Best Practices of Usability Testing Online Questionnaires at the Census Bureau

How rigorous and repeatable testing can improve online questionnaire design

Elizabeth Nichols, Erica Olmsted-Hawala, Temika Holland, and Amy Anderson Riemer U.S. Census Bureau

> QDET2 November 11, 2016 Miami, Florida

This presentation is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.



Concept of Usability Testing

- Usability testing shows us how users perform tasks
 - measures of efficiency, effectiveness, and user satisfaction while accomplishing tasks (ISO Standard 9241-11: 1998)



Example of a Usability Finding





U.S. CENSUS BUREAU census.gov

Overview of Talk

- Top 10 tips
 - How arrived at those practices
 - What still is unknown
- Success example



Tip 1: Participants

10-20 users



- Faulkner (2003) suggests
- 5 users recommended by Virzi (1990, 1992) and Nielsen & Landauer (1993) as described by Fox (2015)
- User groups respondents with particular characteristics



- "Why do you think you qualify for this study?"
- Older respondents
 - Olmsted-Hawala & Holland, 2015; Tullis, 2007; Zaphiris & Savitch, 2008
- Device dependent groups
- Logistical issues
 - Reminder calls and emails



Household





Tip 2: Iterative testing Same survey – multiple rounds

Round 1 of testing	Round 2 of testing	
2015 Census Test 🛛 🚍	2015 Census Test ≡	
Next, we need to record each person's relationship to Jane A Doe. James C Doe is Jane A Doe's (Help)	Next, we need to record each person's relationship to Joe Doe. Jill Doe is Joe Doe's (Help)	
 Opposite-sex husband/wife/spouse Opposite-sex unmarried partner Same-sex husband/wife/spouse Same-sex unmarried partner Biological son or daughter Adopted son or daughter Stepson or stepdaughter Brother or sister Father or mother Grandchild Parent-in-law Son-in-law or daughter-in-law 	 Opposite-sex husband/wife/spouse Opposite-sex unmarried partner Same-sex husband/wife/spouse Same-sex unmarried partner Biological son or daughter Adopted son or daughter 	More spacing betwee respons options



Iterative testing Similar survey, across field periods



Field period 2015

	Ple	ase Log In
Example Access Code	► ####################################	Please enter the 14-digit User ID found below the barcode on the materials we mailed to you User ID: 91313 11211 8674 Login If you do not have a User ID, <u>click here</u>



Iterative testing Similar survey, across field periods



Field period 2016

Census FAQ Instructions Burden	Select Language +		
Welcome to the 2016 Census Test. You will need the materials we mailed to you in order to start. All the information that you provide will remain confidential.	Please Log In Please enter the 14-digit User ID found below the barcode on the materials we mailed to you. User ID: 00000 - 00000 - 0703 Login If you do not have a User ID,click here.		
Terms & Conditions You are accessing a United States Government computer network. Any information you enter into this system is confidential and may be used by the Census Bureau for statistical purposes, as well as for other uses, such as improving the efficiency of our programs. If you want to know more about the use of this system, and how your privacy is protected, visit our online privacy webpage al http://www.census.gov/privacy/privacy_policy/. Use of this system indicates your consent to us collecting, monitoring, recording, and using the information that you provide. So that our website remains accurate and available to you and all other visitors, we monitor network traffic to identify unauthorized attempts to upload or change information or otherwise cause damage to the web service. Your usage of this system is likely to be monitored, recorded, and subject to audit. If you are not using the network connection for authorized purposes, then it is a violation of Federal law and can be pursished with fines or imprisonment (PUBLIC LAW 99-474). Build: ef21ffia. L.OMB.No : 060720889. L.Amproval Expires: 5/312015			



Tip 3: Set up the task

- The survey task
 - "Answer the survey questions as they apply to you in real life."
- Dummy data



- Mocked-up mailing materials
- Pre-filled data
- Example instruction: "If you were to receive the survey at your home, the mailing materials you would get would have your name. Since we cannot replicate that for the lab setting, you will have to pretend that this letter came to your address and that is your address. That is the only part of the study that is pretend."



Tip 4: Probing

Think-aloud



- Think –aloud vs. silence
- 35 question problems vs 9 over 2 rounds of testing Nichols 2016
- Does not affect eye-tracking data with exception of older users

Romano Bergstrom and Olmsted-Hawala,2012

 Concurrent probing or verbal probing as described by Willis (2004; 2015)



Tip 5: Think-aloud



- "How many windows are in your home?"
 - Dillman, 2007; Ericsson and Simon, 1993
- Participants who fall silent
 - "Keep talking" (saying that very quietly)
 - backchannel continuers such as "mm-hm" or "uhhuh"
 - Boren & Ramey, 2000; Ericsson & Simon, 1993; Fox, Ericsson & Best, 2011; Hertzum, Hansen & Anderson, 2009; Nørgaard & Hornbæk, 2006
- Participants who get "stuck" on an issue
 - "We will be sure to tell the developers" or "Thank you, I've made a note of the problem"



Tip 6: Eye-tracking



If there is time to analyze. Much is still unknown.







Tip 7: Measuring Satisfaction Much is still unknown.

- Questionnaire for User Interaction Satisfaction (QUIS) (Chin, Diehl, & Norman, 1988)
 - Use a modification of these items
- System Usability Scale (SUS) (Brooke, 1996)
 - 10-items better for website then web survey
- Still in need of a better survey satisfaction with taking a survey is difficult to measure
- How to present the data



Example of satisfaction scores across 2 rounds

Forward Navigation









Tip 8: Vignettes

- One way to measure accuracy
- Test rare events
 - For example, a break-off and a resume
 - An edit message





Tip 9: Retrospective Debriefing Much is still unknown.

- PowerPoint slide deck to jog memory
- Beware of memory error and retrofitting the problem to something unrelated
- Useful to compare minor changes to design or a before and after





Tip 10: Use pictures to ✓••••• communicate usability issues

- Review video and notes
- Screen shots and PowerPoint with issues and recommendations
- Shortly after sessions end
- Written report can follow



Success Example Return on Investment

American Community Survey 2011-2016



2011 – Usability testing

Instructions FAQs	Save & Logout
	Pagia Info
The following series of questions refer to John T Doe.	Basic Into
	Person Info
	· John T Doe
Where was John T Doe born? (Help)	Jane N Doe
In the United States - Select name of state. Select Name C Outside the United States - Enter name of foreign country, or Puerto Rico, Guam, etc. << Previous Next	• Tommy C Doe
4.0 .35	Cont

2011 ACS Internet Test

- Paradata
 - This question triggered the most alerts of any screen.
 - ~ 10 % of the people visiting this screen receiving an alert

(Horwitz, Tancreto, Zelenak, and Davis, 2013).



2013– Usability testing, version 1

Census Bureau Ame	rican Community Survey
Instructions FAQs	Save & Logout
	Where You Are
The following cories of questions refer to Sarah Joseph	Basic Info
The following series of questions refer to sarah Joseph.	Housing Questions
	Person Info
Where was Sarah Joseph horn? (Help)	 Alice Joseph
	Sarah Joseph
In the United States - Select name of state	Karen Brown
Coutside the United States - Enter name of foreign country, or Puerto Rico, G	uam, etc.
	Contact
	Accessibility Privacy Secu



2013 Usability testing, version 2

Census Bureau Americ	an Community Survey
Instructions FAQS	Save & Logout
	Pasia Infe
The following series of questions refer to Matthew Joseph.	Basic Into
	Housing Questions
	- Sarah Jasanh
Where was Matthew Joseph born? (Help)	Sarah Joseph
	Alice Joseph Matthew Joseph
Select Name COutside the United States - Enter name of foreign country, or Puerto Rico, Guam, e	etc.
	Contact Us



2014 ACS field test

- Paradata for split-panel test
 - Alerts in new design < Alerts in former design</p>

(Zelenak, 2016)



2016 ACS

Census Bureau American Com	munity Survey
Instructions FAQs Save and Log	Out
 The following series of questions refer to John Doe. Where was John Doe born? (Help) In the United States - Select name of state. Select Name Outside the United States - Enter name of foreign country, or Puerto Rico, Guam, etc. Previous Next > 	Where You Are Basic Info Housing Questions Person Info • Jane Doe • John Doe • Jilly Doe
Contact Us	Accessibility Privacy Security



Questions

Elizabeth Nichols

Elizabeth.May.Nichols@census.gov

Erica Olmsted-Hawala

Erica.L.Olmsted.Hawala@census.gov

Temika Holland

Temika.Holland@census.gov

Amy Anderson Riemer

Amy.E.Anderson.Riemer@census.gov



References

- Ashenfelter, K. T., Holland, T., Quach, V., Nichols, E., Lakhe, S. (2012). ACS Internet 2011 Project: Report for Rounds 1 and 2 of ACS Wireframe Usability Testing and Round 1 of ACS Internet Experiment Mailing Materials Cognitive Testing (Survey Methodology #2012-01) http://www.census.gov/srd/papers/pdf/ssm2012-01.pdf
- Boren, T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. IEEE Transactions on Professional Communication, 43(3), 261-278.
- Brooke, J. (1996). SUS A quick and dirty usability scale. Usability Evaluation in Industry. In: P.W. Jordan, B., Thomas, I. McClelland and B.A. Weerdmeester (Eds.), London: Taylor & Francis. Available online at: <u>http://cui.unige.ch/isi/icle-wiki/_media/ipm:test-suschapt.pdf</u>
- Chin, J. P., Diehl, V.A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. Proceedings of CHI '88: HumanFactors in Computing Systems (pp. 213-218). New York: ACM.
- Dillman, D. A. (2007). Mail and internet surveys: The tailored design method (2nd ed.). New York: John Wiley & Sons.
- Ericsson, K. A., & Simon, H. A. (1993). Protocol Analysis: Using Verbal Reports as Data. Cambridge, MA: The MIT Press.
- Faulkner, L. (2003). "Beyond the five-user assumption: Benefits of increased sample sizes in usability testing." Behavior Research Methods, Instruments, & Computers (2003) 35: 379. doi:10.3758/BF03195514
- Fox, J. (2001). Usability Methods For Designing A Computer-Assisted Data Collection Instrument for the CPI. Proceedings of the 2001 Federal Committee on Statistical Methodology (FCSM) Research Conference <u>http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.159.4335&rep=rep1&type=pdf</u> Last accessed 8/30/2016.
- Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. Psychological Bulletin, 137, 316-344.
- Hertzum, M, Hansen, K., and Anderson, H. (2009). Scrutinising usability evaluation: does thinking aloud affect behaviour and mental workload? Behaviour & Information Technology 28 (2), 165-181.
- Horwitz, R., Tancreto, J.G., Zelenak, M.F., Davis, M.C. (2013). Use of Paradata to Assess the Quality and Functionality of the American Community Survey Internet Instrument, Last accessed 8/16/2016 at https://www.census.gov/library/working-papers/2013/acs/2013_Horwitz_01.html
- International Standards of Organization (ISO). (1998). ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) Part 11: Guidance on usability. Last accessed 9/7/2016 at http://www.iso.org/iso/catalogue_detail.htm?csnumber=16883
- Nichols, E. (2016) Cognitive Probing Methods in Usability Testing Pros and Cons Presentation at American Association of Public Opinion Research (AAPOR) annual conference in Austin, Texas, May 15, 2016.
- Nørgaard, M., & Hornbæk, K. (2006). What do usability evaluators do in practice? An explorative study of think-aloud testing. In Proceedings of the 6th Conference on Designing Interactive Systems(pp. 209-218).
- Olmsted-Hawala, E., & Holland, T. (2015). Age-related differences in a usability study measuring accuracy, efficiency, and user satisfaction in using smartphones for Census enumeration: Fiction or Reality? In Proceedings from HCII 2015, Lecture Notes in Computer Science, Human Aspects of IT for the Aged Population, Springer.
- Romano Bergstrom, J., and Olmsted-Hawala, E. (2012). Effects of age and think aloud protocol on eye-tracking data and usability measures, Presentation at International conference EyeTrackUX, Las Vegas, NV, June 2012.
- Tullis, T. (2007). Older adults and the web: Lessons learned from eye-tracking. Universal Access in Human Computer Interaction: Coping with diversity, Volume 4554 of the series Lecture Notes in Computer Science pp 1030-1039
- Willis, G. B. (2015). Analysis of the cognitive interview in questionnaire design. Oxford University Press.
- Willis, G.B., (2004). Cognitive Interviewing Revisited: A Useful Technique, in Theory? Methods for testing and evaluating survey questionnaires. In S. Presser et al. New Jersey: John Wiley & Sons, Inc. pp 149-172.
- Zaphiris, P and Savitch, N. (2008). Age related differences in browsing the Web. In SPARC Workshop on "Promoting independence through new technology", Reading England UK. Available online at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.167.7963&rep=rep1&type=pdf
- Zelenak, M.F., (2016) 2014 American Community Survey Internet Test. U.S. Census Bureau: Decennial Statistical Studies Division 2016 American Community Survey Memorandum Series. #ACS16-MP-05.

