Medium-Fidelity Usability Evaluation for the American Community Survey Website

Using Eye-Tracking Data to Examine Fixation Differences by Task Performance

Temika Holland[(✉)] and Erica Olmsted-Hawala

U.S. Census Bureau, Center for Survey Measurement, 4600 Silver Hill Road, Washington DC 20233, USA
{Temika.Holland,Erica.L.Olmsted.Hawala}@census.gov

Abstract. The American Community Survey (ACS) website provides supplementary information about ACS participation and about ACS data (e.g., data collection, data utilization, survey procedures, etc.). Additionally, the ACS website is a portal to the American Fact Finder (AFF) for access to ACS data. The U.S. Census Bureau is undergoing a new initiative to change the look and feel of Census sites, and various design features have been modified on a web based prototype for the redesigned American Community Survey (ACS) website, including navigational tools and layout. Feedback on whether users of the site would be able to obtain the information they need given the new design features was warranted. The site was tested in its early stages of development using a web-based prototype with limited functionality (i.e., medium-fidelity). Eye tracking was incorporated in the evaluation of the site to gain an in-depth understanding of users' visual interaction and to add support to observed findings. In addition, differences in eye-fixation duration on Areas of Interest during optimal task performance and non-optimal task performance were explored.

Keywords: Usability · Eye tracking · Task performance · Fixation duration

1 Introduction

During the spring of 2014, the U.S. Census Bureau's Human Factors and Usability Research Group conducted a medium-fidelity usability evaluation of a web-based prototype for the redesigned American Community Survey (ACS) website. The ACS website is a public site providing information about the American Community Survey, an ongoing national survey that is administered to nearly 3 million households per year and assists in the allocation of more than $450 billion in federal and state funds.

The site also provides supplemental information about ACS data (e.g., data collection, data utilization, survey procedures) and serves as a portal to the American Fact Finder (AFF) for access to ACS data. The U.S. Census Bureau is undergoing a new initiative to change the look and feel of its websites. As a result, various design features have been modified in a prototype for the redesigned ACS website, including its navigational tools and layout, warranting a usability evaluation.

The web-based prototype used during the medium-fidelity evaluation of the ACS site was limited in its functionality and navigational capabilities (e.g., not all links were functional). The web pages were housed on a CD-ROM and up-loaded to the computer for testing. This type of medium-fidelity testing is often a preferred usability method when exploring design alternatives or when requiring guidance on layout and functionality of a site [1]. This is essential in the user-centered design process because feedback regarding the design can be obtained early on. It allows changes to be made during the design phase, which may in turn be less burdensome for developers of the site.

In addition to obtaining accuracy, efficiency, confidence in navigation decisions, and satisfaction data (common usability metrics) [2], eye-tracking data were also collected in the usability evaluation to gain a deeper understanding of participants' visual interaction with the site. The use of eye-tracking technology has emerged in usability testing as a way to inform the design of an interface [3]. Eye tracking is a method used to study eye movements with the assumption that the data provide information on the human-computer interaction. For example, knowing where users are looking can help developers of a site know which features of a site are being noticed and which are being overlooked. It can also be used to understand emotional states and cognitive processes [3].

Data obtained from eye tracking, such as fixation duration captures how long a users' eyes are relatively still as they look at specific content [3]. Research suggests that greater mean eye-fixation duration can be indicative of information complexity and task difficulty [4]. Other research has found that longer fixation duration was associated with difficult words in a reading task and decreased discriminability of a target [5]. The present study seeks to assess whether these findings will hold for data captured in the usability evaluation of a web-based information-rich prototype. Our hypothesis is that, given findings from previous research, there will be longer fixation duration on links and on the ACS main page overall, during non-optimal task performance-in which users are less accurate in completing the task, take longer to complete a given task, and less confident in their navigation decisions- than during optimal task performance (where users are more accurate in task completion take less time to complete a task, and are more confident in navigation). The rationale for this hypothesis is that poor performance on tasks using the web-based prototype is an indicator of cognitive challenges and task difficulty which would warrant additional eye fixations.

2 Methodology

2.1 Participants

Twelve participants (four males, eight females) took part in the medium-fidelity usability evaluation for the ACS website. All participants were from the public and

recruited from local advertisements (newspapers, Craigslist, etc.). The participants selected for the medium-fidelity evaluation self-reported that they had little to no experience with Census Bureau sites (including the ACS website), had at least one year Internet experience, and spent on average about 44.33 h on the Internet per week. The mean age of participants was 40.67 years (range 23–63 years), and 8 out of 12 participants (67 %) had a Bachelors' degree or greater. (See Table 1 for a complete description of participant demographics.)

Table 1. Characteristics of participants

| Total Participants (n)	12
Gender	
Male	4
Female	8
Average Age (standard deviation)	40.67(14.90)
Education	
High School	3
Some College	1
Bachelor's degree	5
Post Bachelor's degree	3
Race	
White	4
Black	8
Hispanic Origin	
Yes	3
No	9

2.2 Procedure

In each individual usability session, the participant entered the testing area and was informed about the purpose of the study and the uses of data that were to be collected. The moderator asked the participant to read and sign a consent form stating that they understood their rights and were voluntarily taking part in the study. The test administrator began video/audio recording after the participant signed the consent form. The participant completed a demographic and computer and Internet experience questionnaire. The participant was positioned in front of a Tobii X120 computer monitor equipped with infra-red cameras. A brief calibration procedure was performed to ensure the quality of eye-tracking data collected.

The test administrator gave the participant randomized task questions and instructed him/her to think aloud while using the American Community Survey site to complete each task. Tasks consisted of common reasons people would visit the site. For example, *You just received the American Community Survey in the mail, and none of your neighbors did. Find out why your address was selected.* (See Appendix A for a complete listing of tasks). The participant was instructed to read each task aloud, and the test administrator loaded the ACS main page on the screen for the task to begin. At the end of each task, the participant was instructed to rate his/her confidence that their link selection(s) led/would lead them to the correct page to complete the task. After completing all tasks, the participant answered a satisfaction questionnaire about their

overall experience with the site. The test administrator asked the participant a set of debriefing questions to obtain more information about his or her experience. The session concluded, and the participant was given a $40 cash honorarium.

2.3   Metrics

Eye-tracking data were analyzed using t-test procedures to determine whether there were significant differences in eye-fixation duration, on optimal links located on the main page or on the ACS main page overall (Fig. 1), for tasks that had a higher percentage of users perform with accuracy, efficiency, and confidence in navigation decisions (optimal task performance) and those that did not (non-optimal task performance).
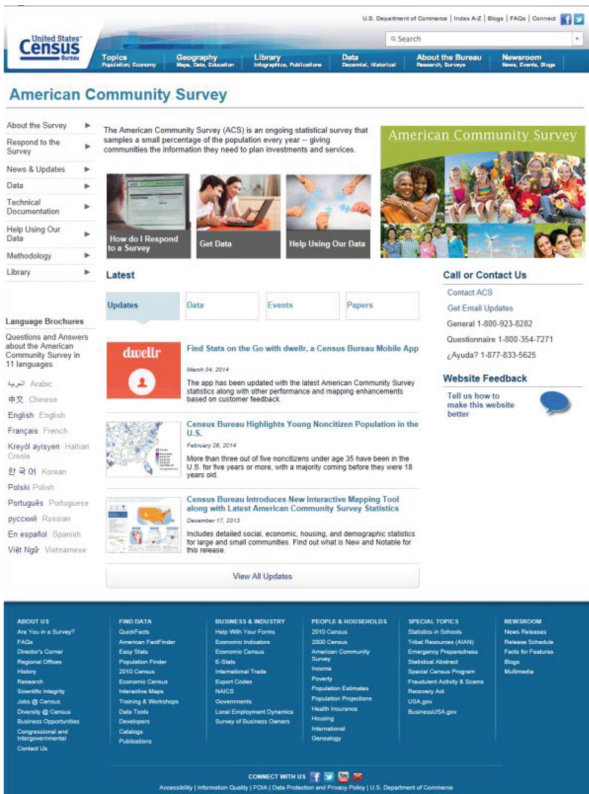
Fig. 1. ACS web-based prototype tested during the medium fidelity usability evaluation (main page)

Fixation Duration. Fixation duration is defined as the amount of time (in seconds) of fixations within predefined areas, also known as Areas of Interest [6]. In this analysis, Areas of Interest consist of optimal links from the main page needed for task

completion and the ACS main page overall. As previous research suggests, users would likely need to fixate longer on these areas during non-optimal task performance because the task is complex and difficult to complete. See Appendix B for highlighted Areas of Interest for the ACS web-based prototype.

Accuracy. Accuracy was based on users' ability to successfully click on the optimal link to successfully complete a given task. A task was coded as having "high accuracy" if 85 % or more of the participants were able to successfully navigate to the optimal link, without assistance. Otherwise, the task was coded as having "low accuracy"

Efficiency. Each task was timed from when the ACS web page prototype loaded until the participant clicked on their last link on the ACS page to complete a given task. Time duration was averaged for each task. A task was coded as a "high efficiency" task if 70 % or more participants were able to complete the task in the average time duration or less. Otherwise, the task was coded as a "low efficiency" task. It is important to note that efficiency is not typically captured in low- and medium-fidelity testing due to the limited functionality of the site, which may not reflect accurate timing of task completion. Therefore, given the range of efficiency scores overall, more leniency was given to what constituted a "high-efficiency" measure than might occur if it were a high-fidelity study.

Confidence. Participants were asked to select links that they felt would direct them to the correct information needed to successfully complete tasks. Following their link(s) selection, they answered the following question to assess the level of confidence they had in their navigation decisions to obtain the information:

On a scale of 1 to 9 where 1 is confident and 9 is not confident, "*How confident are you that you would be able to find the information you were looking for based on your selection?*"

Participants' rating for each task was noted as their level of confidence in their navigation behaviors. Confidence ratings were individually coded such that, those who chose 1 or 2 were confident and those rating a 3 or greater were less confident. Similarly to the accuracy coding, a task was coded as having "high confidence" if 85 % or more participants were confident in their link selection. Otherwise, the task was coded as having "low confidence".

3   Results

Task categorizations are presented below in Table 2. In addition, the number of participants included in the analysis of each task and the average fixation duration for optimal links and the ACS main page (AOI's) is noted.

When examining the average fixation duration on optimal links, there were no significant differences in the fixation duration between tasks that had low accuracy ($M = 1.78, SD = .63$) and high accuracy ($M = 1.48, SD = .64$); $t(7) = 0.72, p = .49$. There were no significant differences in the fixation duration between tasks that had low efficiency ($M = 1.45, SD = .59$) and high efficiency ($M = 2.04, SD = .54$); $t(7) = -1.43, p = .19$. Lastly, there were no significant differences in the fixation

Table 2. Average fixation duration for AOI's and performance categorization by task

[a]Task (n)	Average fixation duration (in seconds) for AOI's		Performance categorization		
	Optimal link(s)	Main page	Accuracy	Efficiency	Confidence
Task 1 (n = 9)	1.25	27.07	Low	Low	Low
Task 2 (n = 10)	2.47	7.43	**High**	**High**	**High**
Task 3 (n = 5)	2.12	14.87	Low	Low	Low
Task 4 (n = 10)	3.08	11.82	Low	**High**	**High**
Task 5 (n = 8)	1.48	9.58	**High**	**High**	**High**
Task 6 (n = 10)	2.12	16.20	Low	Low	Low
Task 7 (n = 10)	0.60	10.31	**High**	Low	**High**
Task 9 (n = 12)	1.79	25.47	**High**	Low	Low
Task 10 (n = 8)	1.17	24.42	Low	Low	Low

[a]Tasks were randomly assigned and not all tasks were given to participants due to time constraints. In addition, task 8 was excluded from the eye tracking analyses as there was no optimal link defined.

duration between tasks that had low confidence (*M = 1.63, SD = .45*) and high confidence (*M = 1.68, SD = .85*); *t(7) = −.11, p = .91.*

When examining the average fixation duration on the ACS main page, there were no significant differences in the fixation duration between tasks that had low accuracy (*M = 18.88, SD = 6.54*) and high accuracy (*M = 13.20, SD = 8.27*); *t(7) = 1.15, p = .28.* However, there were significant differences in the fixation duration between tasks that had low efficiency (*M = 19.72, SD = 6.84*) and high efficiency (*M = 9.61, SD = 2.20*); *t (7) = 2.43, p = .05* and between tasks that had low confidence (*M = 21.60, SD = 5.64*) and high confidence (*M = 9.79, SD = 1.83*); *t(7) = 3.97, p = .01.*

4   Conclusion

While research suggests that increased fixation duration can be indicative of information complexity or task difficulty [4], our results are mixed. Contrary to our hypothesis, we did not find differences in fixation duration on optimal links by task performance. It appears that the links needed to successfully complete a given task captured the same amount of attention from participants regardless of whether the task had a high or low accuracy rate, was high or low in efficiency, or whether there was high or low confidence in navigation decisions.

However, in support of our hypothesis, our results for the fixation duration on the ACS main page showed that increased fixation duration did correspond with task difficulty. Tasks that were more difficult to complete based on non-optimal task performance measures (i.e., low in efficiency, low in confidence) required participants to spend more time looking at content on the ACS main page to find the information they needed. There were significant differences in the fixation duration between tasks that had high efficiency and low efficiency and tasks that were high in confidence and low

in confidence. Tasks for which participants were confident in their navigation decisions and tasks that were higher in efficiency had a shorter fixation duration. For these less challenging tasks, participants were able to quickly sort through the content on the ACS main page and proceed. However, whether or not participants were able to more accurately identify the optimal link needed for successful completion did not vary the amount of time they had to spend looking through the ACS main page content.

Given these mixed results, it may be the case that increased fixation duration is not indicative of only confusion, task difficulty or negative aspects of the user interface design. Perhaps, similar to other studies, fixation duration could be indicative of greater interest and engagement with a target [7]. In addition, perhaps the optimal links themselves may not have been difficult for participants to understand or identify for task completion.

There are various limitations in the present study including, having a small sample size that was based on convenience sampling, using very few tasks in the analyses, and analyzing only fixation duration. Future research should analyze eye-tracking data from a larger sample of participants and perhaps a greater number of tasks or websites. In addition, separate analyses based on participants' individual performance and characteristics should also be considered, as research suggests that there may be differences in eye behaviors based on gender [4] and age [8], to name a few.

Lastly, other types of eye-tracking data can be incorporated in the analyses of eye behaviors to better understand eye patterns as it relates to task performance. For example, gaze time has been shown to be negatively related to task difficulty and saccade rates have been found to decrease when task difficulty or mental load increases [4]. These are two examples of future research with information-rich prototypes that are underway in our lab.

Overall, the present study presents a framework on which future research can build. Our results are preliminary and we hope to elicit future research in exploring the meaning of eye behaviors. Given the emerging uses of eye tracking in usability evaluations, it is important to develop a clearer understanding of what these eye-tracking data mean as it relates to the design of a user-interface and task difficulty.

Appendix A: User Tasks

1. You just received the American Community Survey in the mail, and none of your neighbors did. Find out why your address was selected.

2. Your aunt has been contacted about responding to the American Community Survey (ACS). She would prefer to fill the form out online with your help. Find out if there is an online option.

3. You are interested in seeing what kind of information the American Community Survey has collected. Specifically, does it collect information on the number of people who speak languages other than English in the United States?

4. You are a researcher who will begin using the American Community Survey website on a regular basis, and you want to stay up-to-date on any changes or news items. Is there a way to subscribe to email updates? If so, subscribe to it.

5. Your Vietnamese neighbors received the American Community Survey in the mail and have asked for your help. They want to know if there is any information in their native language. Check the website to see if there is anything available in Vietnamese to help your neighbors understand the basics.

6. You are filling out the survey but feel uneasy about Question 33 - What time did this person usually leave home to go to work? Find out why this question is being asked.

7. You have heard that the American Community Survey has data about your community. Find out how many children are enrolled in elementary education (grades 1–8) in your state.

8. At the request of your Statistics instructor, you are giving a presentation on the American Community Survey. Are there any prepared materials that you can use for this purpose? Based on information on the site, please name at least one example of these materials.

9. Use the site to find:

a. What is the estimated number of grandparents who are caring for their grandchildren in Maryland in 2012?
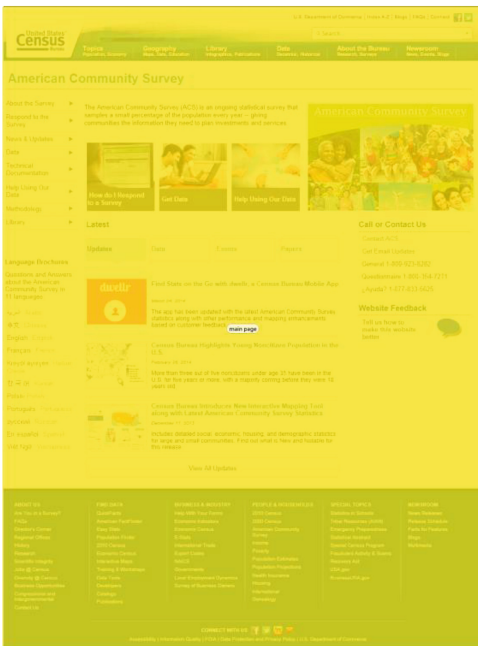
10. You just got mailing materials in the mail inviting you to complete the American Community Survey online, but you aren't sure if this is real or a scam (e.g., is it legitimate?) What would you do?

Appendix B: Defined Areas of Interest on the ACS Web-Based Prototype
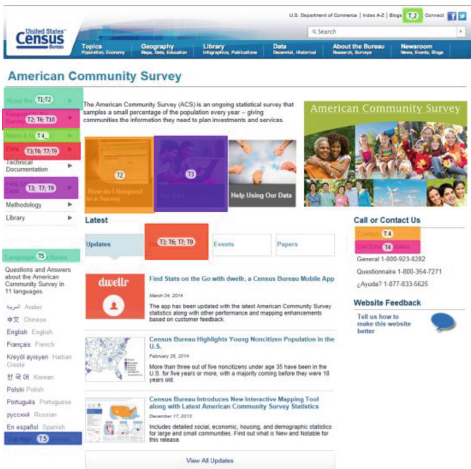
Optimal Links

ACS main page

References

1. Romano-Bergstrom, J., Olmsted-Hawala, E., Chen, J., Murphy, E.: Conducting iterative usability testing on a web site: challenges and benefits. J. Usability Stud. **7**, 9–30 (2011)
2. Frøkjaer, E., Herzum, M., Hornbaek, K.: Measuring usability: are effectiveness, efficiency and satisfaction correlated? In: SIGCHI Conference on Human Factors in Computing Systems, The Haag (2000)
3. Olmsted-Hawala, E., Holland, T., Quach, V.: Usability testing. In: Romano-Bergstrom, J., Schall, A.J. (eds.) Eye Tracking in User Experience Design, pp. 49–80. Morgan Kaufmann, Waltham (2014)
4. Pan, B., Hembrooke, H.A., Gay, G.L., Granka, L.A., Feusner, M.K., Newman, J.K.: The determinants of web page viewing behavior: an eye-tracking study. In: ETRA 2004: Proceedings of the 2004 symposium on Eye tracking research & applications, pp. 147–154 (2004)
5. Guo, K., Mahmoodi, S., Robertson, R.G.: Longer fixation duration while viewing face images. Exp. Brain Res. **17**, 91–98 (2006)
6. Tobii: User Manual: Tobii studio 1.X User Manual (2008)
7. Poole, A., Ball, L.J.: Eye tracking in human-computer interaction and usability research: current status and future prospects. In: Ghaoui, C. (ed.) Encyclopedia of Human Computer Interaction, pp. 211–219. Idea Group, Hershey (2005)
8. Romano-Bergstrom, J., Olmsted-Hawala, E., Jans, M.E.: Age-related differences in eye tracking and usability performance: website usability for older adults. Int. J. Hum.-Comput. Interact. **29**, 541–548 (2013)