/-990
WORLD REGION OF BIRTH OF FOREIGN BORN			
Foreign-born population, excluding population born at sea	**N**		**N**
Europe	N		N
Asia	N		N
Africa	N		N
Oceania	N		N
Latin America	N		N
Northern America	N		N
LANGUAGE SPOKEN AT HOME			
Population 5 years and over	**N**		**N**
English only	N		N
Language other than English	N		N
Speak English less than "very well"	N		N
Spanish	N		N
Speak English less than "very well"	N		N
Other Indo-European languages	N		N
Speak English less than "very well"	N		N
Asian and Pacific Islander languages	N		N
Speak English less than "very well"	N		N
Other languages	N		N
Speak English less than "very well"	N		N
ANCESTRY			
Total population	**63,752**	excellent	**+/-5,536**
American	924	fair	+/-599

Arab	149	poor	+/-250
Czech	0	poor	+/-261
Danish	657	poor	+/-993
Dutch	150	fair	+/-142
English	3,744	good	+/-991
French (except Basque)	433	fair	+/-281
French Canadian	110	poor	+/-133
German	4,206	good	+/-1,083
Greek	54	poor	+/-88
Hungarian	154	poor	+/-200
Irish	6,641	good	+/-1,884
Italian	4,398	good	+/-1,160
Lithuanian	318	poor	+/-361
Norwegian	136	poor	+/-231
Polish	2,568	good	+/-954
Portuguese	51	poor	+/-83
Russian	344	fair	+/-287
Scotch-Irish	481	fair	+/-369
Scottish	584	fair	+/-323
Slovak	70	poor	+/-116
Subsaharan African	163	fair	+/-154
Swedish	226	poor	+/-240
Swiss	173	poor	+/-189
Ukrainian	329	fair	+/-298
Welsh	425	poor	+/-477
West Indian (excluding Hispanic origin groups)	0	poor	+/-261

Source: U.S. Census Bureau, 2006 American Community Survey

Three-Level Indicator Tables

Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Cook County, Illinois

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Reliability Legend based on the Coefficient of Variation (CV)

Range	Reliability
CV <=0.30	good
0.30 < CV <=0.61	fair
CV > 0.61	poor

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**1,932,197**	good	**+/-8,605**
Family households (families)	1,217,733	good	+/-11,714
With own children under 18 years	576,417	good	+/-9,573
Married-couple families	824,922	good	+/-12,260
With own children under 18 years	380,049	good	+/-9,213
Male householder, no wife present	97,227	good	+/-4,702
With own children under 18 years	36,162	good	+/-2,913
Female householder, no husband present	295,584	good	+/-7,577
With own children under 18 years	160,206	good	+/-5,923
Nonfamily households	714,464	good	+/-10,203
Householder living alone	607,726	good	+/-10,372
65 years and over	183,241	good	+/-5,319
Households with one or more people under 18 years	650,125	good	+/-9,711
Households with one or more people 65 years and over	445,793	good	+/-3,588
Average household size	2.69	good	+/-0.01
Average family size	3.47	good	+/-0.03
RELATIONSHIP			
Household population	**5,197,637**	good	**+/-684**
Householder	1,932,197	good	+/-8,605
Spouse	823,694	good	+/-12,275
Child	1,685,734	good	+/-17,554
Other relatives	499,106	good	+/-17,284
Nonrelatives	256,906	good	+/-10,099
Unmarried partner	92,715	good	+/-5,128
MARITAL STATUS			
Males 15 years and over	**1,998,715**	good	**+/-386**
Never married	816,703	good	+/-10,439
Now married, except separated	941,158	good	+/-13,028
Separated	43,162	good	+/-3,652
Widowed	53,199	good	+/-3,758
Divorced	144,493	good	+/-5,678
Females 15 years and over	**2,157,680**	good	**+/-747**
Never married	747,274	good	+/-9,260
Now married, except separated	895,594	good	+/-12,206
Separated	60,098	good	+/-4,302
Widowed	229,664	good	+/-6,024
Divorced	225,050	good	+/-7,804
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**73,259**	good	**+/-4,879**
Unmarried women (widowed, divorced, and never married)	28,928	good	+/-2,897
Per 1,000 unmarried women	37	good	+/-4
Per 1,000 women 15 to 50 years old	54	good	+/-4
Per 1,000 women 15 to 19 years old	27	good	+/-6
Per 1,000 women 20 to 34 years old	94	good	+/-7
Per 1,000 women 35 to 50 years old	28	good	+/-3
GRANDPARENTS			
Number of grandparents living with own grandchildren under 18 years	**130,069**	good	**+/-7,017**
Responsible for grandchildren	50,795	good	+/-4,002
Years responsible for grandchildren			
Less than 1 year	7,767	good	+/-1,467
1 or 2 years	13,077	good	+/-2,214
3 or 4 years	8,088	good	+/-1,910
5 or more years	21,863	good	+/-2,936
Characteristics of grandparents responsible for own grandchildren under 18 years			
Who are female	69.2%	good	+/-2.7
Who are married	58.8%	good	+/-4.5
SCHOOL ENROLLMENT			
Population 3 years and over enrolled in school	**1,446,837**	good	**+/-10,217**
Nursery school, preschool	102,031	good	+/-5,062
Kindergarten	73,301	good	+/-4,494
Elementary school (grades 1-8)	593,479	good	+/-6,521
High school (grades 9-12)	310,359	good	+/-5,369
College or graduate school	367,667	good	+/-8,049
EDUCATIONAL ATTAINMENT			
Population 25 years and over	**3,431,992**	good	**+/-863**
Less than 9th grade	310,367	good	+/-9,618

9th to 12th grade, no diploma	325,249	good	+/-9,764
High school graduate (includes equivalency)	887,087	good	+/-15,012
Some college, no degree	628,589	good	+/-13,472
Associate's degree	216,614	good	+/-7,877
Bachelor's degree	644,383	good	+/-10,701
Graduate or professional degree	419,703	good	+/-10,766
Percent high school graduate or higher	81.5%	good	+/-0.4
Percent bachelor's degree or higher	31.0%	good	+/-0.4
VETERAN STATUS			
Civilian population 18 years and over	**3,929,871**	good	**+/-960**
Civilian veterans	274,364	good	+/-6,597
DISABILITY STATUS OF THE CIVILIAN NONINSTITUTIONALIZED POPULATION			
Population 5 years and over	**4,851,067**	good	**+/-1,356**
With a disability	631,766	good	+/-12,241
Population 5 to 15 years	**817,775**	good	**+/-3,824**
With a disability	41,702	good	+/-3,605
Population 16 to 64 years	**3,439,770**	good	**+/-4,102**
With a disability	350,022	good	+/-10,866
Population 65 years and over	**593,522**	good	**+/-991**
With a disability	240,042	good	+/-5,828
RESIDENCE 1 YEAR AGO			
Population 1 year and over	**5,207,495**	good	**+/-4,411**
Same house	4,408,325	good	+/-24,982
Different house in the U.S.	761,257	good	+/-25,556
Same county	623,124	good	+/-24,407
Different county	138,133	good	+/-9,110
Same state	50,162	good	+/-5,433
Different state	87,971	good	+/-7,316
Abroad	37,913	good	+/-4,743
PLACE OF BIRTH			
Total population	**5,288,655**	good	*****
Native	4,169,346	good	+/-22,439
Born in United States	4,101,596	good	+/-22,931
State of residence	3,266,991	good	+/-24,871
Different state	834,605	good	+/-13,981
Born in Puerto Rico, U.S. Island areas, or born abroad to American parent(s)	67,750	good	+/-5,710
Foreign born	1,119,309	good	+/-22,439
U.S. CITIZENSHIP STATUS			
Foreign-born population	**1,119,309**	good	**+/-22,439**
Naturalized U.S. citizen	496,925	good	+/-15,704
Not a U.S. citizen	622,384	good	+/-17,089
YEAR OF ENTRY			
Population born outside the United States	**1,187,059**	good	**+/-22,931**
Native	67,750	good	+/-5,710
Entered 2000 or later	8,536	good	+/-2,458
Entered before 2000	59,214	good	+/-4,945
Foreign born	1,119,309	good	+/-22,439
Entered 2000 or later	242,954	good	+/-13,228
Entered before 2000	876,355	good	+/-20,483
WORLD REGION OF BIRTH OF FOREIGN BORN			
Foreign-born population, excluding population born at sea	**1,119,309**	good	**+/-22,439**
Europe	283,139	good	+/-11,668
Asia	249,836	good	+/-6,563
Africa	28,920	good	+/-5,289
Oceania	1,479	fair	+/-833
Latin America	548,507	good	+/-13,411
Northern America	7,428	good	+/-1,361
LANGUAGE SPOKEN AT HOME			
Population 5 years and over	**4,899,072**	good	**+/-491**
English only	3,250,628	good	+/-20,303
Language other than English	1,648,444	good	+/-20,346
Speak English less than "very well"	777,445	good	+/-16,850
Spanish	977,765	good	+/-8,542
Speak English less than "very well"	483,054	good	+/-13,307
Other Indo-European languages	432,303	good	+/-17,221
Speak English less than "very well"	192,822	good	+/-10,905
Asian and Pacific Islander languages	167,258	good	+/-8,317
Speak English less than "very well"	80,305	good	+/-5,824
Other languages	71,118	good	+/-7,956
Speak English less than "very well"	21,264	good	+/-3,434
ANCESTRY			
Total population	**5,288,655**	good	*****
American	80,299	good	+/-6,679

Arab	32,159	good	+/-7,153
Czech	51,560	good	+/-4,536
Danish	15,560	good	+/-2,188
Dutch	50,008	good	+/-4,354
English	183,009	good	+/-8,982
French (except Basque)	66,674	good	+/-5,332
French Canadian	11,371	good	+/-1,664
German	603,164	good	+/-15,340
Greek	63,126	good	+/-6,633
Hungarian	25,497	good	+/-3,757
Irish	539,087	good	+/-12,719
Italian	336,430	good	+/-11,828
Lithuanian	48,564	good	+/-4,353
Norwegian	41,840	good	+/-3,788
Polish	507,498	good	+/-13,952
Portuguese	3,026	good	+/-1,012
Russian	81,510	good	+/-6,102
Scotch-Irish	34,284	good	+/-3,786
Scottish	40,053	good	+/-3,531
Slovak	14,545	good	+/-1,970
Subsaharan African	50,587	good	+/-6,574
Swedish	83,743	good	+/-5,633
Swiss	9,306	good	+/-1,805
Ukrainian	27,650	good	+/-3,790
Welsh	13,013	good	+/-2,478
West Indian (excluding Hispanic origin groups)	19,610	good	+/-3,919

Source: U.S. Census Bureau, 2006 American Community Survey

Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Yakima city, Washington

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

Reliability Legend based on the Coefficient of Variation (CV)

Range	Reliability
CV <=0.30	good
0.30 < CV <=0.61	fair
CV > 0.61	poor

The coefficient of variation (CV) is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**30,491**	good	**+/-1,731**
Family households (families)	18,533	good	+/-1,639
With own children under 18 years	9,430	good	+/-1,448
Married-couple families	13,316	good	+/-1,572
With own children under 18 years	6,446	good	+/-1,216
Male householder, no wife present	1,152	good	+/-492
With own children under 18 years	548	fair	+/-388
Female householder, no husband present	4,065	good	+/-747
With own children under 18 years	2,436	good	+/-702
Nonfamily households	11,958	good	+/-1,434
Householder living alone	9,777	good	+/-1,216
65 years and over	4,515	good	+/-804
Households with one or more people under 18 years	10,923	good	+/-1,423
Households with one or more people 65 years and over	8,099	good	+/-908
Average household size	2.66	good	+/-0.14
Average family size	3.46	good	+/-0.22
RELATIONSHIP			
Household population	**81,113**	good	**+/-5,974**
Householder	30,491	good	+/-1,731
Spouse	13,607	good	+/-1,650
Child	25,697	good	+/-3,061
Other relatives	6,208	good	+/-2,044
Nonrelatives	5,110	good	+/-1,690
Unmarried partner	2,344	good	+/-883
MARITAL STATUS			
Males 15 years and over	**31,288**	good	**+/-2,521**
Never married	11,417	good	+/-1,678
Now married, except separated	14,389	good	+/-1,589
Separated	779	fair	+/-466
Widowed	993	good	+/-415
Divorced	3,710	good	+/-889
Females 15 years and over	**32,957**	good	**+/-2,576**
Never married	7,661	good	+/-1,332
Now married, except separated	14,492	good	+/-1,917
Separated	1,462	good	+/-680
Widowed	4,653	good	+/-804
Divorced	4,689	good	+/-1,104
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**1,840**	good	**+/-661**
Unmarried women (widowed, divorced, and never married)	564	fair	+/-389
Per 1,000 unmarried women	54	fair	+/-37
Per 1,000 women 15 to 50 years old	90	good	+/-31
Per 1,000 women 15 to 19 years old	16	poor	+/-29
Per 1,000 women 20 to 34 years old	178	good	+/-59
Per 1,000 women 35 to 50 years old	25	fair	+/-25
GRANDPARENTS			
Number of grandparents living with own grandchildren under 18 years	**N**		**N**
Responsible for grandchildren	N		N
Years responsible for grandchildren			
Less than 1 year	N		N
1 or 2 years	N		N
3 or 4 years	N		N
5 or more years	N		N
Characteristics of grandparents responsible for own grandchildren under 18 years			
Who are female	N		N
Who are married	N		N
SCHOOL ENROLLMENT			
Population 3 years and over enrolled in school	**18,795**	good	**+/-2,291**
Nursery school, preschool	1,935	fair	+/-988
Kindergarten	1,162	fair	+/-647
Elementary school (grades 1-8)	8,266	good	+/-1,513
High school (grades 9-12)	4,829	good	+/-1,108
College or graduate school	2,603	good	+/-839
EDUCATIONAL ATTAINMENT			
Population 25 years and over	**51,030**	good	**+/-3,460**
Less than 9th grade	5,549	good	+/-1,323

9th to 12th grade, no diploma	7,900	good	+/-1,885
High school graduate (includes equivalency)	13,459	good	+/-1,774
Some college, no degree	10,686	good	+/-1,406
Associate's degree	2,453	good	+/-834
Bachelor's degree	7,160	good	+/-1,281
Graduate or professional degree	3,823	good	+/-963
Percent high school graduate or higher	73.6%	good	+/-4.2
Percent bachelor's degree or higher	21.5%	good	+/-3.0
VETERAN STATUS			
Civilian population 18 years and over	**60,420**	good	**+/-3,873**
Civilian veterans	6,135	good	+/-1,046
DISABILITY STATUS OF THE CIVILIAN NONINSTITUTIONALIZED POPULATION			
Population 5 years and over	**72,553**	good	**+/-5,258**
With a disability	14,402	good	+/-2,367
Population 5 to 15 years	**11,655**	good	**+/-1,883**
With a disability	1,138	fair	+/-614
Population 16 to 64 years	**50,239**	good	**+/-3,650**
With a disability	7,809	good	+/-1,700
Population 65 years and over	**10,659**	good	**+/-1,297**
With a disability	5,455	good	+/-1,225
RESIDENCE 1 YEAR AGO			
Population 1 year and over	**82,054**	good	**+/-5,758**
Same house	58,322	good	+/-5,170
Different house in the U.S.	23,308	good	+/-4,711
Same county	17,242	good	+/-3,721
Different county	6,066	good	+/-2,126
Same state	3,819	good	+/-1,588
Different state	2,247	fair	+/-1,512
Abroad	424	poor	+/-433
PLACE OF BIRTH			
Total population	**83,707**	good	**+/-5,980**
Native	70,481	good	+/-5,438
Born in United States	70,168	good	+/-5,437
State of residence	47,203	good	+/-4,554
Different state	22,965	good	+/-2,924
Born in Puerto Rico, U.S. Island areas, or born abroad to American parent(s)	313	fair	+/-233
Foreign born	13,226	good	+/-2,288
U.S. CITIZENSHIP STATUS			
Foreign-born population	**13,226**	good	**+/-2,288**
Naturalized U.S. citizen	3,311	good	+/-893
Not a U.S. citizen	9,915	good	+/-2,023
YEAR OF ENTRY			
Population born outside the United States	**13,539**	good	**+/-2,263**
Native	313	fair	+/-233
Entered 2000 or later	43	poor	+/-70
Entered before 2000	270	fair	+/-218
Foreign born	13,226	good	+/-2,288
Entered 2000 or later	3,780	good	+/-1,403
Entered before 2000	9,446	good	+/-1,726
WORLD REGION OF BIRTH OF FOREIGN BORN			
Foreign-born population, excluding population born at sea	**N**		**N**
Europe	N		N
Asia	N		N
Africa	N		N
Oceania	N		N
Latin America	N		N
Northern America	N		N
LANGUAGE SPOKEN AT HOME			
Population 5 years and over	**N**		**N**
English only	N		N
Language other than English	N		N
Speak English less than "very well"	N		N
Spanish	N		N
Speak English less than "very well"	N		N
Other Indo-European languages	N		N
Speak English less than "very well"	N		N
Asian and Pacific Islander languages	N		N
Speak English less than "very well"	N		N
Other languages	N		N
Speak English less than "very well"	N		N
ANCESTRY			
Total population	**83,707**	good	**+/-5,980**
American	3,707	good	+/-1,813

Arab	0	poor	+/-267
Czech	345	fair	+/-303
Danish	546	fair	+/-285
Dutch	1,768	fair	+/-952
English	9,178	good	+/-1,815
French (except Basque)	2,618	good	+/-823
French Canadian	154	fair	+/-150
German	14,695	good	+/-2,489
Greek	115	poor	+/-137
Hungarian	0	poor	+/-267
Irish	8,105	good	+/-1,849
Italian	550	fair	+/-359
Lithuanian	0	poor	+/-267
Norwegian	3,282	good	+/-1,212
Polish	1,684	fair	+/-1,164
Portuguese	379	poor	+/-460
Russian	498	fair	+/-485
Scotch-Irish	2,137	good	+/-814
Scottish	2,227	fair	+/-1,191
Slovak	0	poor	+/-267
Subsaharan African	210	poor	+/-279
Swedish	2,248	good	+/-1,087
Swiss	37	poor	+/-61
Ukrainian	60	poor	+/-99
Welsh	444	fair	+/-369
West Indian (excluding Hispanic origin groups)	0	poor	+/-267

Source: U.S. Census Bureau, 2006 American Community Survey

Two-Level Indicator Tables

Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: New York city, New York

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

NOTE ABOUT RELIABILITY OF SOME ESTIMATES:

Reliability
use caution

Estimates marked with a yellow "use caution" reliability box (shown above) means that the coefficient of variation (CV) is above 0.30. The CV is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**3,020,284**		**+/-9,874**
Family households (families)	1,823,095		+/-15,027
With own children under 18 years	853,419		+/-12,447
Married-couple families	1,089,597		+/-13,850
With own children under 18 years	502,221		+/-11,080
Male householder, no wife present	164,869		+/-6,286
With own children under 18 years	53,881		+/-4,642
Female householder, no husband present	568,629		+/-11,235
With own children under 18 years	297,317		+/-9,022
Nonfamily households	1,197,189		+/-13,299
Householder living alone	1,005,277		+/-13,313
65 years and over	310,098		+/-7,156
Households with one or more people under 18 years	975,699		+/-13,037
Households with one or more people 65 years and over	727,931		+/-5,676
Average household size	2.66		+/-0.01
Average family size	3.48		+/-0.02
RELATIONSHIP			
Household population	**8,035,586**		**+/-1,713**
Householder	3,020,284		+/-9,874
Spouse	1,087,632		+/-13,872
Child	2,528,284		+/-18,412
Other relatives	906,358		+/-15,851
Nonrelatives	493,028		+/-13,850
Unmarried partner	142,518		+/-6,160
MARITAL STATUS			
Males 15 years and over	**3,085,630**		**+/-790**
Never married	1,340,477		+/-12,623
Now married, except separated	1,388,250		+/-14,189
Separated	90,327		+/-5,221
Widowed	74,161		+/-4,027
Divorced	192,415		+/-6,306
Females 15 years and over	**3,508,904**		**+/-688**
Never married	1,358,014		+/-14,741
Now married, except separated	1,281,085		+/-15,845
Separated	174,075		+/-6,486
Widowed	355,775		+/-8,217
Divorced	339,955		+/-9,116
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**111,755**		**+/-4,906**
Unmarried women (widowed, divorced, and never married)	38,274		+/-3,276
Per 1,000 unmarried women	29		+/-2
Per 1,000 women 15 to 50 years old	50		+/-2
Per 1,000 women 15 to 19 years old	24		+/-4
Per 1,000 women 20 to 34 years old	79		+/-5
Per 1,000 women 35 to 50 years old	31		+/-3
GRANDPARENTS			
Number of grandparents living with own grandchildren under 18 years	**219,494**		**+/-8,382**
Responsible for grandchildren	65,236		+/-5,277
Years responsible for grandchildren			
Less than 1 year	12,973		+/-2,115
1 or 2 years	13,613		+/-2,355
3 or 4 years	11,743		+/-2,129
5 or more years	26,907		+/-3,276
Characteristics of grandparents responsible for own grandchildren under 18 years			
Who are female	74.3%		+/-2.6
Who are married	57.0%		+/-3.5
SCHOOL ENROLLMENT			
Population 3 years and over enrolled in school	**2,139,238**		**+/-14,161**
Nursery school, preschool	135,752		+/-6,507
Kindergarten	110,006		+/-5,564
Elementary school (grades 1-8)	841,229		+/-7,241
High school (grades 9-12)	468,194		+/-6,561
College or graduate school	584,057		+/-11,947
EDUCATIONAL ATTAINMENT			
Population 25 years and over	**5,487,985**		**+/-1,207**
Less than 9th grade	587,388		+/-14,154

9th to 12th grade, no diploma	583,080		+/-13,485
High school graduate (includes equivalency)	1,506,116		+/-20,702
Some college, no degree	719,992		+/-13,530
Associate's degree	332,341		+/-8,968
Bachelor's degree	1,041,909		+/-16,198
Graduate or professional degree	717,159		+/-12,355
Percent high school graduate or higher	78.7%		+/-0.4
Percent bachelor's degree or higher	32.1%		+/-0.4
VETERAN STATUS			
Civilian population 18 years and over	**6,267,449**		**+/-1,181**
Civilian veterans	254,231		+/-6,130
DISABILITY STATUS OF THE CIVILIAN NONINSTITUTIONALIZED POPULATION			
Population 5 years and over	**7,560,751**		**+/-1,799**
With a disability	1,039,458		+/-16,778
Population 5 to 15 years	**1,148,352**		**+/-4,007**
With a disability	62,075		+/-5,362
Population 16 to 64 years	**5,455,470**		**+/-4,248**
With a disability	562,012		+/-12,374
Population 65 years and over	**956,929**		**+/-2,239**
With a disability	415,371		+/-7,571
RESIDENCE 1 YEAR AGO			
Population 1 year and over	**8,092,441**		**+/-5,236**
Same house	7,188,961		+/-21,271
Different house in the U.S.	808,513		+/-19,212
Same county	542,597		+/-18,593
Different county	265,916		+/-9,560
Same state	161,288		+/-8,384
Different state	104,628		+/-7,120
Abroad	94,967		+/-7,852
PLACE OF BIRTH			
Total population	**8,214,426**		*****
Native	5,176,287		+/-28,545
Born in United States	4,858,538		+/-26,032
State of residence	4,074,701		+/-28,292
Different state	783,837		+/-11,062
Born in Puerto Rico, U.S. Island areas, or born abroad to American parent(s)	317,749		+/-9,000
Foreign born	3,038,139		+/-28,545
U.S. CITIZENSHIP STATUS			
Foreign-born population	**3,038,139**		**+/-28,545**
Naturalized U.S. citizen	1,552,173		+/-21,559
Not a U.S. citizen	1,485,966		+/-26,271
YEAR OF ENTRY			
Population born outside the United States	**3,355,888**		**+/-26,032**
Native	317,749		+/-9,000
Entered 2000 or later	27,112		+/-3,425
Entered before 2000	290,637		+/-8,354
Foreign born	3,038,139		+/-28,545
Entered 2000 or later	628,944		+/-18,279
Entered before 2000	2,409,195		+/-25,773
WORLD REGION OF BIRTH OF FOREIGN BORN			
Foreign-born population, excluding population born at sea	**3,038,073**		**+/-28,533**
Europe	520,554		+/-14,121
Asia	777,169		+/-12,664
Africa	111,803		+/-8,506
Oceania	6,762		+/-1,920
Latin America	1,600,051		+/-24,841
Northern America	21,734		+/-2,742
LANGUAGE SPOKEN AT HOME			
Population 5 years and over	**7,637,820**		**+/-458**
English only	3,981,767		+/-28,903
Language other than English	3,656,053		+/-28,927
Speak English less than "very well"	1,792,056		+/-20,814
Spanish	1,883,804		+/-13,185
Speak English less than "very well"	913,085		+/-16,634
Other Indo-European languages	998,354		+/-22,163
Speak English less than "very well"	440,589		+/-13,668
Asian and Pacific Islander languages	596,474		+/-13,130
Speak English less than "very well"	374,760		+/-10,727
Other languages	177,421		+/-12,914
Speak English less than "very well"	63,622		+/-6,244
ANCESTRY			
Total population	**8,214,426**		*****
American	260,396		+/-13,311

Arab	80,680		+/-8,405
Czech	16,990		+/-3,501
Danish	7,011		+/-1,453
Dutch	23,454		+/-2,983
English	143,252		+/-6,678
French (except Basque)	58,096		+/-4,432
French Canadian	9,705		+/-1,793
German	274,950		+/-10,731
Greek	89,279		+/-8,392
Hungarian	53,751		+/-5,561
Irish	415,287		+/-11,776
Italian	670,845		+/-14,166
Lithuanian	14,482		+/-1,814
Norwegian	22,300		+/-3,186
Polish	211,389		+/-10,059
Portuguese	15,905		+/-3,557
Russian	242,351		+/-9,898
Scotch-Irish	26,439		+/-3,613
Scottish	35,538		+/-3,796
Slovak	6,017		+/-1,132
Subsaharan African	183,006		+/-12,857
Swedish	20,910		+/-2,863
Swiss	8,624		+/-2,026
Ukrainian	54,925		+/-5,396
Welsh	11,614		+/-2,426
West Indian (excluding Hispanic origin groups)	615,313		+/-19,289

Source: U.S. Census Bureau, 2006 American Community Survey

Selected Social Characteristics in the United States: 2006
Data Set: 2006 American Community Survey
Survey: 2006 American Community Survey
Geographic Area: Longmont city, Colorado

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is the Census Bureau's Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states, counties, cities and towns and estimates of housing units for states and counties.

NOTE ABOUT RELIABILITY OF SOME ESTIMATES:

Reliability
use caution

Estimates marked with a yellow "use caution" reliability box (shown above) means that the coefficient of variation (CV) is above 0.30. The CV is defined as the standard error of an estimate divided by the mean of that estimate, measured as a percentage. Relatively, a lower CV means a more reliable estimate.

Selected Social Characteristics in the United States: 2006	Estimate	Reliability	Margin of Error
HOUSEHOLDS BY TYPE			
Total households	**31,545**		**+/-1,572**
Family households (families)	20,981		+/-1,687
With own children under 18 years	11,640		+/-1,283
Married-couple families	15,437		+/-1,589
With own children under 18 years	7,924		+/-1,169
Male householder, no wife present	2,566		+/-880
With own children under 18 years	1,668		+/-745
Female householder, no husband present	2,978		+/-791
With own children under 18 years	2,048		+/-680
Nonfamily households	10,564		+/-1,318
Householder living alone	7,542		+/-1,170
65 years and over	2,447		+/-703
Households with one or more people under 18 years	12,726		+/-1,457
Households with one or more people 65 years and over	5,372		+/-894
Average household size	2.68		+/-0.11
Average family size	3.20		+/-0.15
RELATIONSHIP			
Household population	**84,474**		**+/-4,547**
Householder	31,545		+/-1,572
Spouse	15,598		+/-1,506
Child	24,948		+/-2,365
Other relatives	5,657		+/-1,890
Nonrelatives	6,726		+/-2,220
Unmarried partner	2,659		+/-1,012
MARITAL STATUS			
Males 15 years and over	**32,452**		**+/-2,522**
Never married	9,671		+/-1,784
Now married, except separated	17,911		+/-1,750
Separated	686	use caution	+/-528
Widowed	694	use caution	+/-415
Divorced	3,490		+/-905
Females 15 years and over	**33,838**		**+/-2,054**
Never married	7,127		+/-977
Now married, except separated	17,947		+/-1,622
Separated	1,365	use caution	+/-713
Widowed	2,784		+/-726
Divorced	4,615		+/-1,016
FERTILITY			
Number of women 15 to 50 years old who had a birth in the past 12 months	**1,629**		**+/-691**
Unmarried women (widowed, divorced, and never married)	700	use caution	+/-418
Per 1,000 unmarried women	75	use caution	+/-42
Per 1,000 women 15 to 50 years old	73		+/-29
Per 1,000 women 15 to 19 years old	93	use caution	+/-122
Per 1,000 women 20 to 34 years old	109		+/-53
Per 1,000 women 35 to 50 years old	35	use caution	+/-26
GRANDPARENTS			
Number of grandparents living with own grandchildren under 18 years	**N**		**N**
Responsible for grandchildren	N		N
Years responsible for grandchildren			
Less than 1 year	N		N
1 or 2 years	N		N
3 or 4 years	N		N
5 or more years	N		N
Characteristics of grandparents responsible for own grandchildren under 18 years			
Who are female	N		N
Who are married	N		N
SCHOOL ENROLLMENT			
Population 3 years and over enrolled in school	**22,557**		**+/-2,364**
Nursery school, preschool	1,037		+/-441
Kindergarten	1,299		+/-557
Elementary school (grades 1-8)	9,001		+/-1,198
High school (grades 9-12)	5,640		+/-973
College or graduate school	5,580		+/-1,692
EDUCATIONAL ATTAINMENT			
Population 25 years and over	**54,865**		**+/-2,805**
Less than 9th grade	4,146		+/-1,533

9th to 12th grade, no diploma	3,842		+/-1,151
High school graduate (includes equivalency)	13,142		+/-1,615
Some college, no degree	10,794		+/-1,310
Associate's degree	3,204		+/-708
Bachelor's degree	14,234		+/-1,843
Graduate or professional degree	5,503		+/-1,115
Percent high school graduate or higher	85.4%		+/-3.2
Percent bachelor's degree or higher	36.0%		+/-3.6
VETERAN STATUS			
Civilian population 18 years and over	**61,968**		**+/-3,455**
Civilian veterans	6,456		+/-1,361
DISABILITY STATUS OF THE CIVILIAN NONINSTITUTIONALIZED POPULATION			
Population 5 years and over	**77,519**		**+/-4,017**
With a disability	8,058		+/-1,331
Population 5 to 15 years	**13,284**		**+/-1,501**
With a disability	387	use caution	+/-299
Population 16 to 64 years	**57,240**		**+/-3,528**
With a disability	4,984		+/-1,101
Population 65 years and over	**6,995**		**+/-1,156**
With a disability	2,687		+/-743
RESIDENCE 1 YEAR AGO			
Population 1 year and over	**83,595**		**+/-4,433**
Same house	62,556		+/-4,287
Different house in the U.S.	20,832		+/-4,009
Same county	12,583		+/-3,080
Different county	8,249		+/-2,330
Same state	3,990		+/-1,528
Different state	4,259		+/-1,841
Abroad	207	use caution	+/-184
PLACE OF BIRTH			
Total population	**84,880**		**+/-4,551**
Native	72,548		+/-3,995
Born in United States	71,746		+/-3,932
State of residence	28,705		+/-2,278
Different state	43,041		+/-3,354
Born in Puerto Rico, U.S. Island areas, or born abroad to American parent(s)	802	use caution	+/-508
Foreign born	12,332		+/-2,592
U.S. CITIZENSHIP STATUS			
Foreign-born population	**12,332**		**+/-2,592**
Naturalized U.S. citizen	1,937		+/-726
Not a U.S. citizen	10,395		+/-2,567
YEAR OF ENTRY			
Population born outside the United States	**13,134**		**+/-2,633**
Native	802	use caution	+/-508
Entered 2000 or later	0	use caution	+/-271
Entered before 2000	802	use caution	+/-508
Foreign born	12,332		+/-2,592
Entered 2000 or later	3,091		+/-1,518
Entered before 2000	9,241		+/-2,072
WORLD REGION OF BIRTH OF FOREIGN BORN			
Foreign-born population, excluding population born at sea	**N**		**N**
Europe	N		N
Asia	N		N
Africa	N		N
Oceania	N		N
Latin America	N		N
Northern America	N		N
LANGUAGE SPOKEN AT HOME			
Population 5 years and over	**N**		**N**
English only	N		N
Language other than English	N		N
Speak English less than "very well"	N		N
Spanish	N		N
Speak English less than "very well"	N		N
Other Indo-European languages	N		N
Speak English less than "very well"	N		N
Asian and Pacific Islander languages	N		N
Speak English less than "very well"	N		N
Other languages	N		N
Speak English less than "very well"	N		N
ANCESTRY			
Total population	**84,880**		**+/-4,551**
American	2,740		+/-1,053

Arab	46	use caution	+/-78
Czech	263	use caution	+/-203
Danish	664	use caution	+/-457
Dutch	2,365	use caution	+/-1,219
English	10,586		+/-1,734
French (except Basque)	2,604		+/-874
French Canadian	467	use caution	+/-316
German	21,948		+/-3,010
Greek	1,747	use caution	+/-1,497
Hungarian	228	use caution	+/-160
Irish	9,736		+/-1,999
Italian	4,791		+/-1,722
Lithuanian	0	use caution	+/-271
Norwegian	2,608		+/-838
Polish	3,064		+/-1,151
Portuguese	162	use caution	+/-163
Russian	477	use caution	+/-321
Scotch-Irish	2,729		+/-907
Scottish	3,277		+/-867
Slovak	300	use caution	+/-383
Subsaharan African	0	use caution	+/-271
Swedish	2,935		+/-990
Swiss	648	use caution	+/-631
Ukrainian	37	use caution	+/-62
Welsh	650	use caution	+/-381
West Indian (excluding Hispanic origin groups)	0	use caution	+/-271

Source: U.S. Census Bureau, 2006 American Community Survey

Appendix B
General Introduction

Thank you for your time today. My name is <Name>, and I will be working with you today. We will be evaluating a new design of the new ACS data table formats by having you work on several tasks. Your experience with the site is an essential part of our work. We are going to use your comments to give feedback to the developers of the site. Your comments and thoughts will help the developers make changes to improve the site. I did not create the site, so please do not feel like you have to hold back on your thoughts to be polite. Please share both your positive and negative reactions to the site. We are not evaluating you or your skills, but rather you are helping us see how well the site works.

First, I would like to ask you to read and sign this consent form. It explains the purpose of the session and informs you that we would like to videotape the session, with your permission. Only those of us connected with the project will review the tape. We will use it mainly as a memory aid. We may also use clips from the tape to illustrate key points about the design of the Web pages.

[Hand consent form; give time to read and sign; sign own name and date.]

[Start the tape when the participant signs the form.]

So today, you will be helping us test the usability of the new ACS data table formats. Your feedback is valuable, and we appreciate your help. We are going to do some eye tracking as well as have you work on some task scenarios that I will give you.

Before we get started, please take a moment to complete this computer usage and internet experience questionnaire.

[Hand computer experience form.]

For the next 60 minutes, I will ask you to work on 21 tasks. I would like you to tell me your impressions and thoughts about the Tables as you work through the tasks. I would like you to "think aloud" and talk to me about your decisions. So if you expect something to happen, tell me what you expect. If you expect to see some piece of information, tell me about what you expect. This means that as you work on a task, talk to me about what you are doing, what you are going to do, and why. Talk to me about why you clicked on a link or where you expect the link to take you. Also, although these different tables will presented to you one after another today, keep in mind that only one design will be chosen, so users will not have to worry about keeping track of the differences between them.

Finally, during the session, I will remind you to talk to me if you get quiet, not to interrupt your thought process simply to remind you to talk to me. Please focus on verbalizing what you are thinking and expecting to happen. We are interested in the reasoning behind your actions, not just what you are doing.

I ask that each time you start a task, please read the task out load, and once you have found the information you are looking for please state your answer aloud. For example,

say, “My answer is ---” or “This is my final answer.” After each task, I will return you to the next table where you can begin the next task.

Please remember to begin each task by reading the task question aloud as well as stating the final answer. As you work, please remember think aloud.

Before we get started, let’s practice thinking aloud. [This is an example of what this practice might involve: Please go to www.wtop.com. Describe your thought process as you select a new story that you find interesting to read.]

Now I am going to calibrate your eyes for the eye-tracking. I am going to have you position yourself in front of the screen so that you can see your nose in the reflection at the bottom of the monitor. To calibrate your eyes, please follow the blue dot across the screen with your eyes.

[Do Calibration]

Now that we have your eyes calibrated, we are ready to begin.

Appendix C
Consent Form

Consent Form

Usability Study of the ACS Data Tables

Each year the Census Bureau conducts many different usability evaluations. For example, the Census Bureau routinely tests the wording, layout and behavior of products, such as Web sites and online surveys, in order to obtain the best information possible from respondents.

You have volunteered to take part in a study to improve the usability of the ACS data tables. In order to have a complete record of your comments, your usability session will be videotaped. We plan to use the tapes to improve the design of the product. Staff directly involved in the usable design research project will have access to the tapes. Your participation is voluntary and your answers will remain strictly confidential.

This usability study is being conducted under the authority of Title 13 USC. The OMB control number for this study is 0607-0725. This valid approval number legally certifies this information collection.

I have volunteered to participate in this Census Bureau usability study, and I give permission for my tapes to be used for the purposes stated above.

Participant's Name: _____

Participant's Signature: _____ Date: _____

Researcher's Name: _____

Researcher's Signature: _____ Date: _____

Appendix D
Questionnaire on Computer Use and Internet Experience

1. Do you use a computer at home or at work or both?
   (Check all that apply.)
   ___Home
   ___Work
   ___Somewhere else, such as school, library, etc.

2. If you have a computer at home,
   a. What kind of modem do you use at home?
   ___Dial-up
   ___Cable
   ___DSL
   ___Wireless (Wi-Fi)
   ___Other _____
   ___Don't know _____

   b. Which browser do you typically use at home? Please indicate the version if you can recall it.
   ___Firefox
   ___Internet Explorer
   ___Netscape
   ___Other _____
   ___Don't know

   c. What operating system does your home computer run in?
   ___MAC OS
   ___Windows 95
   ___Windows 2000
   ___Windows XP
   ___Windows Vista
   ___Other _____
   ___Don't know

3. On average, about how many hours do you spend on the Internet per day?
   ___0 hours
   ___1-3 hours
   ___4-6 hours
   ___7or more hours

4. Please rate your overall experience with the following:
Circle one number.

	No experience								**Very experienced**
Computers	1	2	3	4	5	6	7	8	9
Internet	1	2	4	5	5	6	7	8	9

5. What computer applications do you use?

Mark (X) for all that apply

___E-mail
___Internet
___Word processing (MS-Word, WordPerfect, etc.)
___Spreadsheets (Excel, Lotus, Quattro, ect.)
___Accounting or tax software
___Engineering, scientific, or statistical software
___Other applications, please specify_____

For the following questions, please circle one number.	**Not Comfortable**				**Comfortable**
6. How *comfortable* are you in learning to navigate new Web sites?	1	2	3	4	5
7. Computer windows can minimize, resized, and scrolled through. How *comfortable* are you in manipulating a window?	1	2	3	4	5
8. How *comfortable* are you using, and navigating through the Internet?	1	2	3	4	5

	Never				**Very Often**
9. How *often* do you work with any type of data through a computer?	1	2	3	4	5
10. How *often* do you perform complex analyses of data through a computer?	1	2	3	4	5
11. How *often* do you use the Internet or Web sites to find information? (e.g., printed reports, news articles, data tables, blogs, etc.)	1	2	3	4	5

	Not familiar				**Very familiar**
12. How *familiar* are you with the Census (terms, data, etc)?	1	2	3	4	5
13. How *familiar* are you with the current American Community Survey (ACS) and American FactFinder (AFF) sites (terms, data, etc.)?	1	2	3	4	5

Appendix E
Task list

The tasks for the baseline condition were identical to the prototype tasks except for the probe questions associated with the first example that each person sees and the questions at the end of the session. This is because we were probing about the Margin of Error (MOE) column and not the reliability indicator column and legend box.

Tasks for Baseline Evaluation [California example]

1. What is the first thing that you noticed about these tables? [Difficulty: Easy]

2. Your supervisor asks you to find some information about the number of women ages 15 to 50 who gave birth in the past 12 months for your hometown of Wilmington, DE [975+/-478]. He wants you to also provide the same number for California as a point of comparison [515,991+/-13,067]. What would you tell your boss based on the tables? [Difficulty: Hard]

- Probe if MOE not mentioned: Would you tell your boss about the differences in reliability or quality between the two estimates?

Tasks for Prototype Evaluation

Except for the first and last task, which remained the same for each block due to their qualitative and position-related content, the order of these blocks of tasks has been randomized for each participant. Some of the questions will appear only with the first prototype that is presented to each participant. These questions appear in ***bolded and italicized*** font below.

Four-level Prototype (Poor: Wilmington, DE and Good: California):

Open both prototypes. They are in HTML format in side-by-side tabs.

1. What is the first thing that you noticed about these tables? [Difficulty: Easy]

- Probe if indicator is mentioned: What do the colors represent?
  - IF quality/reliability: How is quality measured here or what measure is used to determine the level of quality? ***What is a CV? Is a higher or lower CV associated with better data quality/reliability? How is the CV different from a Margin of Error (MOE)? Which would you prefer to use?***
- Probe if indicator is not mentioned:
  - Did you notice the colors in the tables? What do the colors represent?
    - IF quality: How is quality measured here or what measure is used to determine the level of quality? ***What is a CV? Is a higher or lower CV associated with better data***

quality/reliability? How is the CV different from a Margin of Error (MOE)? Which would you prefer to use?

2. Your supervisor asks you to find some information about the number of women ages 15 to 50 who gave birth in the past 12 months for your hometown of Wilmington, DE [good|975+/-478]. He wants you to also provide the same number for California as a point of comparison [excellent|515,991+/-13,067]. What would you tell your boss based on the tables? [Difficulty: Hard]

- Probe if indicator not mentioned: Would you tell your boss about the differences in reliability or quality between the two estimates?

3. You are researching a paper and need to find the number of people of West Indian descent in Wilmington, DE [poor|0+/-261]. What do you report in the paper based on your findings in the tables? [Difficulty: Easy]

4. Find the total number of people with German ancestry [good|4,206|+/-1,083] and the total number of people with Slovak ancestry [poor|70|+/-116] for Wilmington, DE. Which category of ancestry do you think is a better estimate in terms of data quality? [German]. Please explain why you think this is a better estimate of data quality. [Difficulty: Hard]

5. a. For both California and Wilmington, DE, please find three estimates that have the most reliability. [any 3 "excellent" estimates] [Difficulty: Hard]

b. Please find three that are low in reliability. [any 3 "poor" estimates] [Difficulty: Hard]

6. You are asked to report to state leaders the number of people of Italian descent living in California. What answer would you give them? [1,560,870|good|+/- 24,224] Would you recommend using this number? [Yes] Why or why not? [excellent reliability] [Difficulty: Easy]

7. Look at the table for Wilmington, DE and table for California. In general, which area has more reliable estimates? [California]. Explain why you gave this answer. [Difficulty: Easy]

Three-Level Prototype (Poor: Yakima, WA and Good: Cook County, IL)

Open both prototypes side-by-side.

1. What is the first thing that you noticed about these tables? [Difficulty: Easy]

- Probe if indicator is mentioned: What do the colors represent?
  - IF quality/reliability: How is quality measured here or what measure is used to determine the level of quality? ***What is a CV? Is a higher or lower CV associated with better data quality/reliability?***
- Probe if indicator is not mentioned:
  - Did you notice the colors in the tables? What do the colors represent?
    - IF quality: How is quality measured here or what measure is used to determine the level of quality? ***What is a CV? Is a higher or lower CV associated with better data quality/reliability?***

2. Your friend asks you to find some information about the number of males who are married and not separated for her hometown of Yakima, WA [fair|779|+/-466]. She wants to know the same number for Cook County, IL for comparison [good|941,158|+/-13,028]. What would you tell your friend based on the tables? [Difficulty: Hard]

- Probe if indicator not mentioned: Would you tell your friend about the differences in reliability or quality between the two estimates?

3. Find the total number of people with English [good|9,178|+/-1,815] ancestry and the total number of people with Welsh ancestry [poor|60|+/- 99] for Yakima, WA. Which category of ancestry do you think is a better estimate in terms of data quality? [English]. Please explain why you think this is a better estimate of data quality. [Difficulty Hard]

4. a. For both Cook County, IL and Yakima, WA, please find three estimates that have the most reliability. [any 3 "good estimates] [Difficulty: Hard]

b. Please find three that are low in reliability. [any 3 "poor estimates]

5. You are researching a marketing research report for work and need to find the number of people of Swiss descent in Yakima, WA. What do you report in the paper based on your findings in the tables? [Difficulty: Easy]

6. You are asked to report to state leaders the number of people of Polish descent living in Cook County, IL. What answer would you give them? [good|507,498|+/-13,952] Would you recommend using this number? Why or why not? [Difficulty: Easy]

7. Look at the table for Yakima, WA and the table for Cook County, IL. In general, which area has more reliable estimates? [Cook County, IL]. Explain why you gave this answer. [Difficulty: Easy]

Two-Level Prototype (Poor: Longmont, CO and Good: New York, NY)

Open both prototypes side-by-side.

1. What is the first thing that you noticed about these tables? [Difficulty: Easy]

- Probe if indicator is mentioned: What do the colors represent?
  - IF quality/reliability: ***How is quality measured here or what measure is used to determine the level of quality? What is a CV? Is a higher or lower CV associated with better data quality/reliability?***
- Probe if indicator is not mentioned:
  - Did you notice the colors in the tables? What do the colors represent?
    - IF quality: How is quality measured here or what measure is used to determine the level of quality? ***What is a CV? Is a higher or lower CV associated with better data quality/reliability?***

2. You are asked to report to state leaders the number of people of Portuguese descent living in New York, NY. What answer would you give them? [unmarked|15,905|+/-3,557] Would you recommend using this number? Why or why not? [Difficulty: Easy]

3. You are researching a marketing research report for work and need to find the number of people of Dutch descent in Longmont, CO [poor|2,365|+/-1,219]. What do you report in the paper based on your findings in the tables? [Difficulty: Easy]

4. Find the total number of people with Subsaharan African ancestry [poor|0|+/-271and the total number of people with Norwegian [unmarked|2,608|+/-838] ancestry for Longmont, CO. Which category of ancestry do you think is a better estimate in terms

of data quality? [Norwegian]. Please explain why you think this is a better estimate of data quality. [Difficulty: Hard]

5. a. For both New York, NY and Longmont, CO, please find three estimates that have the most reliability. [any 3 unmarked estimates] [Difficulty: Hard]

b. Please find three that are low in reliability. [any 3 "poor estimates] [Difficulty: Hard]

6. Your coworker asks you to find some information about the number of people age 5-15 with a disability for his hometown of Longmont, CO [poor|387|+/-1,501]. He wants to know the same number for New York, NY for comparison [unmarked|62,075|+/-5,362]. What would you tell your coworker based on the tables? [Difficulty: Hard]

- Probe if indicator not mentioned: Would you tell your coworker about the differences in reliability or quality between the two estimates?

7. Look at the table for Longmont, CO and the table for New York, NY. In general, which area has more reliable estimates? [New York, NY]. Explain why you gave this answer. [Difficulty: Easy]

Appendix F: **Questionnaire for User Interaction Satisfaction (QUIS)**

Please circle the numbers that most appropriately reflect your impressions about using the new ACS data tables.

1. Overall reaction to the new ACS data tables:	terrible	1	2	3	4	5	6	7	8	9	wonderful not applicable
2. Overall table format:	confusing	1	2	3	4	5	6	7	8	9	clear not applicable
3. Use of terminology throughout the tables:	inconsistent	1	2	3	4	5	6	7	8	9	consistent not applicable
4. Information displayed in the tables:	inadequate	1	2	3	4	5	6	7	8	9	adequate not applicable
5. Arrangement of information in the tables:	logical	1	2	3	4	5	6	7	8	9	illogical not applicable
6. Tasks can be performed in a straight-forward manner:	never	1	2	3	4	5	6	7	8	9	always not applicable
7. Color-coded reliability indicator for the tables:	confusing	1	2	3	4	5	6	7	8	9	clear not applicable
8. Overall experience of finding information:	difficult	1	2	3	4	5	6	7	8	9	easy not applicable

10. Additional Comments:

Appendix G: Debriefing Questions

Baseline condition: only walk through one time because all tables will be the same

1. Can you walk me through your thinking on why you marked (a particular QUIS item) especially low/high? (Do this for several low/high QUIS ratings).

2. What do you think of the basic screen layout?

    a. Overall?

3. How easy or difficult do you feel it was to complete the tasks? What made a task easy or difficult?

4. What did you like best about the tables?

5. What did you like least about the tables?

6. What is something that you feel should be changed?

7. What is something that you feel should stay the same?

8. Is there anything you'd like to mention that we haven't talked about?

Overall

After the last tables only:

Open a pair of current tables and all of the "lower reliability" prototype tables (Wilmington, DE, Longmont, CO, and Yakima, WA). Allow the participant to look at each prototype for a few seconds before opening the next one.

9. The tables that you have been using for this session represent how the ACS data tables currently look. Here are some new formats that might be used for future versions of the tables. Which version of these tables do you think you would find easiest to use?

- What features of that table did you like best?

Which do you think that you would find the most difficult to use?

- What features of that table did you like least?

Open one pair of current tables.

10. Do you prefer the new tables or the current tables? Please explain your answer.

- Does the MOE or the CV tell you more about the reliability of the estimates? Which would you rather use to answer questions like those you worked with today?

Debriefing Questions -

Prototype conditions: walk through these for each pair of prototypes

1. Can you walk me through your thinking on why you marked (a particular QUIS item) especially low/high? (Do this for several low/high QUIS ratings). Also, do this for each prototype.

2. What do you think of the basic screen layout?

    a. Overall?
    b. The color scheme?

3. How easy or difficult do you feel it was to complete the tasks? What made a task easy or difficult?

4. Do you think the new color-coding scheme for the ACS table helped you find accurate answers?

5. Do you think the new color-coded ACS table helped you to find information quickly ?

6. How satisfied are you with your experiences using the ACS data reliability indicators?

7. What did you like best about the table?

8. What did you like least about the table?

9. What is something that you feel should be changed?

10. What is something that you feel should stay the same?

11. Is there anything you'd like to mention that we haven't talked about?

<u>Overall</u>

After the last prototype only:

Open all pairs of prototypes.
12. Here are three versions of this color-coded reliability indicator. Some have different levels of indicator and some have different text to describe their meaning.

- Which of these tables make it easiest or harder to find information about data quality/reliability? What about this table makes it the easiest to use?

Which did you find the most difficult to use?

- What features of that table did you like least?

Open one pair of current tables.

13. These are how the ACS data tables currently look. Do you prefer the new tables or the current tables? Please explain your answer.

- Does the MOE or the CV tell you more about the reliability of the estimates? Which would you rather use to answer questions like those you worked with today?

Appendix H: Full Accuracy Results Data

Participant	Geography	Task	Order of Appearance	Accuracy
Baseline P1	California	1	3	NA
Baseline P1	California	2	3	1
Baseline P1	California	3	3	1
Baseline P1	California	4	3	1
Baseline P1	California	5	3	1
Baseline P1	California	6	3	0.5
Baseline P1	California	7	3	1
Baseline P1	CookCounty	1	2	NA
Baseline P1	CookCounty	2	2	1
Baseline P1	CookCounty	3	2	0.5
Baseline P1	CookCounty	4	2	1
Baseline P1	CookCounty	5	2	1
Baseline P1	CookCounty	6	2	0
Baseline P1	CookCounty	7	2	1
Baseline P1	New York	1	1	NA
Baseline P1	New York	2	1	1
Baseline P1	New York	3	1	1
Baseline P1	New York	4	1	1
Baseline P1	New York	5	1	1
Baseline P1	New York	6	1	1
Baseline P1	New York	7	1	0
BaselineP2	California	1	2	NA
BaselineP2	California	2	2	1
BaselineP2	California	3	2	0.5
BaselineP2	California	4	2	1
BaselineP2	California	5	2	0.5
BaselineP2	California	6	2	1
BaselineP2	California	7	2	1
BaselineP2	CookCounty	1	1	NA
BaselineP2	CookCounty	2	1	0.5
BaselineP2	CookCounty	3	1	1
BaselineP2	CookCounty	4	1	1
BaselineP2	CookCounty	5	1	0.5
BaselineP2	CookCounty	6	1	1
BaselineP2	CookCounty	7	1	0
BaselineP2	New York	1	3	NA
BaselineP2	New York	2	3	0.5
BaselineP2	New York	3	3	1
BaselineP2	New York	4	3	0.666667
BaselineP2	New York	5	3	0.5
BaselineP2	New York	6	3	0.5
BaselineP2	New York	7	3	1
BaselineP3	California	1	1	NA
BaselineP3	California	2	1	1
BaselineP3	California	3	1	0.5

BaselineP3	California	4	1	0.666667
BaselineP3	California	5	1	0
BaselineP3	California	6	1	1
BaselineP3	California	7	1	0
BaselineP3	CookCounty	1	3	NA
BaselineP3	CookCounty	2	3	1
BaselineP3	CookCounty	3	3	0.666667
BaselineP3	CookCounty	4	3	0.5
BaselineP3	CookCounty	5	3	1
BaselineP3	CookCounty	6	3	1
BaselineP3	CookCounty	7	3	0
BaselineP3	New York	1	2	NA
BaselineP3	New York	2	2	1
BaselineP3	New York	3	2	0
BaselineP3	New York	4	2	1
BaselineP3	New York	5	2	0
BaselineP3	New York	6	2	1
Prototype P1	California	1	2	NA
Prototype P1	California	2	2	1
Prototype P1	California	3	2	1
Prototype P1	California	4	2	1
Prototype P1	California	5	2	1
Prototype P1	California	6	2	1
Prototype P1	California	7	2	1
Prototype P1	CookCounty	1	1	NA
Prototype P1	CookCounty	2	1	0.5
Prototype P1	CookCounty	3	1	0
Prototype P1	CookCounty	4	1	1
Prototype P1	CookCounty	5	1	0
Prototype P1	CookCounty	6	1	1
Prototype P1	CookCounty	7	1	1
Prototype P1	New York	1	3	NA
Prototype P1	New York	2	3	1
Prototype P1	New York	3	3	1
Prototype P1	New York	4	3	1
Prototype P1	New York	5	3	1
Prototype P1	New York	6	3	1
Prototype P1	New York	7	3	1
Prototype P2	California	1	2	NA
Prototype P2	California	2	2	1
Prototype P2	California	3	2	1
Prototype P2	California	4	2	1
Prototype P2	California	5	2	1
Prototype P2	California	6	2	1
Prototype P2	California	7	2	1
Prototype P2	CookCounty	1	3	NA
Prototype P2	CookCounty	2	3	1
Prototype P2	CookCounty	3	3	1

Prototype P2	CookCounty	4	3	1
Prototype P2	CookCounty	5	3	0.5
Prototype P2	CookCounty	6	3	1
Prototype P2	CookCounty	7	3	1
Prototype P2	New York	1	1	NA
Prototype P2	New York	2	1	1
Prototype P2	New York	3	1	1
Prototype P2	New York	4	1	1
Prototype P2	New York	5	1	1
Prototype P2	New York	6	1	1
Prototype P2	New York	7	1	1
Prototype P3	California	1	2	NA
Prototype P3	California	2	2	0.5
Prototype P3	California	3	2	1
Prototype P3	California	4	2	1
Prototype P3	California	5	2	1
Prototype P3	California	6	2	1
Prototype P3	California	7	2	1
Prototype P3	CookCounty	1	3	NA
Prototype P3	CookCounty	2	3	1
Prototype P3	CookCounty	3	3	1
Prototype P3	CookCounty	4	3	1
Prototype P3	CookCounty	5	3	1
Prototype P3	CookCounty	6	3	1
Prototype P3	CookCounty	7	3	1
Prototype P3	New York	1	1	NA
Prototype P3	New York	2	1	0.5
Prototype P3	New York	3	1	1
Prototype P3	New York	4	1	1
Prototype P3	New York	5	1	1
Prototype P3	New York	6	1	1
Prototype P3	New York	7	1	1
Prototype P4	California	1	4	NA
Prototype P4	California	2	4	1
Prototype P4	California	3	4	1
Prototype P4	California	4	4	1
Prototype P4	California	5	4	1
Prototype P4	California	6	4	1
Prototype P4	California	7	4	0
Prototype P4	CookCounty	1	3	NA
Prototype P4	CookCounty	2	3	1
Prototype P4	CookCounty	3	3	1
Prototype P4	CookCounty	4	3	1
Prototype P4	CookCounty	5	3	1
Prototype P4	CookCounty	6	3	1
Prototype P4	CookCounty	7	3	0
Prototype P4	New York	1	1	NA
Prototype P4	New York	2	1	1

Prototype P4	New York	3	1	1
Prototype P4	New York	4	1	1
Prototype P4	New York	5	1	1
Prototype P4	New York	6	1	1
Prototype P4	New York	7	1	1
Prototype P5	California	1	3	NA
Prototype P5	California	2	3	1
Prototype P5	California	3	3	0.5
Prototype P5	California	4	3	1
Prototype P5	California	5	3	1
Prototype P5	California	6	3	1
Prototype P5	California	7	3	1
Prototype P5	CookCounty	1	1	NA
Prototype P5	CookCounty	2	1	1
Prototype P5	CookCounty	3	1	1
Prototype P5	CookCounty	4	1	1
Prototype P5	CookCounty	5	1	0.5
Prototype P5	CookCounty	6	1	1
Prototype P5	CookCounty	7	1	1
Prototype P5	New York	1	2	NA
Prototype P5	New York	2	2	1
Prototype P5	New York	3	2	0.5
Prototype P5	New York	4	2	1
Prototype P5	New York	5	2	0.5
Prototype P5	New York	6	2	1
Prototype P5	New York	7	2	1
Prototype P6	California	1	1	NA
Prototype P6	California	2	1	1
Prototype P6	California	3	1	0.5
Prototype P6	California	4	1	0.333333
Prototype P6	California	5	1	1
Prototype P6	California	6	1	1
Prototype P6	California	7	1	1
Prototype P6	CookCounty	1	3	NA
Prototype P6	CookCounty	2	3	1
Prototype P6	CookCounty	3	3	1
Prototype P6	CookCounty	4	3	1
Prototype P6	CookCounty	5	3	1
Prototype P6	CookCounty	6	3	1
Prototype P6	CookCounty	7	3	1
Prototype P6	New York	1	2	NA
Prototype P6	New York	2	2	0.5
Prototype P6	New York	3	2	1
Prototype P6	New York	4	2	1
Prototype P6	New York	5	2	1
Prototype P6	New York	6	2	1
Prototype P6	New York	7	2	1
Prototype P7	California	1	1	NA

Prototype P7	California	2	1	1
Prototype P7	California	3	1	0
Prototype P7	California	4	1	1
Prototype P7	California	5	1	1
Prototype P7	California	6	1	1
Prototype P7	California	7	1	1
Prototype P7	CookCounty	1	3	NA
Prototype P7	CookCounty	2	3	1
Prototype P7	CookCounty	3	3	1
Prototype P7	CookCounty	4	3	1
Prototype P7	CookCounty	5	3	1
Prototype P7	CookCounty	6	3	0
Prototype P7	CookCounty	7	3	1
Prototype P7	New York	1	2	NA
Prototype P7	New York	2	2	1
Prototype P7	New York	3	2	0.5
Prototype P7	New York	4	2	1
Prototype P7	New York	5	2	1
Prototype P7	New York	6	2	0.5
Prototype P7	New York	7	2	1
Prototype P8	California	1	2	NA
Prototype P8	California	2	2	1
Prototype P8	California	3	2	1
Prototype P8	California	4	2	1
Prototype P8	California	5	2	1
Prototype P8	California	6	2	0.5
Prototype P8	California	7	2	1
Prototype P8	CookCounty	1	1	NA
Prototype P8	CookCounty	2	1	1
Prototype P8	CookCounty	3	1	1
Prototype P8	CookCounty	4	1	1
Prototype P8	CookCounty	5	1	1
Prototype P8	CookCounty	6	1	1
Prototype P8	CookCounty	7	1	1
Prototype P8	New York	1	3	NA
Prototype P8	New York	2	3	0.5
Prototype P8	New York	3	3	1
Prototype P8	New York	4	3	0.666667
Prototype P8	New York	5	3	0
Prototype P8	New York	6	3	1
Prototype P8	New York	7	3	0
Prototype P9	California	1	1	NA
Prototype P9	California	2	1	0
Prototype P9	California	3	1	1
Prototype P9	California	4	1	1
Prototype P9	California	5	1	1
Prototype P9	California	6	1	1
Prototype P9	California	7	1	1

Prototype P9	CookCounty	1	3	NA
Prototype P9	CookCounty	2	3	1
Prototype P9	CookCounty	3	3	1
Prototype P9	CookCounty	4	3	1
Prototype P9	CookCounty	5	3	0.5
Prototype P9	CookCounty	6	3	1
Prototype P9	CookCounty	7	3	1
Prototype P9	New York	1	2	NA
Prototype P9	New York	2	2	1
Prototype P9	New York	3	2	1
Prototype P9	New York	4	2	0.333333
Prototype P9	New York	5	2	1
Prototype P9	New York	6	2	1
Prototype P9	New York	7	2	1

Appendix I: Full Efficiency Data Results

Participant	Geography	Task	Order of Appearance	Efficiency
Baseline P1	California	1	3	1:47
Baseline P1	California	2	3	1:07
Baseline P1	California	3	3	0:48
Baseline P1	California	4	3	0:42
Baseline P1	California	5	3	1:28
Baseline P1	California	6	3	0:30
Baseline P1	California	7	3	0:29
Baseline P1	CookCounty	1	2	0:21
Baseline P1	CookCounty	2	2	1:47
Baseline P1	CookCounty	3	2	1:02
Baseline P1	CookCounty	4	2	2:24
Baseline P1	CookCounty	5	2	1:03
Baseline P1	CookCounty	6	2	0:54
Baseline P1	CookCounty	7	2	0:29
Baseline P1	New York	1	1	1:47
Baseline P1	New York	2	1	1:25
Baseline P1	New York	3	1	1:34
Baseline P1	New York	4	1	1:04
Baseline P1	New York	5	1	4:08
Baseline P1	New York	6	1	1:15
Baseline P1	New York	7	1	1:51
BaselineP2	California	1	2	0:54
BaselineP2	California	2	2	0:43
BaselineP2	California	3	2	0:45
BaselineP2	California	4	2	0:30
BaselineP2	California	5	2	4:24
BaselineP2	California	6	2	0:35
BaselineP2	California	7	2	2:00
BaselineP2	CookCounty	1	1	0:53
BaselineP2	CookCounty	2	1	0:51
BaselineP2	CookCounty	3	1	1:12
BaselineP2	CookCounty	4	1	3:55
BaselineP2	CookCounty	5	1	0:20
BaselineP2	CookCounty	6	1	0:27
BaselineP2	CookCounty	7	1	1:21
BaselineP2	New York	1	3	0:39
BaselineP2	New York	2	3	0:31
BaselineP2	New York	3	3	0:26
BaselineP2	New York	4	3	1:35
BaselineP2	New York	5	3	3:29
BaselineP2	New York	6	3	1:53
BaselineP2	New York	7	3	1:00
Prototype P1	California	1	2	0:28
Prototype P1	California	2	2	1:12
Prototype P1	California	3	2	0:46

Prototype P1	California	4	2	0:59
Prototype P1	California	5	2	1:39
Prototype P1	California	6	2	0:35
Prototype P1	California	7	2	0:24
Prototype P1	CookCounty	1	1	2:23
Prototype P1	CookCounty	2	1	1:55
Prototype P1	CookCounty	3	1	1:57
Prototype P1	CookCounty	4	1	2:22
Prototype P1	CookCounty	5	1	0:33
Prototype P1	CookCounty	6	1	0:47
Prototype P1	CookCounty	7	1	0:44
Prototype P1	New York	1	3	0:45
Prototype P1	New York	2	3	1:00
Prototype P1	New York	3	3	0:39
Prototype P1	New York	4	3	0:51
Prototype P1	New York	5	3	1:27
Prototype P1	New York	6	3	1:44
Prototype P1	New York	7	3	0:34
Prototype P2	California	1	2	0:43
Prototype P2	California	2	2	1:01
Prototype P2	California	3	2	1:50
Prototype P2	California	4	2	0:29
Prototype P2	California	5	2	2:10
Prototype P2	California	6	2	1:02
Prototype P2	California	7	2	0:37
Prototype P2	CookCounty	1	3	0:33
Prototype P2	CookCounty	2	3	0:55
Prototype P2	CookCounty	3	3	0:35
Prototype P2	CookCounty	4	3	1:14
Prototype P2	CookCounty	5	3	1:02
Prototype P2	CookCounty	6	3	0:14
Prototype P2	CookCounty	7	3	0:35
Prototype P2	New York	1	1	1:30
Prototype P2	New York	2	1	3:09
Prototype P2	New York	3	1	1:15
Prototype P2	New York	4	1	1:03
Prototype P2	New York	5	1	1:50
Prototype P2	New York	6	1	1:35
Prototype P2	New York	7	1	0:38
Prototype P3	California	1	2	0:34
Prototype P3	California	2	2	0:48
Prototype P3	California	3	2	0:30
Prototype P3	California	4	2	0:45
Prototype P3	California	5	2	1:41
Prototype P3	California	6	2	0:31
Prototype P3	California	7	2	0:26
Prototype P3	CookCounty	1	3	0:41
Prototype P3	CookCounty	2	3	1:18

Prototype P3	CookCounty	3	3	0:39
Prototype P3	CookCounty	4	3	1:26
Prototype P3	CookCounty	5	3	0:30
Prototype P3	CookCounty	6	3	0:19
Prototype P3	CookCounty	7	3	0:26
Prototype P3	New York	1	1	5:33
Prototype P3	New York	2	1	1:53
Prototype P3	New York	3	1	0:56
Prototype P3	New York	4	1	1:14
Prototype P3	New York	5	1	4:29
Prototype P3	New York	6	1	2:09
Prototype P3	New York	7	1	0:48
Prototype P4	California	1	4	3:35
Prototype P4	California	2	4	4:13
Prototype P4	California	3	4	1:28
Prototype P4	California	4	4	1:04
Prototype P4	California	5	4	1:26
Prototype P4	California	6	4	0:29
Prototype P4	California	7	4	0:06
Prototype P4	CookCounty	1	3	1:42
Prototype P4	CookCounty	2	3	2:01
Prototype P4	CookCounty	3	3	0:40
Prototype P4	CookCounty	4	3	1:04
Prototype P4	CookCounty	5	3	0:37
Prototype P4	CookCounty	6	3	0:36
Prototype P4	CookCounty	7	3	1:17
Prototype P4	New York	1	1	2:11
Prototype P4	New York	2	1	1:35
Prototype P4	New York	3	1	1:17
Prototype P4	New York	4	1	1:14
Prototype P4	New York	5	1	2:27
Prototype P4	New York	6	1	1:32
Prototype P4	New York	7	1	0:16
Prototype P5	California	1	3	0:18
Prototype P5	California	2	3	0:43
Prototype P5	California	3	3	0:27
Prototype P5	California	4	3	0:37
Prototype P5	California	5	3	1:13
Prototype P5	California	6	3	0:24
Prototype P5	California	7	3	0:11
Prototype P5	CookCounty	1	1	0:19
Prototype P5	CookCounty	2	1	1:49
Prototype P5	CookCounty	3	1	1:27
Prototype P5	CookCounty	4	1	1:57
Prototype P5	CookCounty	5	1	0:22
Prototype P5	CookCounty	6	1	0:35
Prototype P5	CookCounty	7	1	0:52
Prototype P5	New York	1	2	2:11

Prototype P5	New York	2	2	0:36
Prototype P5	New York	3	2	1:00
Prototype P5	New York	4	2	0:59
Prototype P5	New York	5	2	4:51
Prototype P5	New York	6	2	1:22
Prototype P5	New York	7	2	0:21
Prototype P6	California	1	1	2:03
Prototype P6	California	2	1	1:29
Prototype P6	California	3	1	0:56
Prototype P6	California	4	1	1:30
Prototype P6	California	5	1	2:43
Prototype P6	California	6	1	1:09
Prototype P6	California	7	1	1:15
Prototype P6	CookCounty	1	3	0:21
Prototype P6	CookCounty	2	3	3:30
Prototype P6	CookCounty	3	3	0:40
Prototype P6	CookCounty	4	3	2:40
Prototype P6	CookCounty	5	3	0:35
Prototype P6	CookCounty	6	3	1:04
Prototype P6	CookCounty	7	3	0:33
Prototype P6	New York	1	2	2:11
Prototype P6	New York	2	2	1:53
Prototype P6	New York	3	2	0:52
Prototype P6	New York	4	2	1:00
Prototype P6	New York	5	2	3:13
Prototype P6	New York	6	2	2:00
Prototype P6	New York	7	2	0:26
Prototype P7	California	1	1	0:54
Prototype P7	California	2	1	2:15
Prototype P7	California	3	1	1:22
Prototype P7	California	4	1	0:53
Prototype P7	California	5	1	1:55
Prototype P7	California	6	1	0:42
Prototype P7	California	7	1	0:36
Prototype P7	CookCounty	1	3	1:37
Prototype P7	CookCounty	2	3	3:09
Prototype P7	CookCounty	3	3	0:39
Prototype P7	CookCounty	4	3	2:32
Prototype P7	CookCounty	5	3	1:01
Prototype P7	CookCounty	6	3	1:05
Prototype P7	CookCounty	7	3	1:13
Prototype P7	New York	1	2	2:11
Prototype P7	New York	2	2	1:18
Prototype P7	New York	3	2	0:43
Prototype P7	New York	4	2	0:36
Prototype P7	New York	5	2	1:01
Prototype P7	New York	6	2	1:22
Prototype P7	New York	7	2	0:36

Prototype P8	California	1	2	5:36
Prototype P8	California	2	2	3:53
Prototype P8	California	3	2	2:06
Prototype P8	California	4	2	4:13
Prototype P8	California	5	2	3:28
Prototype P8	California	6	2	0:59
Prototype P8	California	7	2	2:48
Prototype P8	CookCounty	1	1	1:03
Prototype P8	CookCounty	2	1	1:33
Prototype P8	CookCounty	3	1	1:24
Prototype P8	CookCounty	4	1	3:22
Prototype P8	CookCounty	5	1	0:56
Prototype P8	CookCounty	6	1	0:35
Prototype P8	CookCounty	7	1	1:04
Prototype P8	New York	1	3	2:11
Prototype P8	New York	2	3	1:06
Prototype P8	New York	3	3	0:00
Prototype P8	New York	4	3	0:02
Prototype P8	New York	5	3	0:01
Prototype P8	New York	6	3	0:05
Prototype P8	New York	7	3	0:00
Prototype P9	California	1	1	1:40
Prototype P9	California	2	1	1:56
Prototype P9	California	3	1	1:05
Prototype P9	California	4	1	1:38
Prototype P9	California	5	1	1:47
Prototype P9	California	6	1	0:46
Prototype P9	California	7	1	0:08
Prototype P9	CookCounty	1	3	1:49
Prototype P9	CookCounty	2	3	1:09
Prototype P9	CookCounty	3	3	0:58
Prototype P9	CookCounty	4	3	2:25
Prototype P9	CookCounty	5	3	0:27
Prototype P9	CookCounty	6	3	0:29
Prototype P9	CookCounty	7	3	0:46
Prototype P9	New York	1	2	2:11
Prototype P9	New York	2	2	1:51
Prototype P9	New York	3	2	1:16
Prototype P9	New York	4	2	0:23
Prototype P9	New York	5	2	1:23
Prototype P9	New York	6	2	1:53
Prototype P9	New York	7	2	0:16