

ARTICLES

Do Interviewers with High Cooperation Rates Behave Differently? Interviewer Cooperation Rates and Interview Behaviors

Kristen Olson¹, Antje Kirchner¹, Jolene Smyth¹

¹ University of Nebraska-Lincoln

Keywords: survey practice

<https://doi.org/10.29115/SP-2016-0011>

Survey Practice

Vol. 9, Issue 2, 2016

Interviewers are required to be flexible in responding to respondent concerns during recruitment, but standardized during administration of the questionnaire. These skill sets may be at odds. Recent research has shown a U-shaped relationship between interviewer cooperation rates and interviewer variance: the least and the most successful interviewers during recruitment have the largest interviewer variance components. Little is known about why this association occurs. We posit four hypotheses for this association: 1) interviewers with higher cooperation rates more conscientious interviewers altogether, 2) interviewers with higher cooperation rates continue to use rapport behaviors from the cooperation request throughout an interview, 3) interviewers with higher cooperation rates display more confidence which translates into different interview behavior, and 4) interviewers with higher cooperation rates continue their flexible interviewing style throughout the interview and deviate more from standardized interviewing. We use behavior codes from the Work and Leisure Today Survey (n=450, AAPOR RR3=6.3%) to evaluate interviewer behavior. Our results largely support the confidence hypothesis. Interviewers with higher cooperation rates do not show evidence of being "better" interviewers.

Introduction

The two most important tasks for interviewers in sample surveys are gaining cooperation and administering the survey questionnaire. Gaining cooperation requires flexibility, tailoring statements, and maintaining interaction with the sampled householder (Groves and Couper 1998). Standardized interviewing requires precisely following the script by reading questions exactly as worded, using nondirective probing, nondirective clarification, and neutral feedback procedures (Fowler and Mangione 1990). Thus, interviewers are instructed to be flexible on the doorstep but standardized in question administration.

How well can interviewers follow this dual role (flexibility followed by standardization)? Can interviewers switch from flexible recruitment to standardized question administration or are they good at only one of these tasks (e.g., those who are good at recruitment are bad at standardized question administration)? To examine this issue, this article develops and examines four hypotheses – conscientiousness, rapport, confidence, and flexibility – for how interviewer-level cooperation rates in a telephone survey

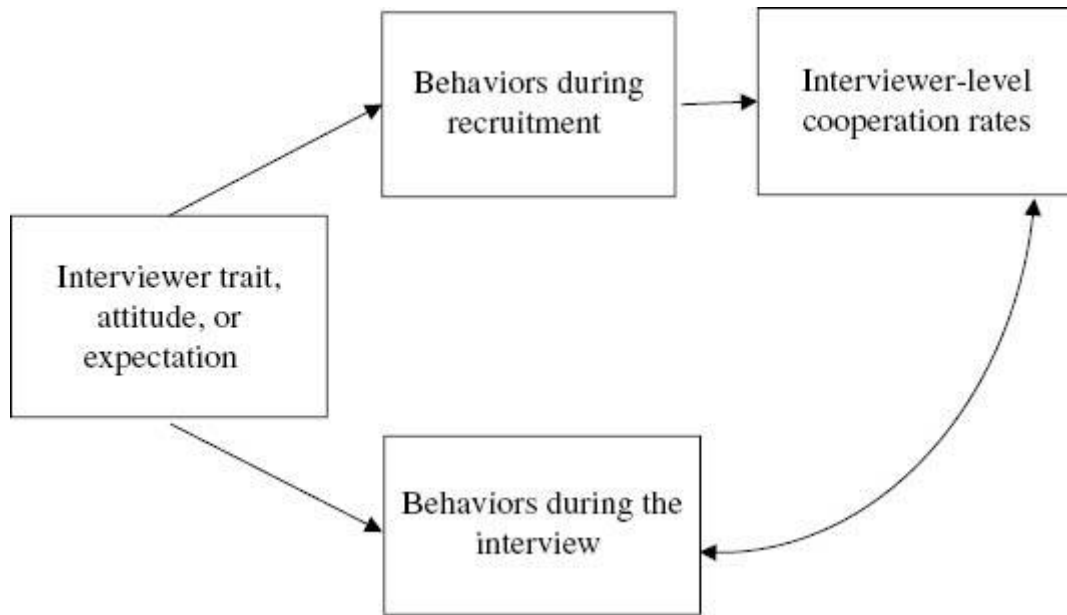


Figure 1 Conceptual model for correlation between interviewer cooperation rate and interview behaviors.

may be associated with interviewer behaviors during question administration. These interviewer behaviors provide a unique insight into potential correlates of data quality.

Hypothesized Mechanisms

An association between survey nonresponse and interviewer-related measurement error occurs when there is a common cause such as an interviewer personality trait, attitude, or expectation that affects interviewer behaviors during both recruitment and question administration (Brunton-Smith, Sturgis, and Williams 2012; Figure 1).

There are four general mechanisms through which recruitment and question administration behaviors may be jointly influenced. We assert that these interviewer traits, attitudes, or expectations manifest through the interactional processes between interviewers and sampled persons at the recruitment and question administration stage. These behaviors are generally constrained by training and monitoring and are likely to vary by interview mode. We also assume that higher response rates indicate more flexible recruitment behaviors, although we do not empirically examine this here. Table 1 shows how we operationalize each of these concepts.

Conscientiousness

Conscientiousness, one of the Big Five personality traits, includes such characteristics as being organized and following rules (John and Srivastava 1999). Interviewers who have higher cooperation rates may be more conscientious interviewers all around because they can follow the rules of training – i.e., they are able to be flexible and tailor during recruitment but standardized during question administration (e.g., Brunton-Smith, Sturgis,

Table 1 Definitions of interviewer behaviors.

Behavior	Definition
Conscientiousness	
Read exact	Read question exactly as worded
Adequate probe	Nondirective probing behavior, such as repeat the question exactly as written, repeat response options, take your best guess etc.
Exact verification	Repeat respondent's answer exactly as given
Appropriate clarification	Clarifying behavior, such as whatever it means to you, read a definition exactly as worded, clarify the unit in question
Appropriate feedback	Feedback behavior, such as short acknowledgements of response without any other statement, longer motivational feedback
Rapport	
Rapport-related feedback	Interviewer flatters or compliments the respondent, digresses in off-topic conversation or discloses personal information or attitudes
Task-related feedback	Task-, time- and telephone quality-related feedback
Laughter	Interviewer laughs
Affirmative feedback	Interviewer provides an affirmative statement
Confidence	
Stuttering	Interviewer restarts or repeats part of a word or phrase in question or response options
Disfluency	Whether there are any disfluencies, stutters, or repairs
Interruptions	Interviewer interrupts the respondent
Flexibility	
Read major change	Interviewer changes the meaning of the question, omits, or adds multiple words to the question wording
Inadequate probe	Directive probing behavior, such as repeat the question or response options with changes, other directive probes that lead the respondent toward a particular response
Inadequate verification	Repeats respondent's answer with changes but does not change the meaning

and Williams 2012). Thus, we would expect these interviewers to exhibit higher rates of standardized behaviors such as reading the question exactly as worded; nondirective probes (e.g., repeating the entire question); verifying responses appropriately; using appropriate clarification (e.g., “whatever it means to you”); and providing appropriate feedback (e.g., “thanks”).

Existing literature shows mixed results for whether more conscientious interviewers achieve higher response rates than less conscientious interviewers (e.g., Dutwin et al. 2014). We know of no studies that have examined conscientiousness and data quality.

Rapport

Rapport, although inconsistently defined in the literature, is generally thought of as interviewer friendliness or motivating behaviors (Garbarski, Schaeffer, and Dykema 2016). Interviewers who have higher cooperation rates may carry rapport-building behaviors from recruitment into question administration. Rapport can be measured by nontask behaviors such as laughter, off-script talk to put the respondent at ease, and non-neutral

feedback. If rapport is the mechanism linking cooperation rates to measurement error, we would expect interviewers with high cooperation rates to also have higher rates of these interview behaviors.

Existing literature shows that interviewers vary in rapport behaviors, such that increased verbal communication, friendliness, and projecting a positive self-image are related to response rates (e.g., Jäckle et al. 2013; Schaeffer et al. 2013), although, the relationship between the personality trait of agreeableness and cooperation are less conclusive (e.g., Dutwin et al. 2014). Rapport behaviors occur during question administration but are inconsistently linked to data quality (Schaeffer, Dykema, and Maynard 2010).

Confidence

The third mechanism that might link cooperation rates with measurement error is interviewer confidence or self-assurance. More confident interviewers may engage in behaviors that convey credibility of the requests, thus increasing cooperation rates, but also affecting question administration. Confidence may be conveyed through paralinguistic cues such as fewer disfluencies and less stuttering, shorter delays in responding to questions, and more interruptions of other speakers (Ketrow 1990; Kollock, Blumstein, and Schwartz 1985). If confidence is the mechanism at work, we would expect interviewers with high cooperation rates to have lower rates of disfluencies and stutters and more interruptions.

There is mixed evidence about the role of confidence and response rates. Overall confidence or assertiveness may be effective at first contact, but not at later contacts (Jäckle et al. 2013). Being confident that households can be persuaded tends to be associated with higher response rates in face-to-face surveys (e.g., Durrant et al. 2010). We know of no studies that explicitly examine interviewer confidence as a predictor of interview behaviors or measurement error.

Flexibility

The final possible mechanism linking cooperation rates with measurement error is interviewer flexibility during recruitment and question administration. Flexibility is different from rapport in that it reflects a general ability to tailor verbal behaviors to address individuals' concerns at recruitment (Groves and Couper 1998). Here, we would expect that interviewers with higher cooperation rates use a more conversational or flexible form of interviewing (Schober and Conrad 1997) resulting in higher rates of reading questions with major changes, inadequate probes, and verifications.

The approach to tailoring is generally positively associated with response rates (e.g., Groves and McGonagle 2001, but see Schaeffer et al. 2013). Additionally, there is ample evidence of interviewers adapting behaviors

during the interview to fit a situation at hand and ease the response task for the respondent (e.g., Maynard et al. 2002; Schaeffer, Dykema, and Maynard 2010). Whether these behaviors are related to doorstep tailoring is relatively unexplored.

Data

The data come from the Work and Leisure Today survey, a 15 minute RDD CATI survey of U.S. adults in landline telephone households fielded by AbtSRBI during summer 2013 (n=450, AAPOR RR3=6.3 percent) (see Olson and Smyth 2015 for details). Each interview was audio recorded, transcribed, and behavior coded. A team of trained undergraduate students was used for behavior coding, and two trained graduate students served as master coders to evaluate coding reliability for a random subset of 10 percent of the cases. Where there were disagreements, the master coders' codes were used.

Eight behavior codes were assigned to each conversational turn: the actor (e.g., interviewer); the initial action (e.g., question asked); an assessment of the initial action (e.g., question asked with changes); a more specific assessment of this action (e.g., question asked with slight changes); problems reading words in parentheses; laughter; disfluencies; and interruptions. The reliability of these codes was high ($\kappa > 0.90$ for most codes). The lowest kappa values were for the detailed assessments of the interactions; we focus only on those behaviors that meet a minimum kappa requirement of 0.40.

To increase the stability of our estimates, we exclude three interviewers with fewer than 10 interviews and two partial audio recordings, leaving a sample of 433 coded interviews conducted by 19 interviewers.

Dependent Variables

To construct our dependent variable, we first identified whether a given behavior occurred at least once during the question-answer sequence. We then summed the total number of questions on which a particular behavior occurred over all of the questions in an interview. Analytically, we account for the number of questions asked through an offset term to estimate the rate of occurrence. Table 2

provides descriptive statistics for these behaviors. For example, half of the questions were read exactly as worded, and interviewers provided affirmative feedback on approximately 40 percent of the questions.

Table 2 Rate of interviewer behaviors by interviewer-level cooperation rate.

	Overall	Lower interviewer cooperation rate	Higher interviewer cooperation rate
Conscientiousness			
Read exact	0.50	0.50	0.51
Adequate probe	0.15	0.14 ⁺	0.16 ⁺
Exact verification	0.17	0.20***	0.12***
Appropriate clarification	0.04	0.04	0.04
Appropriate feedback	0.07	0.08 ⁺	0.06 ⁺
Rapport			
Rapport-related feedback	0.01	0.01	0.01
Task-related feedback	0.02	0.02	0.02
Laughter	0.05	0.05 [*]	0.06 [*]
Affirmative feedback	0.39	0.38	0.40
Confidence			
Stuttering	0.06	0.06	0.05
Disfluency	0.29	0.35***	0.21***
Interruptions	0.12	0.11 [*]	0.14 [*]
Flexibility			
Read major change	0.12	0.11***	0.14***
Inadequate probe	0.06	0.05	0.06
Inadequate verification	0.07	0.07	0.07
Total	433	242	191

Note: 46.7 questions asked on average. ⁺ $p \leq 0.10$; ^{*} $p \leq 0.05$; ^{**} $p \leq 0.01$; ^{***} $p \leq 0.001$.

Independent Variables

Our primary independent variable is the interviewer cooperation rate. In this study, cases were randomly assigned to interviewers with no explicit refusal conversion attempts. Across the 19 interviewers, cooperation rates range from 3.9 percent to 10.5 percent, with a mean of 6.8 percent. We use a centered linear term for the cooperation rate in our models.

Control Variables

It is necessary to control for interviewer and respondent characteristics because interviewers may recruit nonrandom sets of respondents, even when assigned a random set of phone numbers (West and Olson 2010). Interviewer sex (47.4 percent female), race (47.4 percent white), and overall experience (73.7 percent with 1+ years of experience) are included to account for any potential interviewer effects on both response rate and interviewer behaviors. Respondent characteristics of sex (64.0 percent female), age (70 percent age 51+), education (41.8 percent college+), race (12.7 percent nonwhite), marital status (47.8 percent married), presence of children in the household (18.0 percent with children), employment status (40.4 percent employed), income (41.8 percent \$50,000+), and Internet use (69.3 percent internet users) are included.

Methods

We start with descriptive statistics evaluating differences in the occurrence of behaviors for interviewers with cooperation rates above or below the median. The bivariate analyses do not account for interviewer clustering.

Next, we evaluate the association between interviewer cooperation rates and interviewer behaviors during question administration using a two-level negative binomial multilevel model with the interviewer cooperation rate included as a linear term, the control variables, and an interviewer random effect (using the `menbreg` command in Stata 14). To model the dependent variable as a rate rather than a count, we include the number of questions as an offset so that the model predicts the rate of the behavior's occurrence over the total number of question asked (which varies across respondents; full models available from authors). When we evaluate interviewer variance in a null linear hierarchical model, virtually all interviewer-related intraclass correlation coefficients (except for directive probing) are significantly different from zero ($p < 0.05$) and account for between 6 to 73 percent of the total variance in behaviors.

Findings

Contrary to our initial expectations for the conscientiousness hypothesis, there are no significant differences between high and low cooperation rate interviewers for reading questions exactly as worded, providing adequate feedback, or probing behaviors (Table 2). Interviewers with a high cooperation rate tend to verify a response exactly as given on fewer questions (difference = -0.08 ; $p < 0.001$). In the multivariate models, verification is not related to cooperation rates, although interviewers with higher cooperation rates provide appropriate feedback (coef = -0.25 ; $p < 0.001$) significantly less often than interviewers with lower cooperation rates. Interviewers do not differ in any other conscientiousness behaviors. Thus, there is no support for the conscientiousness hypothesis – all significant associations are opposite the hypothesized direction.

Turning to the rapport hypothesis, bivariate and multivariate analyses show no differences by cooperation rate for rapport-related feedback or providing affirmative feedback. For laughter, high cooperation rate interviewers laugh on more questions than low cooperation rate interviewers (difference = 0.013 ; $p < 0.05$), although this association does not hold in the multivariate models. The interviewer cooperation rate is negatively associated with providing task-related feedback (coef = -0.29 , $p < 0.01$) in the multivariate models. Overall, there is little consistent evidence that interviewers with different cooperation rates vary in rapport behaviors.

All behaviors conveying confidence differ between high and low cooperation rate interviewers as predicted in either the bivariate or multivariate analyses. In the bivariate analyses, high cooperation rate interviewers have fewer

disfluencies (difference= -0.14 ; $p<0.001$) and more interruptions of respondents (difference= 0.02 ; $p<0.05$). These results hold in the multivariate analyses – both stuttering (coef= -0.18 ; $p<0.05$) and disfluencies (coef= -0.15 ; $p<0.05$) occur less often for high cooperation rate interviewers. Thus, there is strong support for the confidence hypothesis.

For the final flexibility hypothesis, our bivariate analyses suggest that interviewers with higher cooperation rates have significantly more occurrences of major changes in question reading (difference= 0.03 ; $p<0.001$), but this effect does not hold in the multivariate analyses. There is no difference in the occurrence of inadequate probing or verification in either bivariate or multivariate analyses. Thus, we have no consistent evidence to support the flexibility hypothesis.

Discussion

In this article, we examine the link between telephone survey recruitment and question administration. We theorize four common causes linking interviewer cooperation rates to measurement error and then examine whether interviewers' cooperation rates are associated with question administration behaviors indicative of each cause.

We find no support for the conscientiousness hypothesis; interviewers with high cooperation rates are no better or worse at standardized question administration than those with low cooperation rates. We also find little support for the rapport and flexibility hypotheses. This finding suggests that high cooperation rate interviewers do not undermine measurement by using rapport building and flexible question administration behaviors.

The hypothesis that receives the greatest empirical support is that of confidence. Interviewers with higher cooperation rates have fewer stutters and disfluencies during question administration than interviewers with lower cooperation rates. This finding suggests that these nonverbal mannerisms play important roles during both recruitment and measurement. Stuttering and other disfluencies can help give respondents more time to process survey questions. As such, our results suggest that interviewers with higher cooperation rates may inadvertently be reducing data quality by providing (subtly) less time to process and answer survey questions. This needs to be explored in more detail.

The data in this study has limitations. First, with only 19 interviewers, it is difficult to identify nonlinear relationships between an interviewer's cooperation rate and behaviors during the interview. In sensitivity analyses, the squared cooperation rate in our models was never significant ($p<0.05$). Second, some behaviors occur rarely, limiting our ability to examine them. Another possible limitation is the range of the cooperation rates between 3 and 10 percent. With such a low range, no interviewer had a "high" cooperation rate. However, this is also a strength. The low cooperation rates

occurred because the survey did not use multiple follow-up attempts or refusal conversions. This means that the recruitment interaction and question administration generally were conducted by the same interviewer. Thus, the cooperation rate was unaffected by multiple other interviewers attempting to recruit the household. Third, interviewer behaviors may change over the course of the field period, potentially affecting both recruitment and question-administration outcomes. This could lead to endogeneity between the independent and dependent variables. One possible solution would be to use response rates from a prior study, although this would limit inference to more experienced interviewers, or to prospectively cumulate response rates over the field period. Finally, this study used a telephone administration. We do not know whether these findings will translate to other modes, although we expect that in-person administration would amplify the association between recruitment and measurement behaviors.

Survey practitioners can use these findings to pinpoint the type of issues to focus on in interviewer hiring and training. Both low and high cooperation interviewers need training on how to administer standardized questions and how to avoid carrying rapport building and flexible conversation behaviors into the measurement process. Our results suggest that there is no need to target additional (re)training at interviewers with high cooperation rates in telephone interviews. Additionally, these results suggest that survey organizations could screen individuals for use of stutters and disfluencies at hiring time and evaluate interviewers, especially those with high cooperation rates, for their use of these nonverbal vocal mannerisms during the interview itself. If missing stutters and disfluencies are not allowing respondents to fully process and answer questions, interviewers may need to be trained to slow down in other ways. The association between these behaviors and data quality should also be examined.

Author Notes

An earlier version of this paper was presented at the Joint Statistical Meetings, August 2015, Seattle, WA, and at the Total Survey Error Conference, Baltimore, MD, September 2015. This material is based upon work supported by the National Science Foundation under Grant No. SES-1132015. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Thanks to Jill Dever and the Special Issue editors for comments.

Published: May 31, 2016 EDT.

REFERENCES

- Brunton-Smith, I., P. Sturgis, and J. Williams. 2012. "Is Success in Obtaining Contact and Cooperation Correlated With the Magnitude of Interviewer Variance?" *Public Opinion Quarterly* 76 (2): 265–86. <https://doi.org/10.1093/poq/nfr067>.
- Durrant, G. B., R. M. Groves, L. Staetsky, and F. Steele. 2010. "Effects of Interviewer Attitudes and Behaviors on Refusal in Household Surveys." *Public Opinion Quarterly* 74 (1): 1–36. <https://doi.org/10.1093/poq/nfp098>.
- Dutwin, D., J.D. Loft, Darling J.E., A.L. Holbrook, T. Johnson, R.E. Langley, P.J. Lavrakas, et al. 2014. "Current Knowledge and Considerations Regarding Survey Refusals." https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/RefusalTF_FINAL090814.pdf.
- Fowler, F.J., and T.W. Mangione. 1990. *Standardized Survey Interviewing: Minimizing Interviewer-Related Error*. Newbury Park: Sage Publications.
- Garbarski, D., N.C. Schaeffer, and J. Dykema. 2016. "Interviewing Practices, Conversational Practices, and Rapport: Responsiveness and Engagement in the Standardized Survey Interview." *Sociological Methodology*. https://www.researchgate.net/publication/283153466_Interviewing_Practices_Conversational_Practices_and_Rapport_Responsiveness_and_Engagement_in_the_Standardized_Survey_Interview.
- Groves, R.M., and M. Couper. 1998. *Nonresponse in Household Interview Surveys*. New York: John Wiley & Sons, Inc.
- Groves, R.M., and K.A. McGonagle. 2001. "A Theory-Guided Interviewer Training Protocol Regarding Survey Participation." *Journal of Official Statistics* 17 (2): 249–66. <http://www.jos.nu/Articles/abstract.asp?article=172249>.
- Jäckle, A., P. Lynn, J. Sinibaldi, and S. Tipping. 2013. "The Effect of Interviewer Experience, Attitudes, Personality and Skills on Respondent Co-Operation with Face-to-Face Surveys." *Survey Research Methods* 7 (1): 1–15. <https://ojs.ub.uni-konstanz.de/srm/article/view/4736/4990>.
- John, O.P., and S. Srivastava. 1999. "The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspectives." In *Handbook of Personality: Theory and Research*, edited by L.A. Pervin and O.P. John. New York: Guilford Press.
- Ketrow, S.M. 1990. "Attributes of a Telemarketer's Voice and Persuasiveness." *Journal of Direct Marketing* 4 (3): 7–21. <https://doi.org/10.1002/dir.4000040304/epdf>.
- Kollock, P., P. Blumstein, and P. Schwartz. 1985. "Sex and Power in Interaction: Conversational Privileges and Duties." *American Sociological Review* 50 (1): 34–46. <https://doi.org/10.2307/2095338>.
- Maynard, D.W., H. Houtkoop-Steenstra, N.C. Schaeffer, and J. Zouwen. 2002. *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*. New York: John Wiley & Sons, Inc.
- Olson, K., and J.D. Smyth. 2015. "The Effect of CATI Questions, Respondents, and Interviewers on Response Time." *Journal of Survey Statistics and Methodology* 3 (3): 361–96. <https://doi.org/10.1093/jssam/smv021>.
- Schaeffer, N.C., J. Dykema, and D.W. Maynard. 2010. "Interviewers and Interviewing." In *Handbook of Survey Research*, edited by P.V. Marsden and J.D. Wright. Bingley, UK: Emerald Group Publishing.
- Schaeffer, N.C., D. Garbarski, J. Freese, and D.W. Maynard. 2013. "An Interactional Model of the Call for Survey Participation: Actions and Reactions in the Survey Recruitment Call." *Public Opinion Quarterly* 77 (1): 323–51. <https://doi.org/10.1093/poq/nft006>.

- Schober, M.F., and F.G. Conrad. 1997. "Does Conversational Interviewing Reduce Survey Measurement Error?" *Public Opinion Quarterly* 61 (4): 576–602. <http://poq.oxfordjournals.org/content/61/4/576.full.pdf>.
- West, B.T., and K. Olson. 2010. "How Much of Interviewer Variance Is Really Nonresponse Error Variance?" *Public Opinion Quarterly* 74 (5): 1004–26. <http://poq.oxfordjournals.org/content/74/5/1004.full.pdf+html>.