A comparison of remote asynchronous usability testing and in-house lab testing of a Census Bureau website

Elizabeth Nichols, Erica Olmsted-Hawala, and Marylisa Gareau Center for Survey Measurement, U.S. Census Bureau

March 20, 2014 – 2014 FedCASIC



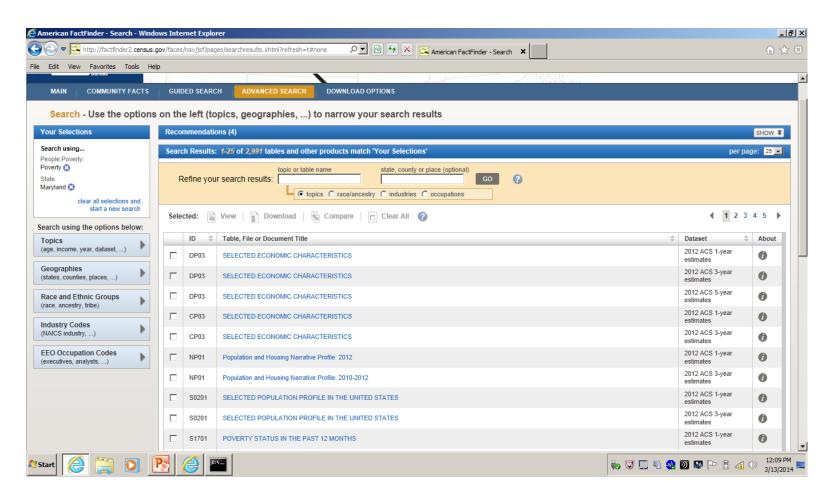
Question:

Does remote asynchronous usability testing provide the same results as in-house lab testing for a data dissemination website?

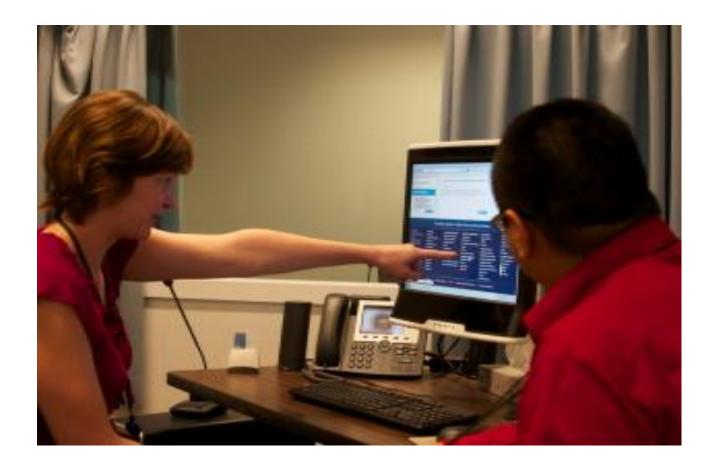
Why does this question matter?

- -Recruit a variety of users to get different experiences
- -Own devices
- -Own time

U.S. Census Bureau data dissemination website: American FactFinder

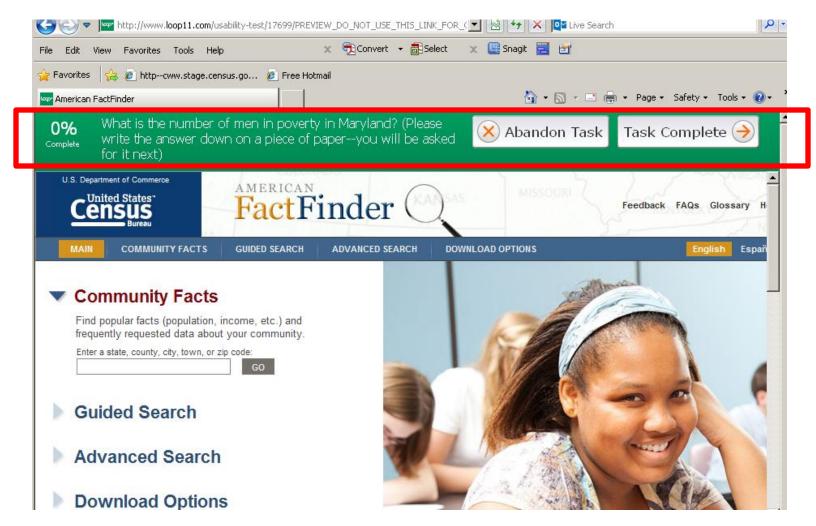


In-house lab testing at the Census Bureau





Loop11: Remote testing product





Methods: Participants

- In-house lab
 - 10 novice and 4 experienced American FactFinder users
 - Recruited from metro DC area
 - Database
 - Word of mouth, flyers
 - Newspaper, craigslist ads
 - Participants were given \$40 for their travel expenses / time
- Remote
 - 120 American FactFinder users
 - Census Information Centers and State Data Centers (experienced users)
 - 2 email notifications were sent
 - No monetary compensation

Methods: Timing

- In-house lab
 - 14 participants: December 10, 2012 January 14, 2013
 - Analyzing and writing report: December 19, 2012 through February 2013
- Remote
 - 24 participants: September 9 27, 2013
 - Webinar to client: November 13, 2013

Methods: Tasks and Measures

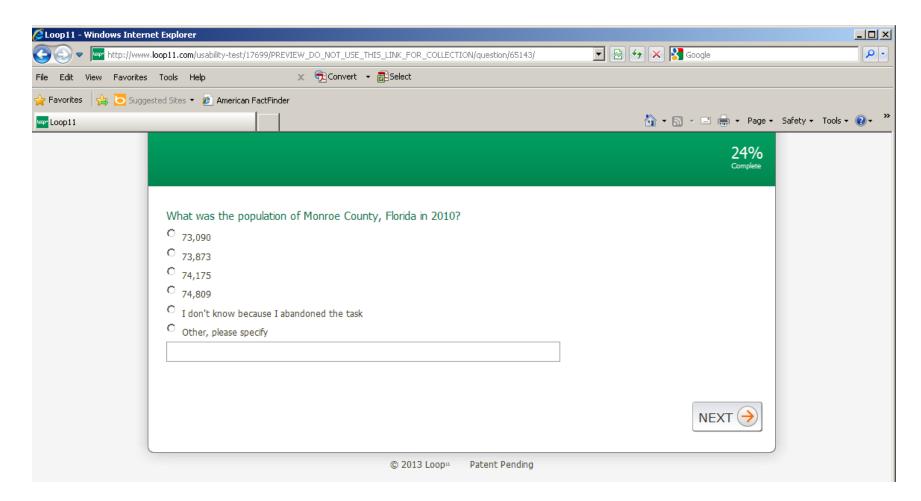
In-house lab

- 8 tasks
- Think-aloud
- Accuracy
- Time-on-task
- Eye-tracking data collected
- Satisfaction questions for each task
- Debriefing satisfaction questions at end

Remote

- 4 tasks (identical to 4 used in in-house)
- Accuracy
- Time-on-task
- Satisfaction questions for each task
- Debriefing satisfaction questions at end

Accuracy: Follow-up question



Task 1: What is the number of men in poverty in Maryland?

Method	Accuracy	Notes
Remote: By post-task question	69% (18/26)	2011 answers: 1 year=247,081 (12) 5 year=216,558 (4) 2012 answers: 3 year=252,726 (2 write-ins)
In-house: Test administrator marked correct or incorrect	40% (6/14)	Included experts and novice users

Task 2: What is the population of Monroe County Florida in 2010?

Method	Accuracy	Notes
Remote: By post-task question	96% (23/24)	Decennial Estimate (22) ACS Estimate (1)
In-house: Test administrator marked correct or incorrect	100% (14/14)	Answer from Community Facts, novice and expert users

Task 3: On average, how old are the people in Monroe County, Florida?

Method	Accuracy	Notes
Remote: By post-task question	88% (20/24)	2010 demographic profile from Census (18) 2012 Pop estimates (2) 2012 ACS (1) 3 people abandoned task
In-house: Test administrator marked correct or incorrect	86% (7/8)	Most from Community Facts number

Task 4: When will the population estimates in Montgomery County be released?

Method	Accuracy	Notes
Remote: By post-task question	50% (12/24)	3 said July 1 (we counted as a failure, was a misreading of table)
In-house: Test administrator marked correct or incorrect	50% (7/14)	

Satisfaction scores of AFF Ease of use*

	Task 1	Task 2	Task 3	Task 4
Remote	5.5 (26)	6.4	6.2	5.3
In-house	4.2 (4)	4 (4)	5.9 (10 novice)	5.3 (4)

Scale: 7-Very easy, 1-Very difficult

* There are some differences between the in-house and remote ease of use question

Satisfaction scores of AFF Frustrating to use*

	Task 1	Task 2	Task 3	Task 4
Remote	3.1 (26 participants)	1.8 (24)	2.4 (24)	3.4 (24)
In-house	3.5 (4)	4 (4)	1.3 (10 novice participants)	3.3 (4)

Scale: 7-Extremely frustrating, 1-Not frustrating at all

* There are some differences between the in-house and remote frustrating to use question

Time-on-Task - Efficiency

	Task 1	Task 2	Task 3	Task 4
Remote*	8 min. (27 participants)	2 min. 40 sec. (24)	3 min. 30 sec. (24)	3 min. 30 sec. (24)
In-house	6 min. 46 sec. (14)	2 min. 32 sec. (14)	1 min. 54 sec. (10)	2 min. 33 sec. (14)

^{*} Includes success, failures, abandonments

Staff time

	OMB letter	Set up test (a)	Collect data	Analyze
Remote	2 hours	5 hours	0	4 hours
In-house	2 hours	16 hours	2 hours per participant	30 hours

(a) For In-house: includes creating protocol, recruiting participants, photocopying materials & dry run.

For remote: includes entering tasks, success URLs and satisfaction questions into application.

Collected in-house: Eye-tracking





Collected in remote: Pages visited

	Task 1	Task 2	Task 3	Task 4
Average # of pages visited*	9.1	3.7	5	4.4
Min/Max pages	(1/20)	(3/6)	(2/24)	(1/14)

^{*} Includes success, failures, abandonments

Using remote usability testing

	In-House	Remote
High level problems were identified	√	√
Accuracy	√	√
Time-on-task	✓	✓
Satisfaction	✓	✓
Detailed problems were identified	√	

When to use remote testing?

- In-house should come first
- Quick turn around
- Need performance measures to compare against
- Design comparisons
- Iterative testing

Questions?

Beth Nichols
Elizabeth.May.Nichols@census.gov

