Joint Cognitive and Usability Testing of Translated Instruments: Identification of Translation Problems through eye tracking

Patricia Goerman, Mikelyn Meyers, Sabin Lakhe, and Kathleen Kephart, U.S. Census Bureau

Presented at the Thirteenth International Workshop on Comparative Survey Design and Implementation
City University, London
March 26-28, 2014

Introduction

- More U.S. Census Bureau internet surveys in Spanish
  - American Community Survey, 2013
  - Decennial Census, testing in preparation for 2020
- New testing opportunities
  - Joint usability and cognitive testing
  - Comparison of eye tracking across languages

Research Questions

- Is it possible to identify translation issues through comparison of respondent eye tracking data across languages?
- Do problems identified through cognitive interview probes also appear in respondents' eye movements?

Review of the Literature

- Eye tracking, basic definitions
  - Fixation counts
  - Fixation duration
  - Saccades
  - Regression
- Sources: Olmsted-Hawala, et al. 2011; Galesic et al., 2008; Bax, 2013, Rayner, 1998

Example issues studied through eye tracking

- Optimal placement of instructions
- Optimal location of answer boxes
- Timing to respond to questions for self v. proxy
- Testing in different modes, including paper
- Joint cog and usability testing
- Sources: Kunz and Fuchs, 2012; Lenzner, et al. 2014; Olmsted-Hawala, et al. 2011; Walton et al. 2014; Bergstrom, et al. 2014

Eye tracking and language research

- Cognitive processes during translation
  - Jensen, 2008; Doherty et al. 2010
- Language affects gaze and fixation time
  - Rayner, 1998
- Research out of Spain: Areas of interest in web design; gaze falling longer on incorrect text
  - Lopez-Gil, et al. 2010; Marcos and Rello, 2013
- Malaysian students taking Eng proficiency test
  - Bax, 2013

Evaluation of questionnaire translation through eye tracking

- Lack of literature on the topic
- Relevant literature: eye tracking to identify indications of difficulty (Bax, 2013)
  - Eye fixations typically 200-250ms
  - Mean saccade size 7-9 letter spaces
  - Rightward saccades, moving forward
  - Regressions
  - Return sweeps v. backtracking

Two Census Bureau studies

- 2014 Census test internet instrument
  - Testing on desktop and laptop computers
  - 11 English speakers
  - 10 Spanish speakers
- 2015 Census test internet instrument
  - Testing on respondents' smartphones/tablets
  - 30 English speakers (15 each in 2 mini-rounds)
  - 16 Spanish speakers (10 & 6 in 2 mini-rounds)
- Tobii eye tracking, X120, T120, X2-60

Plans

- Exploratory research
- Compare:
  - Usability issues seen through observation
  - Cognitive issues found through probing
  - Eye tracking fixations

Study 1 (2014)

- Eye tracking issues
  - Desktop in usability lab (English)
  - Laptops remote testing with (Spanish)
  - % accuracy eye tracking (ideal about 75%)
    - English n=9 (range: 59%-88%; average: 75.3%)
    - Spanish n=9 (range: 18%-77.5% average : 51.2%)

Study 2 (2015)

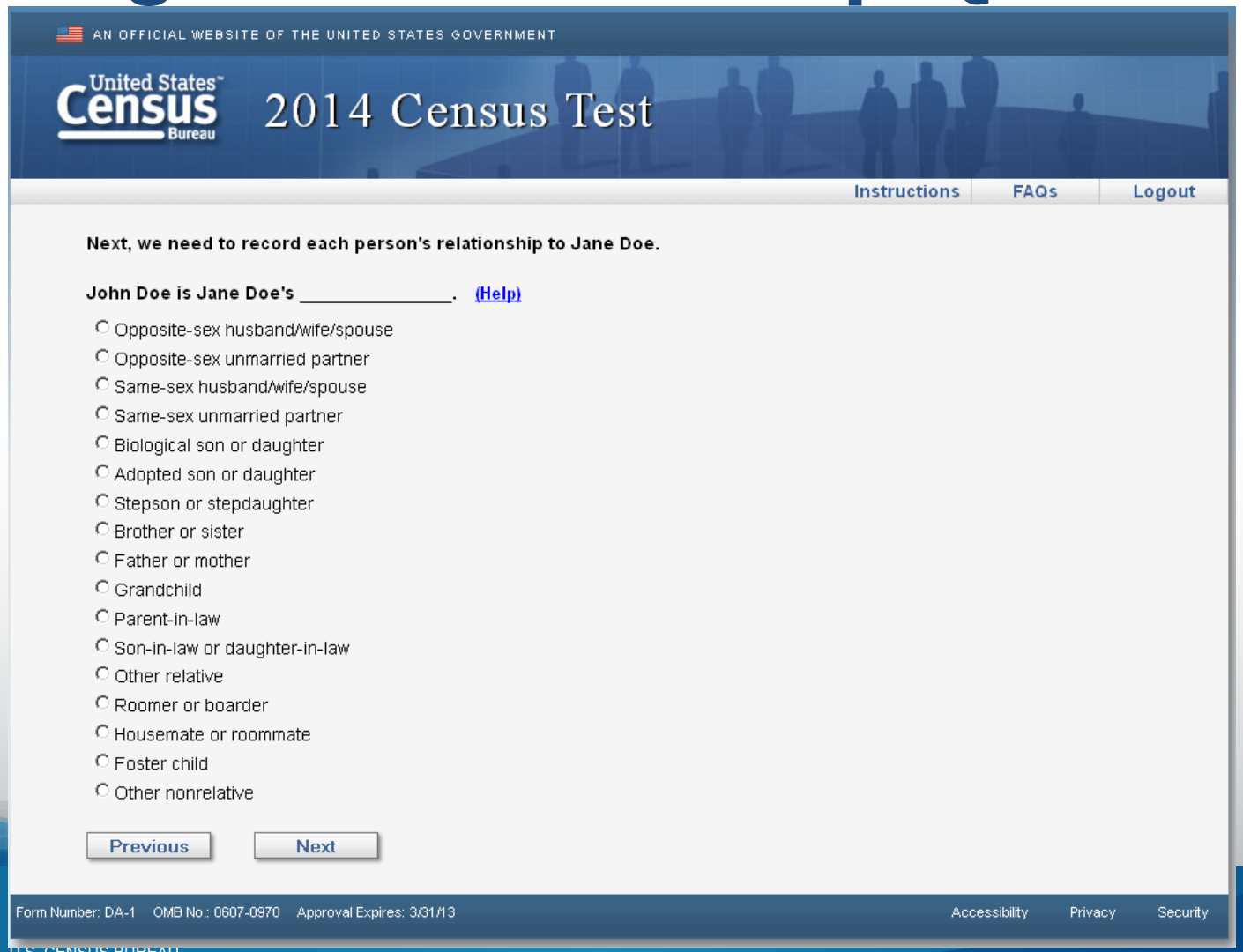
- Eye tracking issues
  - Mobile stand in usability lab and field (English)
  - Mobile stand in field (Spanish)
  - % accuracy eye tracking (ideal 75%, can be lower)
    - English n=15 (range: 0%-66%; average: 24.7%) (mini-round 2)
    - Spanish n=10 (range: 0%-35.5%; average: 10.7%) (mini round 1)

Eye tracking mobile stand

Steps taken to correct issues

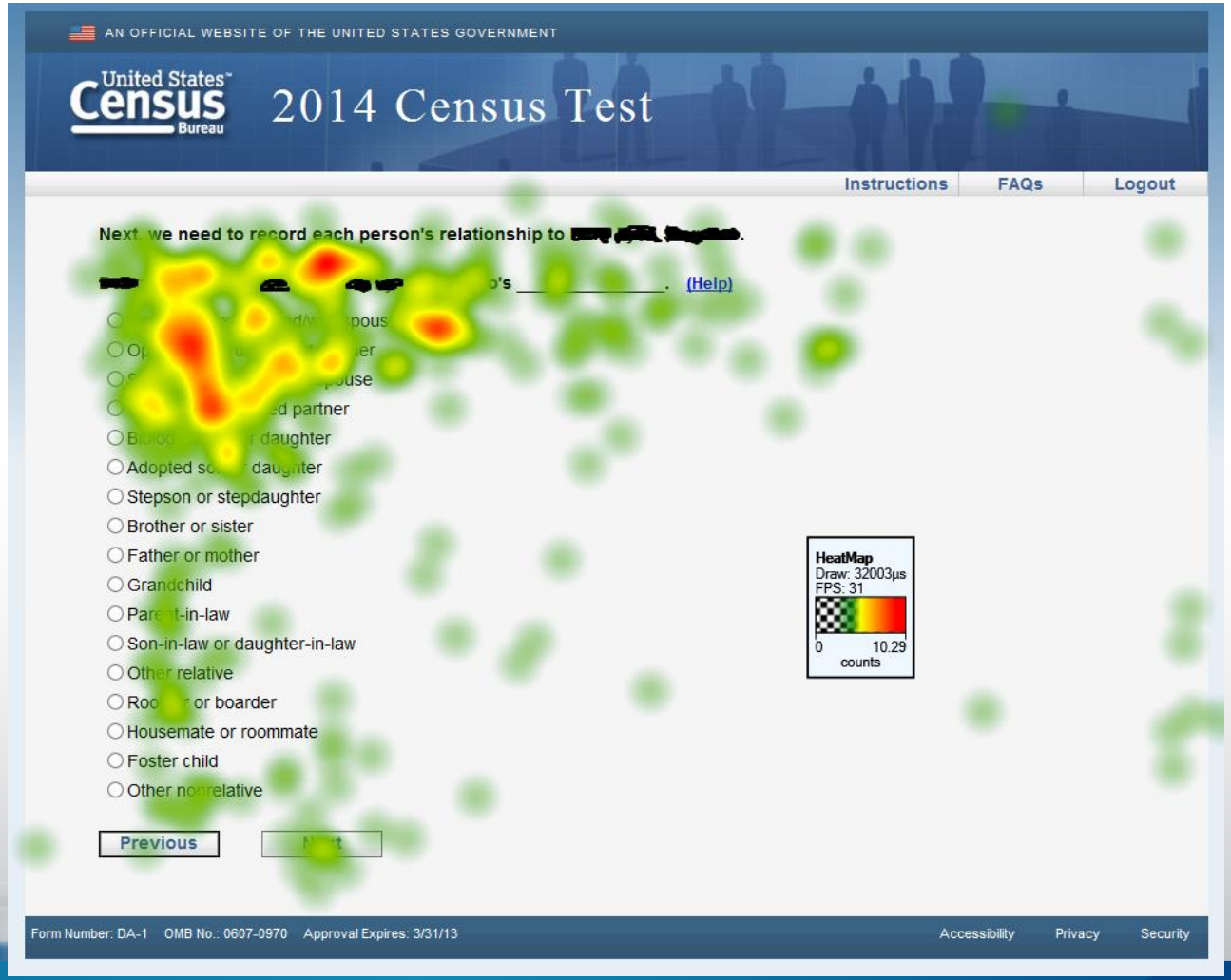
- Met with Tobii representatives
- Practiced eye tracking set up, calibration
- Eye tracking success still extremely poor
- Not enough data to compare across languages

Preliminary findings 2014: English relationship Q. v. 1



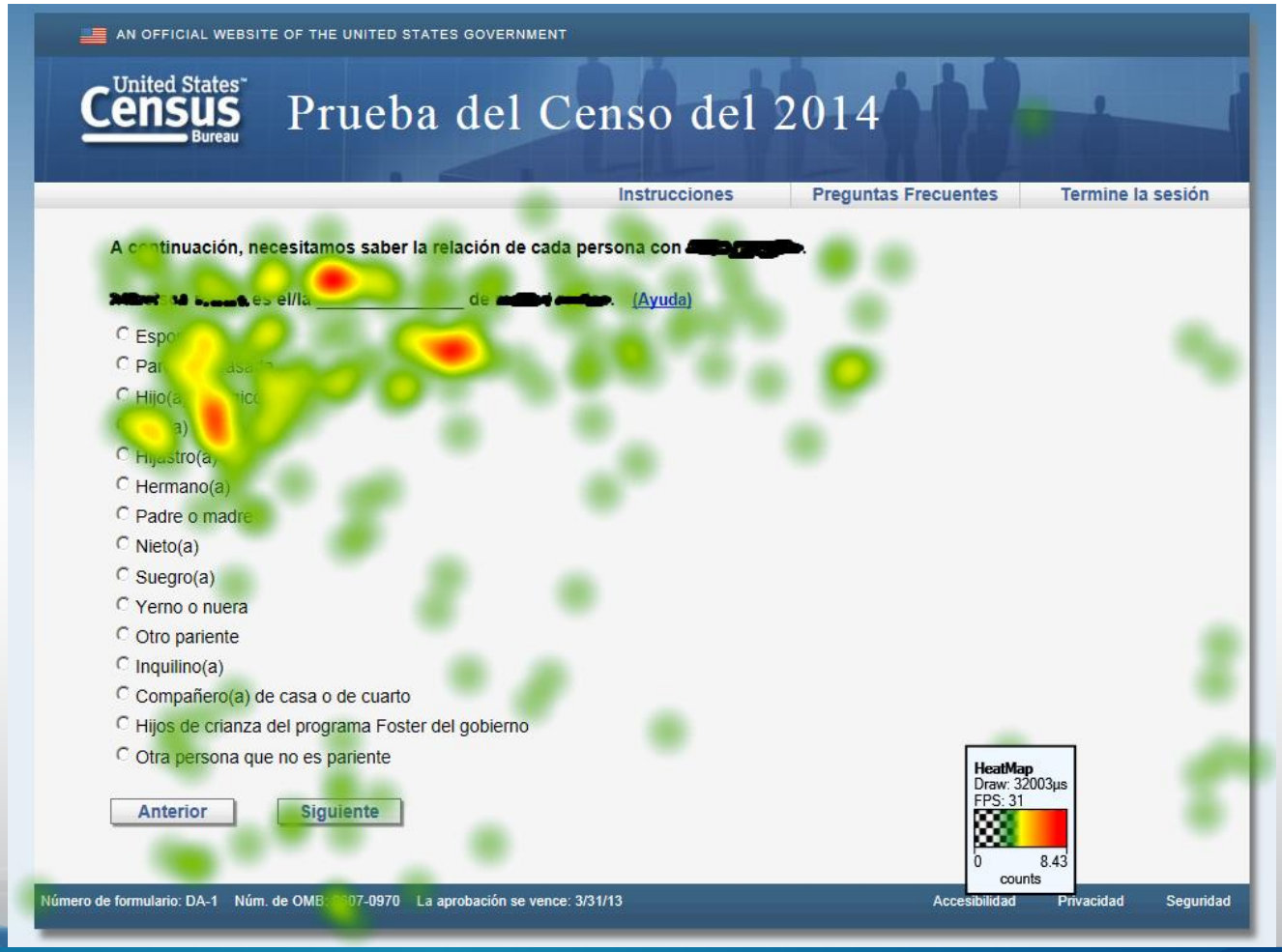
English eye tracking heat map v. 1

- Combined heat map for 6 English speakers

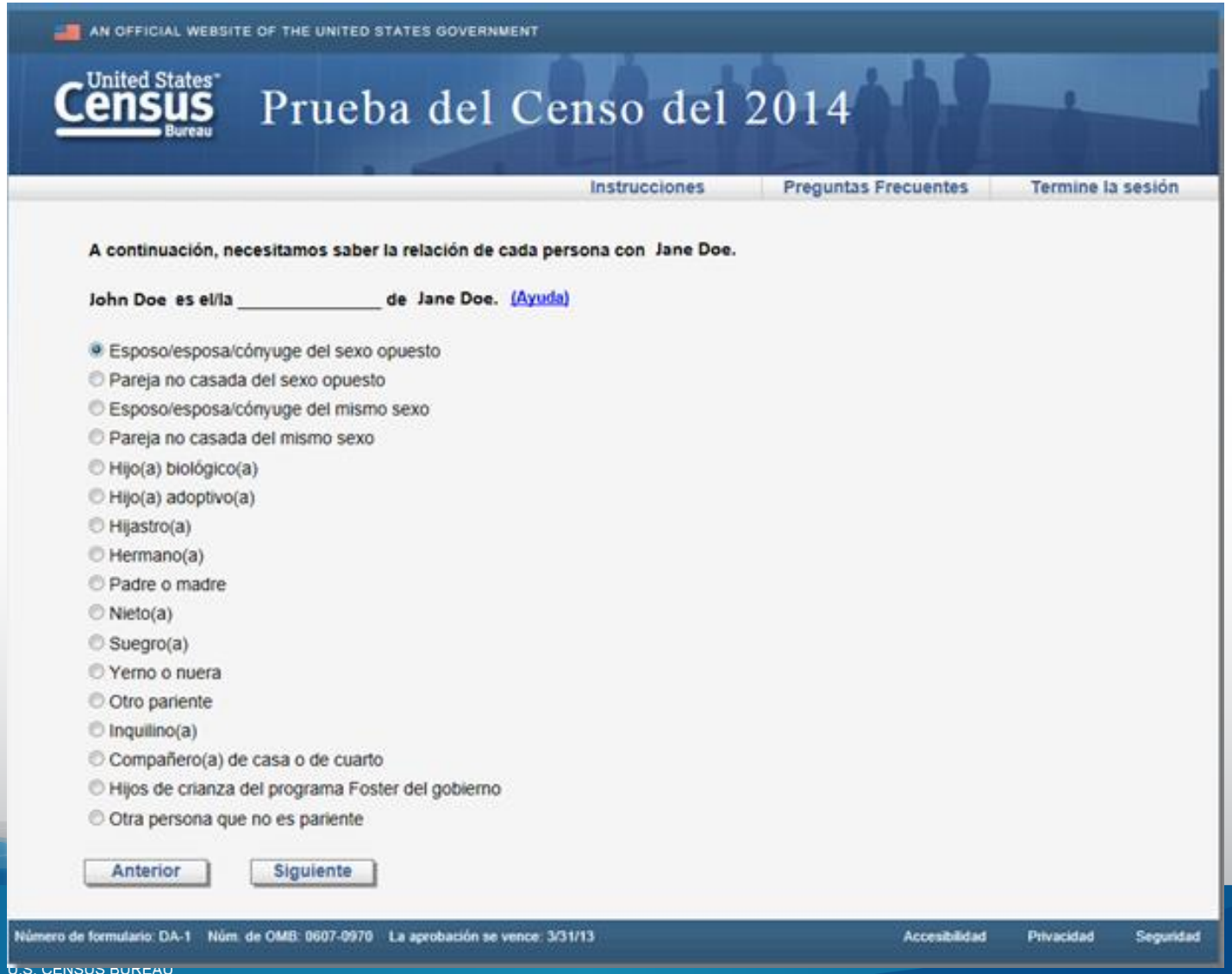


Spanish eye tracking heat map v. 1

- Combined heat map for 2 Spanish speakers

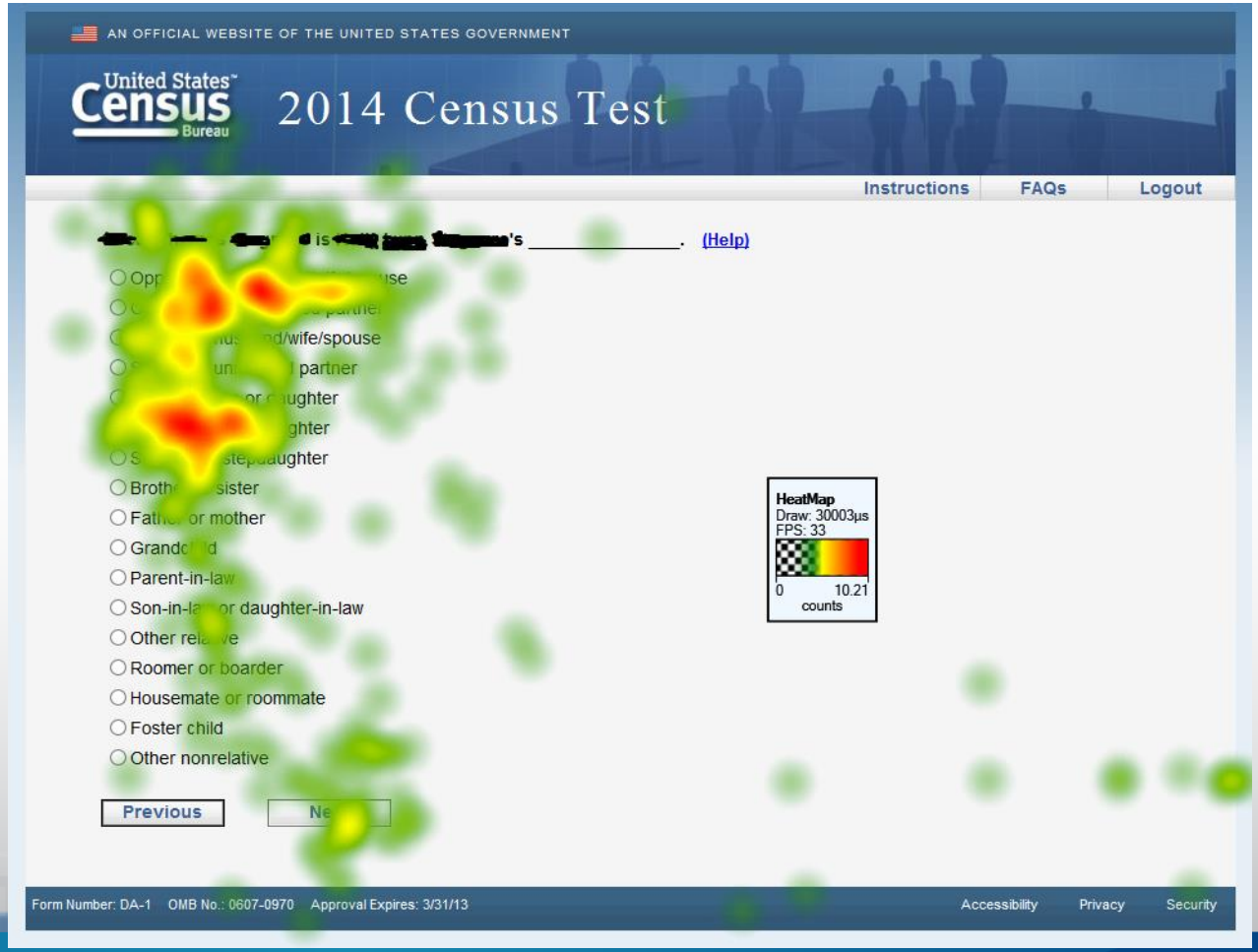


Spanish relationship Q. v. 1



English relationship Q v.2

- Combined heat map for 8 English speakers

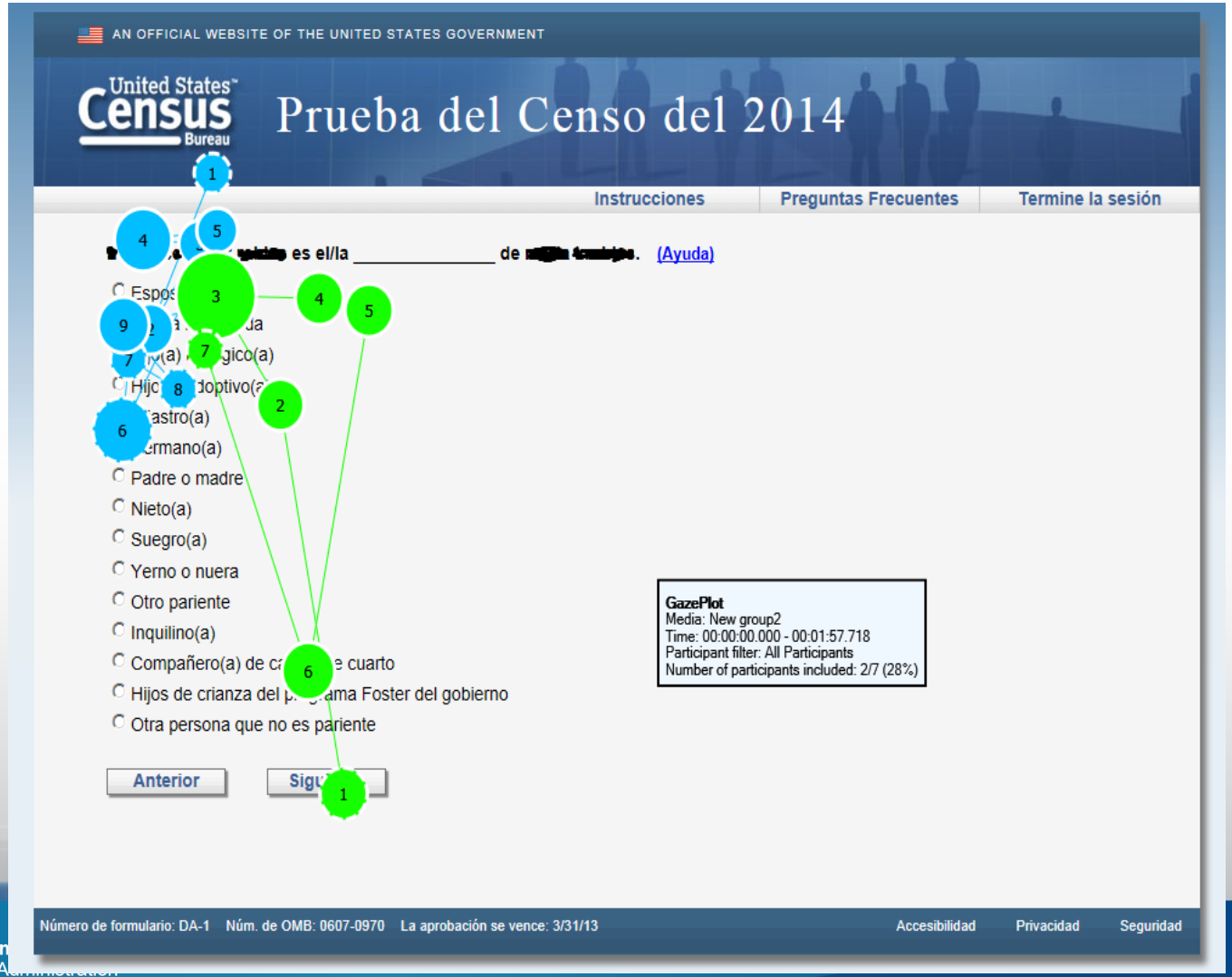


Spanish relationship Q v. 2

- Combined heat map for 2 Spanish speakers

Spanish gaze plot

- 2 Spanish speakers



Future Plans

- Investigate different types of eye tracking equipment in our lab
- Wait for future data to be gathered, possibly next project that does not involve mobile devices
- More research on best ways to measure differences across language
- Controlled experiment

Joint Cognitive and Usability Testing of Translated Instruments: Identification of Translation Problems through eye tracking

Patricia Goerman, Mikelyn Meyers, Sabin Lakhe, and Kathleen Kephart, U.S. Census Bureau

For more information:
E-mail: Patricia.L.Goerman@census.gov

Disclaimer:

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