Compendium of Current Pretesting Methods

Emily Geisen, Joe Murphy
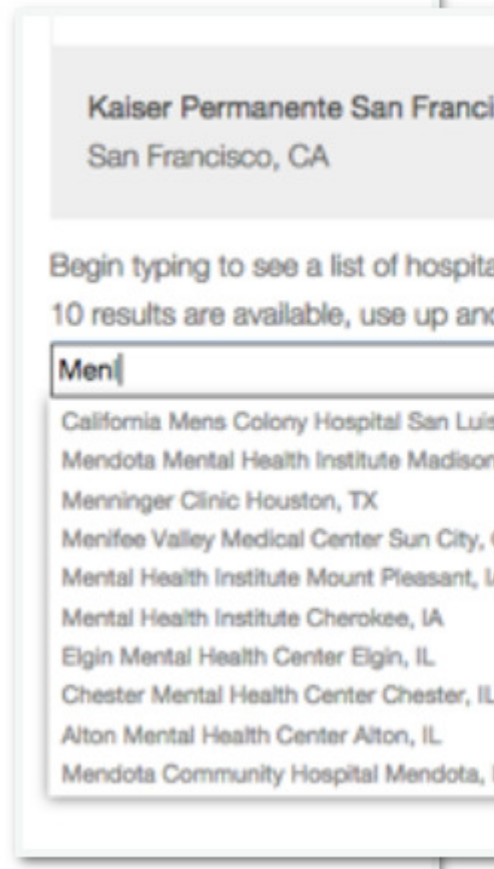
Introduction

- Technology is changing interviewers interact with
  - Choice of device
  - Survey fun

Introduction

- **Technology is changing how respondents and interviewers interact with surveys**
  - Choice of device
  - Survey functionality
  - Pretesting methodologies

- **What's the impact on survey quality?**
  - Some known effects
    - Longer completion times, nonresponse (Antoun, 2015)
  - Many unknown effects or unknown reasons for known effects

Purpose

- Compare traditional pretesting methods with emerging methods for evaluating "modern" surveys

  - Traditional methods (expert review, cognitive interviewing, pilot testing)
    - Limitations
    - Effect of technology on these methods

  - Emerging methods (usability testing, eye tracking, crowdsourcing)
    - Overview
    - Advantages
    - Examples

Traditional Pretesting Methods

Expert Review

- Content evaluation by substantive expert or survey methodologist
  - Subjective based on reviewer's knowledge experience
  - Standardized guidelines
    - QAS focuses mainly on cognitive process, not on visual design, navigation, etc
    - Web survey guidelines (Dillman, Smyth, & Christian, 2009; Couper, 2008): Constantly changing – some become obsolete, some not addressed

- Researchers must design aspects of questionnaires where no guidelines exist.

Cognitive Interviewing

- Identifies potential problems in surveys by evaluating the cognitive processes used to answer questions.
  - If testing not done on programmed instrument, does not evaluate visual design or respondent-survey interaction
  - Primarily focuses on cognitive processes

- Technology's impact on cognitive testing
  - Recruitment strategies (newspaper advertisements vs craigslist and Facebook)
  - Remote and virtual cognitive testing

Pilot Testing

- A "dry-run" to test out the procedures that are being considered for the full study

  - Paradata: time spent per question, error/warning messages activated, number of times back button was pushed, proportion of mobile respondents
  - Item missing rates, question distributions

- Otherwise, limited for determining issues with respondent-survey interaction

  - Does not identify all questions that are not performing well
  - For questions with poor performance indicators, don't know what the problem is (usability, comprehension)

Emerging Pretesting Methods

Usability Testing (UT) Overview

- Watching a user (respondent or interviewer) as they complete a task or goal
  - Obtain observational data, self-report, implicit data
  - Evaluate accuracy, efficiency and satisfaction
  - Revise and test again (iterative, use to guide design)

- Used in late 1990s as use of CAI increased, mainly for interviewer-administered (IA) surveys (Hansen, Fuchs, & Couper, 1997; Marquis, Nichols, & Tedesco 1998; Couper, 2000)

- Particularly useful for self-administered (SA) surveys, web-based surveys and mobile surveys to address quality concerns
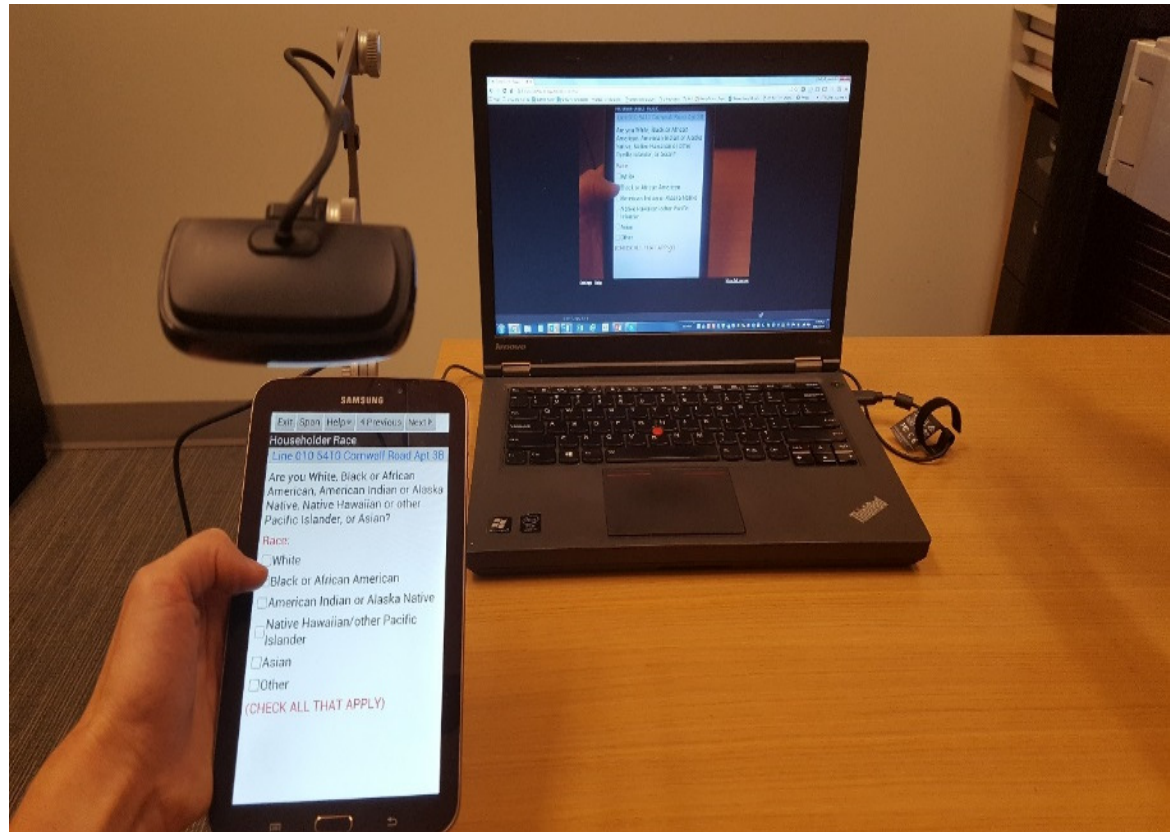
UT overview cont.

- Now, easy and inexpensive to apply UT
  - Screen recording software
  - Screening sharing
  - Web camera
  - Mobile sleds

- Supports
  - Lab or field testing
  - remote testing
  - Off-site observers

Advantage: explore respondent-survey interaction

Usability Model for Surveys

- **Interpreting the design**:
  - What meaning do respondents assign to visual design and layout?
  - How do respondents believe the survey works?

- **Completing actions and navigating**:
  - How well does the survey support respondents' ability to complete tasks and goals?
  - How well do respondents follow navigational cues and instructions?

- **Processing feedback:**
  - How do respondents interpret and react to the survey feedback in response to their actions?
  - How well does the survey help respondents identify, interpret, and resolve errors?

Source: Geisen and Romano Bergstrom, 2017

UT Examples

Romano Bergstrom & Strohl (2013)

	1 Never	2	3	4	5 All of the time
Teach others about something you have learned					
	1 Never	2	3	4	5 All of the time
Get better at doing something					
	1 Never	2	3	4	5 All of the time
Give updates throughout the day					
	1 Never	2	3	4	5 All of the time
Have fun					
	1 Never	2	3	4	5 All of the time
See/hear something entertaining					

UT Examples continued

Geisen, Olmsted, Goerman, Lakhe (2014)

- On mobile, screens with multiple questions had high omission rates or error rates
- Important information was skipped or ignored
- For questions with text entry, screen blocked questions

Romano & Chen, 2011

- Next/Previous buttons
- Participants prefer "Next" on right
- When previous on right, got used more

Eye Tracking Overview

- Tracks where an individual looks within a visual field
- Modern eye tracking is unobtrusive – technology built into computer monitors and stand-alone devices
- Records fixations and saccades

Image source: Geisen & Romano Bergstrom (2017)

Eye Tracking Advantages & Limitations

- **What does it tell us** (Olmsted-Hawala, et al.2014):
  - What people look at when completing a survey
  - How many times they look at various things
  - The order in which they look at things
  - How long they look at things

- **Does not work well when participant's gaze and attention are not together** (Jarrett & Gaffney, 2008)
  - Thinking up answers on the spot
  - Need to look up answer (e.g., refer to a receipt)
  - Third-party ("Dear, how much is our monthly mortgage?)

Eye tracking examples

- People don't read instructions (Romano & Chen, 2011)
- People read pages with questions differently from other pages
  (Jarrett & Romano Bergstrom, 2016)
- Participants do not read linearly, and skipped around the survey (Redline and Lankford, 2001)
- Participants spend more time looking at options on the top of a list of responses (Galesic, Tourangeau, Couper, and Conrad, 2008)
- Longer fixation time on question stems indicates comprehension issues (Lenzner, Kaczmirek, and Galesic, 2011)

Crowdsourcing

"Tapping into the collective intelligence of the public to complete a task"

- King, 2009

Crowdsourcing Overview

- What is it?
  - Remote, unmoderated usability or cognitive testing
  - Willing pool of individuals (crowd)
  - Participant panel or direct outreach
  - Perform tasks that researcher provides
  - Embedded scripted probes to assess tasks
  - All items are self-administered
  - Examples platforms include Mechanical Turk, TryMyUI, userzoom

Crowdsourcing advantages

Edgar, Murphy & Keating (2016)

- Lower costs (incentives, travel, set up)
- Timeliness: Some platforms can recruit 1000+ participants per day
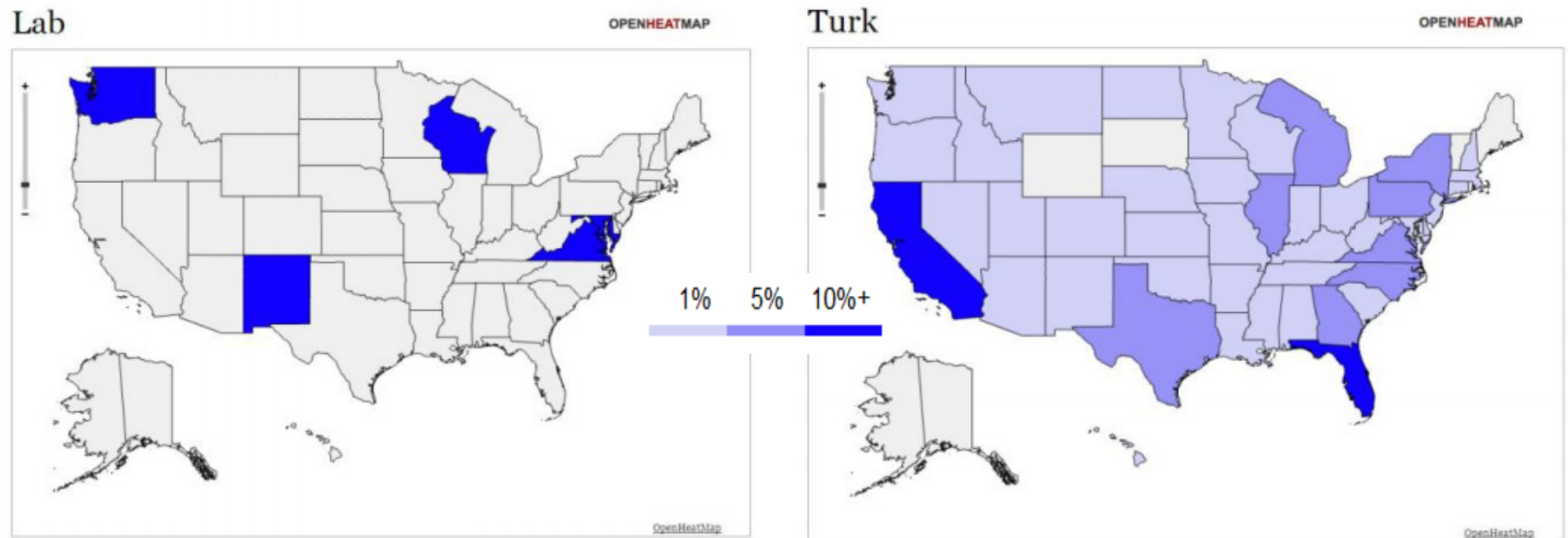- Wide geographic reach

Image source: Murphy, Keating, Edgar, 2014

Crowdsourcing examples

- Edgar, Murphy, & Keating (2016):
  - Worked as well as cog testing for simple tasks
    - Examples of sportswear
    - Define flu season
  - Cog testing better for complex tasks: explaining thought process for remembering clothes purchased recently
    - Major drawback: no spontaneous probing
- Murphy, Mayclin, Richards, & Roe (2015)
  - Used crowdsourcing to test alternative versions of a question that did not perform well in pilot test
- Can obtain quantitative data

Summary and Conclusions

Advantages and Limitations

Approach	Advantages	Limitations
Usability Testing	• Iterative process that guides design • Behavioral data • Performance measures • Evaluate mobile surveys	• Small sample sizes • Findings are mostly qualitative
Eye-tracking	• Implicit data not affected by self-report • Evaluate mobile surveys	• Requires specialized equipment to conduct • Does not work when participants' gaze and attention are not together
Crowd-sourcing	• Geographic diversity • Larger samples • Obtain quantitative data • Quick to conduct and implement • Less expensive	• Not great for complex tasks / tasks must be short • Potential for bias depending on how respondents recruited • No spontaneous probing

Conclusions

Emerging methods are not intended to replace existing methods; a multimethod approach works best.
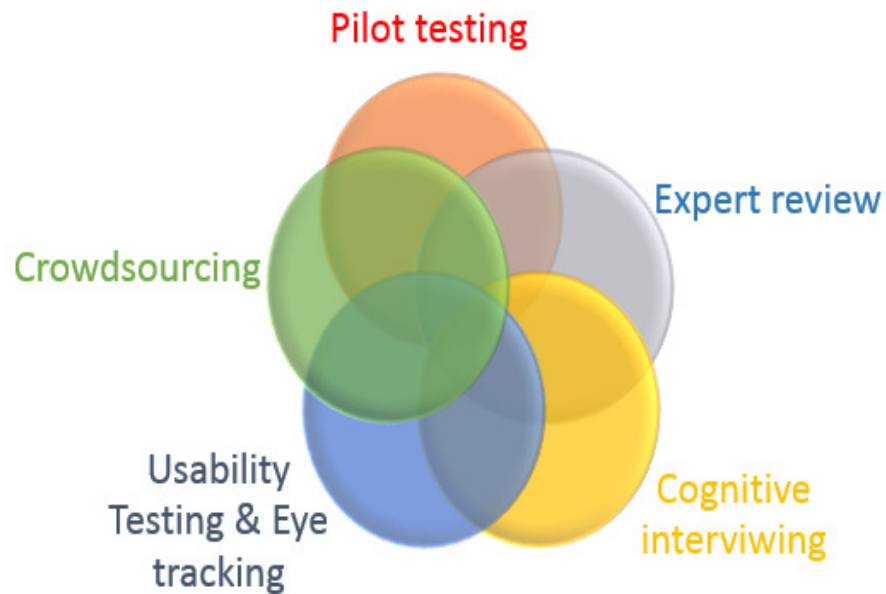
How different pretesting methods interact to identify potential issues with survey quality
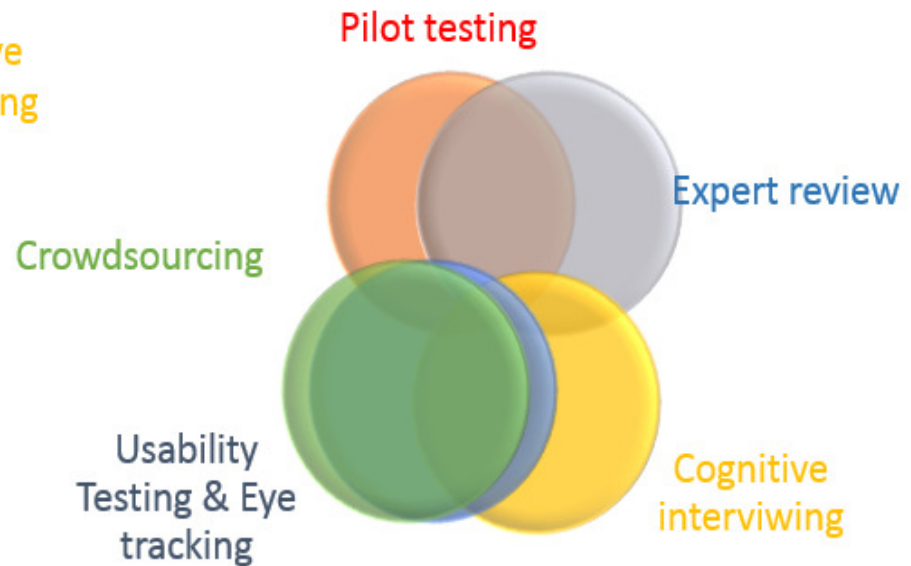
Depending on the survey, 1 or 2 methods may offer more advantage compared with others.

More Information

Emily Geisen

Survey Methodologist

919-541-6566

egeisen@rti.org

Joe Murphy

Survey Methodologist

312-456-5261

jmurphy@rti.org