INTERVIEWER REFUSAL AVERSION TRAINING TO INCREASE SURVEY PARTICIPATION[1]

Thomas S. Mayer and Eileen O'Brien, United States Census Bureau
Tom Mayer, United States Census Bureau, Washington DC, 20233

Key Words: Interviewer training; Refusal Aversion Training; Nonresponse; Interviewer-respondent interaction

INTRODUCTION

Initial nonresponse rates for ongoing Federal household surveys are increasing (e.g., Atrostic, Bates, Burt, Silberstein, & Winters, 1999). Survey researchers are now focusing on the development of innovative methods to improve survey response rates. One area of interest is the interviewer-respondent interaction and its influence on survey cooperation. The Census Bureau, among other organizations, is beginning to consider the merits of specialized training programs that build on an understanding of this interviewer-respondent interaction to help interviewers effectively deal with respondent reluctance and avoid refusals. The paper reports on a pilot study of one training protocol designed to do so.

The interaction between interviewer and respondent up to the point of the start or termination of the interview is generally short, between 1-5 minutes for face-to-face interviews (Groves & Couper, 1998), and even shorter for telephone interviews (Oksenberg, Coleman, & Cannell, 1986). Groves, Cialdini, & Couper (1992) suggested that inexperienced interviewers often create "soft-refusals" in these first interactions by pressing the respondent to make a decision too quickly. Groves & Couper (1998) also posited that such soft-refusals happen because interviewers do not give an effective response to a specific respondent concern. Furthermore, their work showed that inexperienced interviewers often misconstrue respondent questions as an indication of reluctance. Questions asked by a respondent, however, are often predictive of positive interview outcomes for several reasons: a) it shows the respondent is motivated and attending to the conversation; b) it allows the conversation to continue; and c) it provides the interviewer with information to use to tailor the interaction. Thus, their theoretical considerations regarding the nature of initial interviewer-respondent interactions and their influence on respondent cooperation suggest this: interviewer refusal avoidance training would do better to focus on quickly identifying a concern and quickly delivering a brief, situation-appropriate response.

Groves & Couper (1998) described the theoretical motivation for such a training protocol. Because experienced interviewers are more adept at *tailoring* (adapting their approach to the specific characteristics of a household) and *maintaining interaction* (continuing the conversation with the respondent), they are more successful in obtaining cooperation from respondents. In their model, the following components are necessary for tailoring to be successful: a) the interviewer must have a repertoire of techniques, strategies, phrases, etc., related to the particular survey request; b) the interviewer must be adept at reading the verbal and nonverbal cues from the respondent; c) the interviewer must be able to apply the appropriate strategy according to the cues received from the respondent; and d) the interaction between the interviewer and respondent must be long enough so that tailoring can be applied. This indicates a strong relationship between tailoring and maintaining interaction. Maintaining interaction, say Groves and Couper, works because as a conversation grows longer, it becomes less socially acceptable to break it off. In addition, tailoring and maintaining interaction work together because the longer the conversation lasts, the more cues are provided to the interviewer, thus increasing the opportunity to apply tailoring techniques (Groves, Cialdini, & Couper, 1992; Groves & Couper, 1998).

An independent review of the Census Bureau's interviewer training program for the Current Population Survey identifies both the absence and the importance of these components in interviewing success. In their evaluation, Doughty et al. (2000) reported:

> Insufficient training, practice, and feedback is provided for the most challenging tasks required of FRs. (These) are converting refusals, gaining cooperation of respondents, troubleshooting computer problems, communicating purpose of survey/use

---

[1] This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

of data, listing and coverage, establishing and maintaining rapport with respondent, understanding the job of an interviewer, and answering respondents' questions (p. 10).

Thus, CPS interviewers appeared to be ill-prepared for the tasks most important to the constructs of tailoring and maintaining interaction as described by Groves and Couper (1998). That is, Doughty et al. found that CPS interviewers were not sufficiently trained in converting refusals, gaining cooperation of respondents, communicating the purpose of the survey, establishing and maintaining rapport with the respondent, and answering respondent's questions.

In addition to the fairly specialized task of dealing with reluctance, Doughty et al. (2000) also reported that "FRs do not feel prepared to perform their job independently before N-1 [field training]" (p. 10). Thus, more general concerns about interviewers' level of self-confidence with their role may be an important issue. Earlier research demonstrates a strong association between confidence-levels and interviewing success. Singer, Frankel, & Glassman (1983) reported that interviewers with optimistic expectations about the ease of gaining respondent cooperation had significantly higher response rates than those with less optimistic expectations. Similarly, Groves and Couper (1998) reported that interviewers having higher levels of self-confidence about their ability to gain respondents' cooperation tended to have longer interactions with respondents. One of the strengths of the Refusal Aversion Training is its focus on behaviors that build interviewers self-confidence by addressing the difficult job responsibilities described by Doughty et al.

Groves and Couper's (1998) concepts of tailoring and maintaining interaction suggested that developing refusal aversion training is a four step process: a) collecting respondent concerns in their own words; b) developing "alternative kinds of information relevant to those concerns" (p. 266); c) training interviewers to group respondent concerns into categories; and d) training interviewers to quickly provide responses to the respondent, using wording appropriate for that particular respondent. Building on this work, Groves & McGonagle (in press) implemented specialized "refusal aversion training" for survey interviewers, including concepts and skills associated with the constructs of tailoring and maintaining interaction. In two independent studies they found that interviewer groups who received training increased cooperation rates by 10.0

and 13.6 percentage points, respectively. Both tests, however, were conducted in an establishment survey.

The current study expanded upon the research of Groves & McGonagle (in press) by implementing their design in a household survey environment. The Questionnaire Design Experimental Research Survey, (QDERS), is an omnibus RDD survey designed by researchers in the Center for Survey Methods Research (CSMR) at the Census Bureau. The current Refusal Aversion Training project functioned as a pilot study to test the feasibility of using the Groves-McGonagle training protocol in Census Bureau demographic survey training.

Although the motivation for this training protocol is built upon a theory of householder cooperation in a face-to-face survey environment, QDERS offered a useful and cost-effective bridge to the household context from earlier tests in establishment surveys. One important dimension of the Groves-Couper model, however, is missing from the QDERS experiment. In using a telephone survey to test the protocol, interviewers' ability to identify cues from which to tailor and maintain interaction with respondents are restricted to aural cues only (Groves & Couper, 1998). Thus, how these telephone survey results might generalize to a face-to-face survey is not clear. Groves and Couper do suggest the training effect with field enumerators might be greater precisely because they are afforded a richer set of cues (e.g., household environment, respondents' facial expressions, posture, gestures) to better implement tailoring, build rapport, and maintain interaction with respondents. For the current study we expected Refusal Avoidance Training to increase interviewers' ability to gain householder cooperation. This would be reflected in higher cooperation rates and decreases in refusal rates. Because of the limited cues inherent in a telephone survey interaction described above, we would expect these results to represent a conservative estimate of the training effect, (i.e., we would expect a greater response rate increase in a face-to-face survey environment).

METHOD
Participants
Participants included 24 interviewers from the Hagerstown Telephone Center (HTC) recruited to collect survey data for the QDERS. Interviewer tenure ranged from five months to 15 years. All interviewers had previously received the initial training conducted by HTC, and 18 interviewers had at some time attended a Refusal Avoidance Workshop conducted by HTC staff.

Materials

Focus groups. Two focus groups were held in order to create the materials used for the Refusal Aversion Training. To avoid any confounds with the actual training, these focus groups were conducted via video-conferencing with a total of eighteen experienced telephone interviewers from the two other Census Bureau's telephone centers (Tucson and Jeffersonville Telephone Centers). The purpose was to obtain verbatim respondent concerns and interviewer rebuttals that proved successful for these staff in surveys of similar topics (health care, residential finance, income, etc.)

The first focus group served to identify a list of verbatim respondent concerns, providing detailed terminology used by respondents. Prior to the second focus group, researchers grouped respondent concerns into thematic sets. Systematically working through concerns within each set, researchers prompted interviewers in the second focus group to provide situation-appropriate rebuttals for each concern. Interviewers were repeatedly reoriented toward providing exact words and phrases that had proved successful in the past in gaining cooperation from respondents with specific concerns.

Refusal Aversion Handbook. An interviewer handbook was developed based on the training materials used by Groves & McGonagle (in press) and material gathered in the aforementioned focus groups. It was composed of three components: a thorough description of the refusal aversion process, a catalog of ten dominant themes of respondent reluctance for the particular survey, and a 'toolkit' section where interviewers could record new concerns and rebuttals identified in training or in ongoing work with the survey.

Drawn from Groves and Couper (1998) and Groves & McGonagle (in press), the essential building block of this Refusal Aversion Training was described in the handbook's first section as The Five Basic Steps to Encouraging Survey Response. Here interviewers learned to identify distinct parts of the refusal aversion process: a) Prepare for the call; b) Engage in active listening; c) Diagnose the main barrier; d) Quickly identify a counter response; and e) Quickly deliver a clear, brief rebuttal. This component is generalizable to all surveys.

The second section of the handbook offered ten common themes of reluctance expressed by respondents as they might relate to a hypothetical RDD, omnibus survey such as QDERS. The themes included: a) Legitimacy Concerns; b) Time and Burden Concerns; c) "Why me?" Concerns; d) Purpose Concerns; e) Confidentiality Concerns; f) Government Concerns; g) Voluntary Survey; h) Burn-out; i) Refusals; and k) Income.

Under each main theme, the handbook noted several representative examples of verbatim respondent concerns and successful, verbatim interviewer rebuttal strategies. Each example was derived from the focus groups, previous Bureau research or from other surveys with survey concepts similar to QDERS.

We feel that whereas some themes of respondent concerns may be common across surveys, (e.g. Time and Burden, Confidentiality, etc.), it would be necessary to rebuild this section of the handbook for use in other surveys. Unique design features of a survey can influence a householder's ability to assess the interviewer's intent, (e.g., mode of contact, use of advance letters, etc.), and how an interviewer may use them to tailor an effective rebuttal strategy for a particular respondent. The emphasis of the Refusal Aversion Training is not simply drilling interviewers on what strategies to use to address reluctance, but when and how to best use them. As such, both the composition of what is said and/or done must be relevant for the particular survey and tailored as needed to individual respondents.

Exercises. A number of exercises were designed to familiarize interviewers with the refusal aversion process and systematically increase their ability to use the process in a production setting. Therefore, the first exercises were fairly relaxed and gave participants a chance to consciously go through the steps in the process. Concluding exercises focused on increasing the speed and accuracy with which interviewers used the process.

Procedure

Interviewers from the Hagerstown Telephone Center (HTC) collected data for the QDERS. Two independent Random Digit Dialing (RDD) samples were used. Each sample included 4,000 cases and was in the field for approximately two weeks. The first half sample was fielded in August 2000; the second in September 2000.

Participants were divided into three groups containing eight interviewers each, (one control group and two test groups). The Control Group only attended the training sessions for the content of the QDERS and did not receive Refusal Aversion Training. The Before Group received Refusal Aversion Training before the beginning of the first data collection period. The Between Group received the Refusal Aversion Training before the beginning of the second data collection period. This design allowed for both between subject and within subject analysis. Using two independent samples eliminated any confounds that might be associated with a within-subject analysis conducted using only one sample (e.g., differences in the number of cases, etc.).

Refusal Aversion Training for both the Before and Between Groups consisted of eight hours of training given on two consecutive days. The first day of training included lectures on the process of encouraging survey response and basic exercises designed to reinforce specific steps of this process. The second day of training consisted almost entirely of exercises designed to increase the speed with which interviewers could effectively implement the process.

RESULTS

Interviewer-level First Contact Cooperation Rates (completed interviews and sufficient partial interviews/eligible contacted households) were calculated for each QDERS interviewer[2]. CATI case management practices resulted in cases being randomly assigned to the next available interviewer. If contact was established with a household but the interview was not completed, subsequent attempts could be made by any interviewer working the survey. It was therefore necessary to calculate these rates using only the outcome of the first contact with an eligible household, thereby avoiding any confounds associated with different interviewers (possibly from different experimental groups) conducting follow-up interviews with households. Rates were calculated for each interviewer for both data collection periods. Comparisons were made not only between experimental groups, but also within experimental groups over time.

A two factor analysis of variance with a repeated measure on one factor for the First Contact Cooperation Rate revealed a significant main effect for Group, $\underline{F}(2, 17) = 4.79$, $\underline{p} < 0.05$, and significant main effect for Time, $\underline{F}(1, 17) = 14.15$, $\underline{p} < 0.01$, but no interaction between Group and Time, $\underline{F}(2,17) = 1.57$, $\underline{p} > 0.10$. As can be seen from Figure 1, the

---

[2] Whereas interviewer-level First Contact Cooperation rates were used for the analysis, overall QDERS response rates were also calculated using the guidelines of the American Association for Public Opinion Research (AAPOR, 2000). The total number of completed cases for QDERS was 1,862 with a final response rate of 42-59%. The lower estimate (42%) includes in the denominator all cases of unknown eligibility (i.e., never-contacted cases whose status as working residential telephone numbers is uncertain), which corresponds to definition "RR2" in the AAPOR guidelines. The higher estimate (59%) excludes such uncertain cases, in correspondence with definition "RR6." Both estimates include in the numerator 54 cases designated as "sufficient partial" interviews, as well as the 1,808 which were fully completed.

Before Group (N=7) had the highest First Contact Cooperation Rates, followed by the Between Group (N=6), with the Control Group (N=6) showing the lowest cooperation rates. It also shows that for all three groups, First Contact Cooperation Rates increased from data collection period one to data collection period two. The lack of a significant interaction between Group and Time is most likely a factor of the small sample sizes of the groups, but could also be influenced by the fact that training was administered at different times in the production schedule for the two treatment groups. Because of the importance of this pilot study, and the potential practical significance of these results, follow-up comparisons were made to further explore the effects of the training.

Tukey multiple comparison tests ($\underline{p} = 0.10$) did not reveal any significant differences in the First Contact Cooperation Rates between any of the experimental groups for the first data collection period. Because neither the Control nor the Between Groups had received any Refusal Aversion Training, however, the groups were combined in order to increase the statistical power. The First Contact Cooperation Rate for the combined non-trained groups ($\bar{x} = 26.2$) was significantly lower than that of the Before Group, $\underline{t}(18) = -2.01$, $\underline{p} < 0.10$. This suggests that the Refusal Aversion Training was successful in increasing the First Contact Cooperation Rates of trained versus untrained interviewers.

All three groups showed an increase in First Contact Cooperation Rates from the first data collection period to the second. This suggests that some natural learning (maturation) took place over time, contributing to higher cooperation rates in the second data collection period for all groups. Both the Before and Between Groups, however, showed significant increases in First Contact Cooperation Rates from the end of the first data collection period to end of the second, $\underline{t}(6) = -5.61$, $\underline{p} < 0.01$ and $\underline{t}(5) = -3.07$, $\underline{p} < 0.05$, respectively, whereas the Control Group showed no significant increase, $\underline{t}(6) = -0.81$, $\underline{p} > 0.10$. These results suggest that the Refusal Aversion Training contributed to the learning process over and above what might have occurred naturally over time for interviewers with no training. As can be seen from Figure 1, First Contact Cooperation Rates increased about 3.7 percentage points for the Between Group, and about 7.2 percentage points for the Before Group, but only increase about 2.4 percentage points for the Control Group.

For the second data collection period, as expected, Tukey multiple comparison tests ($\underline{p} = 0.10$) showed that the Before Group achieved significantly

higher cooperation rates than the Control Group. This supports the notion that the Refusal Aversion Training was successful in increasing the First Contact Cooperation Rates of the interviewers who participated in the training. In addition, the multiple comparisons also showed that the Before Group had significantly higher cooperation rates than the Between Group. Since both groups had received training, this result suggests that there might be a "kick in" effect of the Refusal Aversion Training such that it takes interviewers a period of time (e.g., about a week) to apply the knowledge that they have learned in the Refusal Aversion Training. That is, whereas the Refusal Aversion Training did seem to help interviewers increase their cooperation rates for the first two weeks after training, it appears that the benefit increases for the third and forth weeks after training. This was echoed in interviewers' remarks in a post-data collection debriefing session. In the words of one interviewer, "It took about a week of using it (the Refusal Aversion Training) before it all came together and I got comfortable with it."

Because the two treatment groups received training at different times we might speculate as to how such a "kick in" effect might operate. One caveat is that conclusions drawn from comparisons between these groups do not account for individual differences between the interviewers (interviewer bias). That said, however, it is useful to explore the pattern of results for First Contact Cooperation Rates as they might occur over time. Step one is to note that for the first data collection period the combined non-trained group of interviewers shows cooperation rates of about 26 percent (see Figure 1). Step two is to examine the First Contact Cooperation Rates for the two-week period after training (this would be the rate for the first data collection period for the Before Group and the rate for the second data collection period for the Between Group). We see that the First contact Cooperation Rates are 30.7 percent and 33.0 percent for the Between Group and Before Group, respectively. In the final step we see that the rates for the second data collection period for the Before Group (40.2 percent) represent the First Contact Cooperation Rate for a two to four week period after training. Assuming that the cooperation rates for the Between and Control groups for the first data collection period represents the baseline cooperation rates for untrained interviewers, and that the rates for the Before group represent the potential improvement after up to four weeks post-training, it might therefore be possible to speculate gains of up to 14 percent (from 26 to 40 percent) from no training to a period two to four weeks after receiving Refusal Aversion Training. In future research it would be interesting to further examine the effect of time to explore this possible "kick in" effect, and also determine if there is a point after training where learning levels off, and perhaps, even starts to decline.

Conclusions

As hypothesized, First Contact Cooperation rates increased in the range of 3-7 percentage points for interviewers who participated in the Refusal Aversion Training and as much as 14 percentage points over time compared to those who did not receive the training. We believe that these results support further testing of such Refusal Aversion Training in a face-to-face survey environment. We recommend, however, that follow-up research include: a) a larger sample of interviewers; b) a longer data collection period; c) baseline data for all participant groups; and d) the design of an interviewer evaluation as a management tool. Ideally, such a test would be conducted on a more focused, complex survey data collection with a history of response rate problems. In addition, such a test would permit further investigation of the effects of the training over time, such as when it "kicks in" and how it may erode over time. Such evidence could suggest when best to deliver the training, how often, on which staff and so on. Ultimately, this would yield important information that could guide management decisions regarding interviewer training.

REFERENCES

The American Association for Public Opinion Research (2000). Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Ann Arbor, Michigan: AAPOR.

Atrostic, B. K., Bates, N., Burt, G., Silberstein, A., & Winters, F. (1999). Nonresponse in household surveys: New measures and new insights. Paper presented at the International Conference on Survey Nonresponse, Portland, OR, October, 1999.

Billiet, J., & Loosveldt, G. (1988). Improvement of the quality of responses to factual survey questions by interviewer training. Public Opinion Quarterly, 52, 190-211.

Couper, M. P. (1997). Survey introductions and data quality. Public Opinion Quarterly, 61, 317-338.

Couper, M. P., & Groves, R. M. (1992). The role of the interviewer in survey participation. Survey Methodology, 18, 263-277.

Doughty, P., Foley, A., Spuches, C., & Yonai, B. (2000). Current Population Survey interviewer training evaluation study: Final Report. Report submitted to the US Bureau of the Census, Januray, 2000.

Fowler, F. J., & Mangione, T. W. (1990). Standardized survey interviewing: Minimizing

interviewer-related error. Newbury Park: Sage Publications.

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. Public Opinion Quarterly, 56, 475-495.

Groves, R. M., & Couper, M. P. (1998). Nonresponse in household interview surveys. New York: John Wiley.

Groves, R. M., & McGonagle, K. (in press). A theory-guided interviewer training protocol regarding survey participation.

Kahn, R. L., & Cannell, C. F. (1957). The Dynamics of Interviewing: Theory, technique, and cases. New York: John Wiley & Sons, Inc.

Maynard, D. W. (1997). Keeping the gate: Declinations of request to participate in a telephone survey interview. Sociological Methods & Research, 26, 34-79.

Miller-Steiger, D., & Groves, R. M. (1997). Interviewer training techniques: Current practice within survey organizations. Ann Arbor, MI: Survey Research Center.

Morton-Williams, J. (1993). Interviewer Approaches. Aldershot, England: Dartmouth Publishing Company Limited.

Morton-Williams, J., & Young, P. (1987). Obtaining the survey interview – an analysis of tape recorded doorstep introductions. Journal of the Market Research Society, 29, 35-54.

Oksenberg, L., & Cannell, C. (1988). Effects of interviewer vocal characteristics on nonresponse. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, and J. Waksberg (Eds.), Telephone Survey Methodology (pp. 257-269). New York: John Wiley & Sons.

Oksenberg, L., Coleman, L., & Cannell, C. F. (1986). Interviewers' voices and refusal rates in telephone surveys. Public Opinion Quarterly, 50, 97-111.

Singer, E., Frankel, M. R., & Glassman, M. B. (1983). The effect of interviewer characteristics and expectations on response. Public Opinion Quarterly, 47, 68-83.

Figure 1

Mean First Contact Cooperation Rates for Groups by Data Collection Period
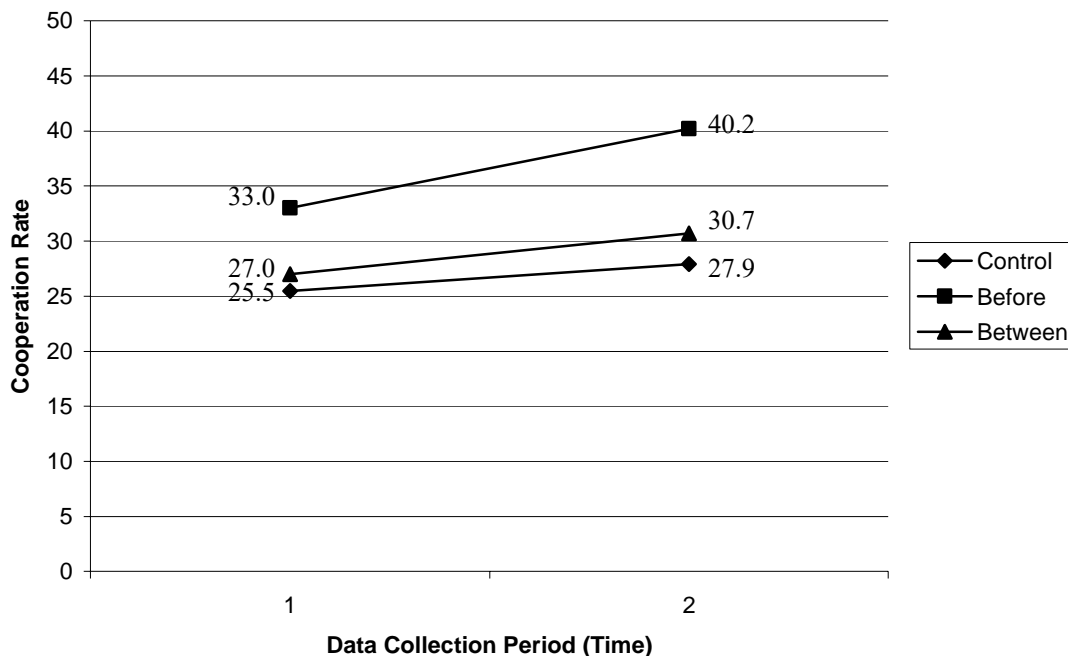
Note: Average number of first contacts = 159