Is a Proxy Response Good Enough? Using Paired Cognitive Interviews to Assess the Accuracy of Proxy Responses

Mary C. Davis, Jenna Fulton and Amber Henderson
Center for Survey Measurement, U.S. Census Bureau

Presented at DC-AAPOR/WSS Summer Preview/Review Conference
July 24, 2017

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

Outline

- Within household proxy reporting overview
- Recent CSM research using paired cognitive interviews
- Proposed guidelines for conducting paired cognitive interview research
- Examples and suggestions of analysis of paired cognitive interview data
- Conclusions and next steps

WITHIN HOUSEHOLD PROXY REPORTING – AN OVERVIEW

Within Household Proxy Reporting

- One household member responds to a survey about self and provides information for all other household members
- Advantages
  - Reduced costs (compared to self response) due to fewer visits to sampled household; faster data collection
  - Some evidence of greater data quality compared to self reports
- Disadvantages
  - Proxy motivation may differ, may lead to differences in effort, and higher response error
  - Potentially a tradeoff between reduced costs and data quality

Existing Research

- The proxy and target respondent are rarely both interviewed to determine if their responses align (Cobb & Krosnick, 2009)
- Limited existing research -
  - Explores the types of questions that can be more accurately answered by proxies (Krosnick et. al, 2015)
  - Evaluates pretesting approaches to determine a question's suitability for proxy response
- Blair et al. (1991) propose using a cognitive interviewing approach to evaluate questions by asking identical questions to pairs of respondents, and assessing their response strategies

RECENT CSM RESEARCH

Recent CSM Research

- Two supplements of the Current Population Survey:
  - Civic Engagement & Volunteerism (CIVIC)
  - Computer and Internet Use (COMPUTER)
- Sponsors
  - Bureau of Labor Statistics
  - Corporation for National and Community Service
  - National Telecommunications and Information Administration

CIVIC	COMPUTER
<u>Survey Content</u> • Interactions with friends, family, community • Political involvement; group membership & volunteering • Donations	<u>Survey Content</u> • Internet-connected devices • Locations of Internet use • Activities conducted using the Internet • Privacy and security concerns

<u>Survey Administration</u>
• Respondent first answers entire questionnaire about self
  • Provides response for each additional adult household member (age 18+)
	<u>Survey Administration</u>
	• Respondent first answers household level questions
	  • Provides person-based responses for one section (internet activities)

<u>Cognitive Testing Design</u> • 2 rounds • 21 pairs: 11 related / 10 unrelated	<u>Cognitive Testing Design</u> • 2 rounds • 14 pairs: 7 related / 7 unrelated

PROPOSED GUIDELINES FOR RESEARCH USING PAIRED COGNITIVE INTERVIEWS

Recruiting and Screening

- Difficult to recruit unrelated household members
  - Some respondents would represent their household structure as roommates when they were non-marital partners, and thus considered related pairs by our definition
- Screened participants who called in and got contact numbers for adult household members
  - Should also ask screening questions of household member to monitor recruiting goals and to collect demographic data

Scheduling and Interviewing Procedures

- Initial scheduling plan: interview pairs on the same day with the same interviewer
  - Respondents did not always have similar availability
  - May have affected our ability to recruit
- Revised scheduling plan: optional scheduling extension allowing respondents to be interviewed within 3 days of each other and by different interviewers
  - Increased potential for cancellations and ‘’no shows’’
  - Possibility of discussion of survey content with other member of pair

Interview Protocol

- Self-Response Probes
  - Question comprehension
  - Meaning of terms, phrases
- Proxy-Response Probes
  - How did respondent come up with answers for other household members?
  - How confident was respondent in answers for other household members?
  - Which questions were more challenging to provide a proxy response for?
  - Was it more challenging to provide a proxy response for one household member vs. another?

SUGGESTIONS AND EXAMPLES OF ANALYSIS OF PAIRED COGNITIVE INTERVIEW DATA

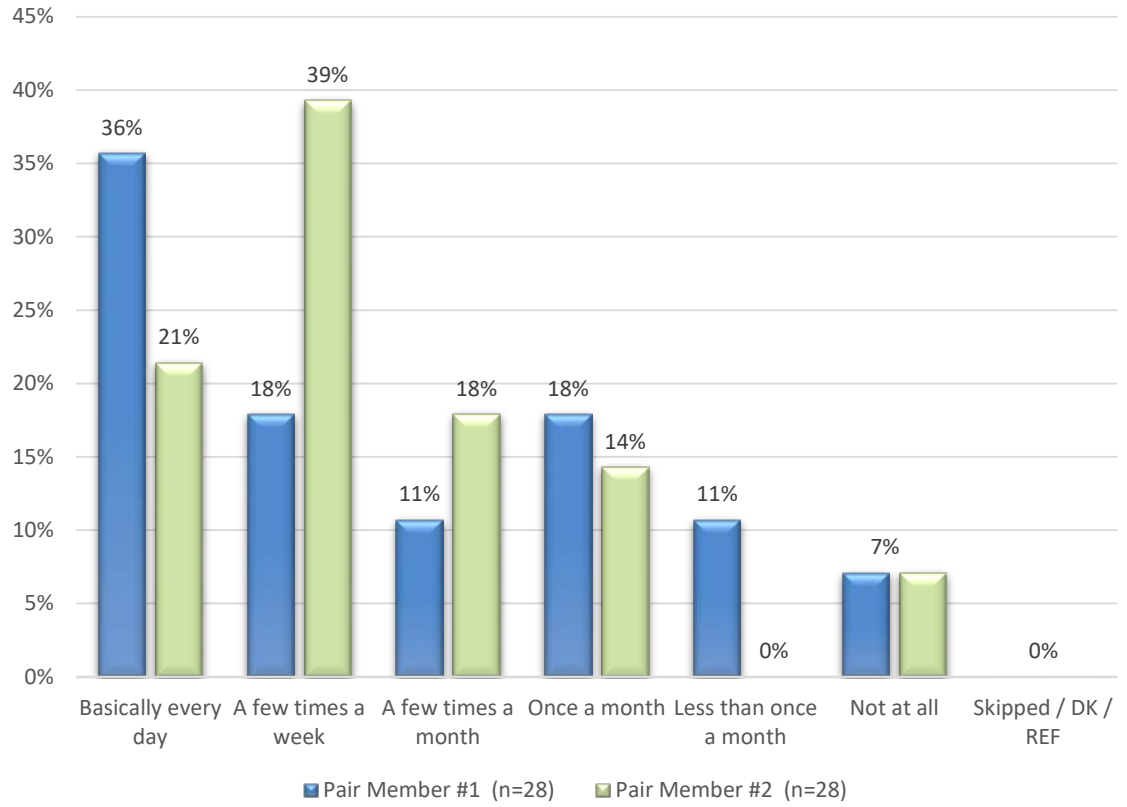
Analysis of Paired Cognitive Interview Data
- An Overview -

- Relies on quantitative, qualitative, and other roster and respondent data
- A variety of analyses can help demonstrate the degree of similarity between self and proxy reports
  - Frequency distributions among pair members
  - Rates of matching and near matching responses
- Note that assessing rates of matching responses does not provide information about the accuracy of these responses

Frequency Distributions Among Pair Members

- Example from CIVIC Round 2 Data -

- Demonstrates how the response distribution would vary depending on who reported for the household
- Pair Member #1s' responses for themselves and other household members are tallied for each question
  - Responses for Pair Member #2s are tallied separately and compared



Q2: How often did [you/[NAME]] discuss political, societal, or local issues with friends or family?

Calculating Exact Match Rates

- Calculate at the question level -
  - For each person-level response, code whether two reports match
  - Divide the number of matching responses by the total number of responses for that question to determine the question's exact match rate
- Can be calculated among pairs of respondents and for other household members who did not provide a self response, but for whom a response was provided by both respondents in the pair
- Match rates can be calculated for questions with closed or open response options

Exact Match Rates by Question Response Format
- Related and Unrelated Pairs -

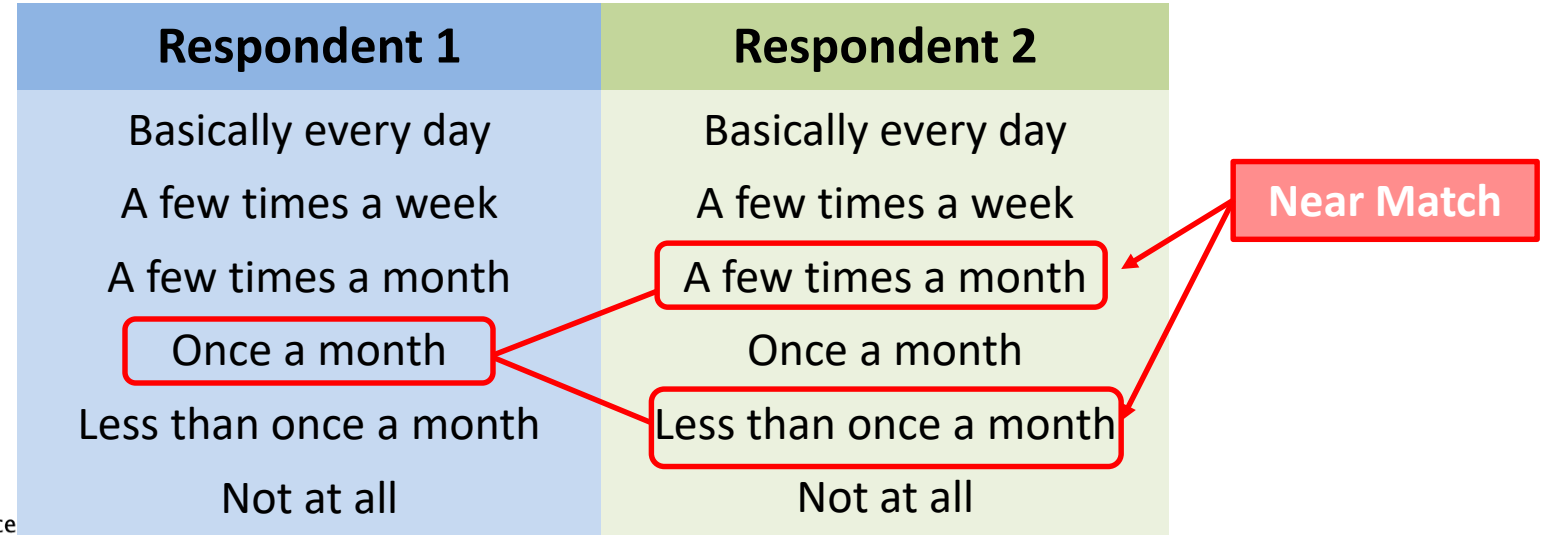
Question Response Format	# of Questions	Avg. Match Rate*	Related Pairs Match Rate*	Unrelated Pair Match Rate*
CIVIC response option formats	21	51%	49%	52%
6-Category, close-ended	11	41%	36%	46%
Yes / No	8	70%	72%	67%
Open-ended numeric	2	25%	26%	24%
COMPUTER Yes / No response format	40	72%	70%	75%

** Match rate data is calculated across both rounds of testing*

6-Category, Close-Ended Response Options	
(10 Questions) (1) Basically every day (2) A few times a week (3) A few times a month (4) Once a month (5) Less than once a month (6) Not at all	***(1 Question)*** (1) All activity is in person (2) Activity is more in-person than online (3) Activity is evenly split between in-person and online (4) Activity is more online than in-person (5) All activity is online (6) I am a member but not active with any groups, organizations or associations.

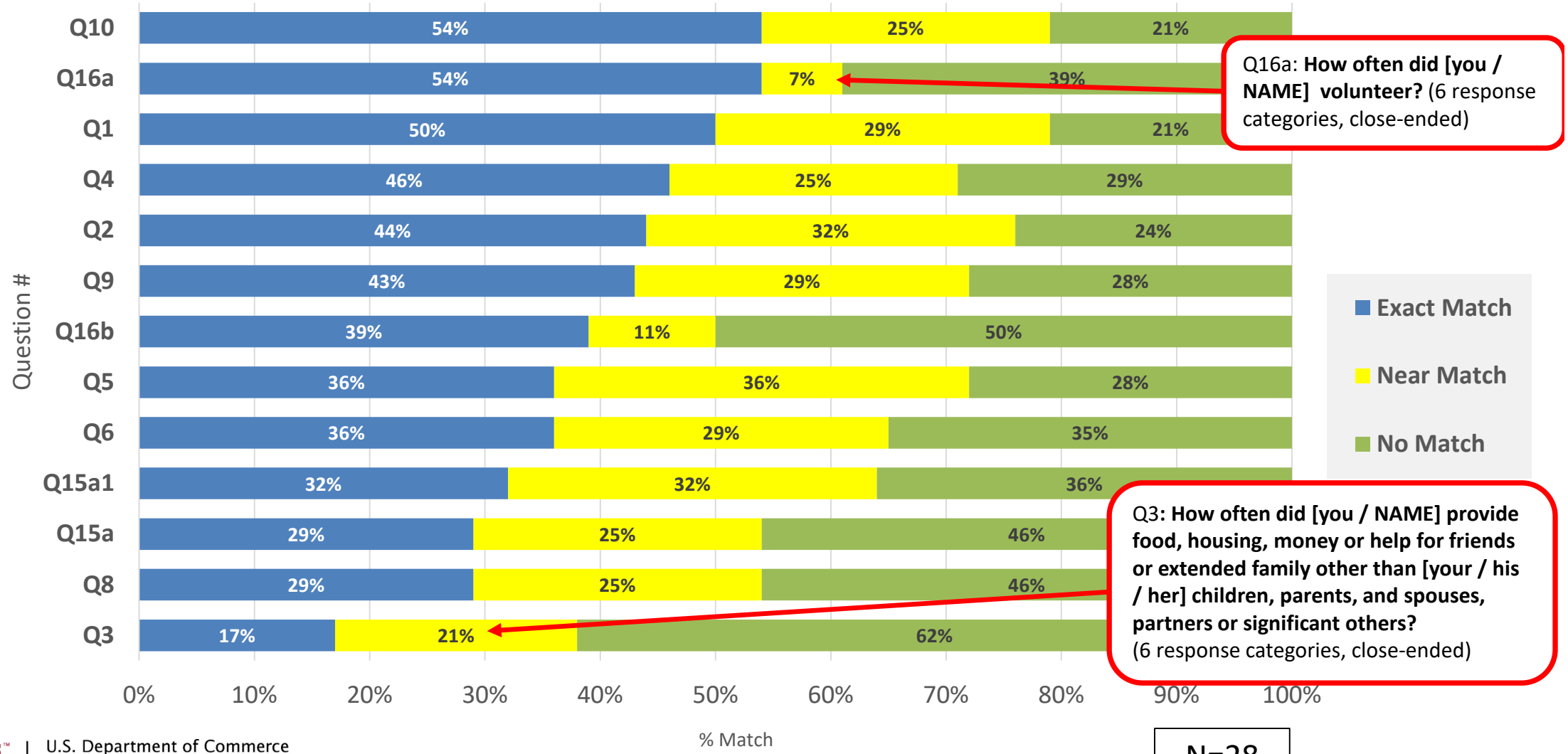
Rates of Near Matches using Quantitative Data

- For questions with more than two response options
- Near matches are responses provided for the same question by two respondents *that are not identical*, but are only one or more response categories apart
  - Number of categories apart is at the discretion of the researcher
  - Should reflect question design and goals
- May be especially useful if researchers are considering collapsing some response options during survey administration or analysis



Rates of Exact Matches & Near Matches

- CEV Round 1 Categorical and Open-Ended Questions -



Conclusions

- Pairs from related households appear to be no better able to provide matching responses than those from unrelated households
- Across two rounds of testing, the overall match rate was slightly above 50% for all questions in the Civic Engagement & Volunteerism survey
  - Match rates were lowest for open-ended and 6-category response options
- Across two rounds of testing for Computer & Internet Use, only 4 of 40 yes/no questions had matching responses below 50%
  - The overall match rate was 72% for questions with yes / no response options

Next Steps for Paired Cognitive Interview Research

- Explore additional analyses for data from both projects
  - Perceived levels of confidence and difficulty
  - Characteristics of proxy respondents that influence ability to provide matching responses
  - Characteristics of questions that are easier / more difficult to answer
- Examine use of administrative records to evaluate accuracy of responses for self and proxy

Contact Information

Mary.C.Davis@census.gov

Jenna.Fulton@census.gov

Amber.Henderson@census.gov

APPENDIX

References

1. Bickart, B.A., Blair, J., Menon, G. and Sudman, S. (1990). Cognitive Aspects of Proxy Reporting of Behavior, in NA - Advances in Consumer Research Volume 17, eds. Marvin E. Goldberg, Gerald Gorn, and Richard W. Pollay, Provo, UT: Association for Consumer Research, Pages: 198-206.
2. Blair, J., Menon, G., & Bickart, B. (1991). Measurement effects in self vs. proxy responses to survey questions: An information-processing perspective. In P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, & S. Sudman (Eds.), *Measurement Errors in Surveys* (145-166). Hoboken, NJ: John Wiley & Sons, Inc
3. Krosnick, J. A., Presser, S., Fealing, K. H., & Ruggles, S. (2015). The future of survey research: Challenges and opportunities. The National Science Foundation Advisory Committee for the Social, Behavioral and Economic Sciences Subcommittee on Advancing SBE Survey Research. http://www.nsf.gov/sbe/AC_Materials/The_Future_of_Survey_ Research.pdf. Accessed 10 Jul 2015.
4. Mathiowetz N.A., and Groves R.M. (1985)The effects of respondent rules on health survey reports. *Am JPublic Health.* 1985;75:639–644.
5. Moore, J. C. (1988). Self/proxy response status and survey response quality. *Journal of Official Statistics, 4*(2), 155–172.