



## **VI Jornadas de Investigación Científica**

*"15 años de la Facultad de Ciencias Sociales"*

**11 y 12 de setiembre de 2007**

**URUGUAY**

# Experiments in producing nonresponse bias

Robert M. Groves  
Mick P. Couper  
Stanley Presser  
Eleanor Singer  
Roger Tourangeau  
Giorgina Piani  
Acosta Lindsay Nelson

## **EXPERIMENTS IN PRODUCING NONRESPONSE BIAS**

---

ROBERT M. GROVES  
MICK P. COUPER  
STANLEY PRESSER  
ELEANOR SINGER  
ROGER TOURANGEAU  
GIORGINA PIANI ACOSTA  
LINDSAY NELSON

**Abstract** While nonresponse rates in household surveys are increasing in most industrialized nations, the increasing rates do not always produce nonresponse bias in survey estimates. The linkage between nonresponse rates and nonresponse bias arises from the presence of a covariance between response propensity and the survey variables of interest. To understand the covariance term, researchers must think about the common influences on response propensity and the survey variable. Three variables appear to be especially relevant in this regard: interest in the survey topic, reactions to the survey sponsor, and the use of incentives. A set of randomized experiments tests whether those likely to be interested in the stated survey topic participate at higher rates and whether nonresponse bias on estimates involving variables central to the survey topic is affected by this. The experiments also test whether incentives disproportionately increase the participation of those less interested in the topic. The experiments show mixed results in support of these key hypotheses.

ROBERT M. GROVES, MICK P. COUPER, STANLEY PRESSER, and ROGER TOURANGEAU are with the University of Michigan and Joint Program in Survey Methodology. ELEANOR SINGER, GIORGINA PIANI ACOSTA, and LINDSAY NELSON are with the University of Michigan. This is a revised version of a paper prepared for presentation at the sixteenth annual International Workshop on Household Survey Nonresponse, Tällberg, Sweden, August 28–31, 2005. We appreciate the support provided by a grant from the U.S. National Science Foundation (SES-0207435). Address correspondence to Robert M. Groves; e-mail: bgroves@isr.umich.edu.

doi:10.1093/poq/nfl036

© The Author 2006. Published by Oxford University Press on behalf of the American Association for Public Opinion Research. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org.

## Introduction

Although nonresponse rates to sample surveys are increasing in most industrialized countries (de Leeuw and de Heer 2002), recent case studies show little relationship between nonresponse rates and nonresponse bias (Curtin, Presser, and Singer 2000; Keeter et al. 2000; Merkle and Edelman 2002). Further, a recent analysis of some 30 nonresponse bias studies shows that although nonresponse bias may be common, the nonresponse rate by itself does not predict well the amount of nonresponse bias (Groves 2006).

Nonresponse has several harmful effects beyond potential nonresponse bias, however. Response rates continue to be used as quality indicators in many disciplines. Thus, when probability samples are used, survey organizations use follow-up efforts and persuasion attempts to increase the participation of difficult-to-contact and reluctant sample persons. Because these efforts cost money, fixed research budgets result in a smaller numbers of interviews. Smaller data sets, in turn, produce estimates with higher standard errors, and the standard error estimates are usually biased as well because they fail to reflect the stochastic nature of the response mechanisms. Survey statisticians have developed a variety of postsurvey adjustments to ameliorate nonresponse bias, but all of these techniques involve assumptions about the relationship between response propensity and the specific survey estimate in question. Thus, insight into the mechanisms linking response propensity to nonresponse bias is sorely needed to guide survey practice.

### THEORETICAL FRAMEWORK

We see the decision to respond to a survey as influenced by a set of predispositions of the sample person that, in turn, are based on personal experiences, social statuses, group norms, personality traits, and personal interests. Some of these predispositions might be quite central to the sample person (e.g., "My job requires 100 percent of my attention"). Others may be less central (e.g., "Science is good"). In interviewer-administered surveys, sample persons may vary in what they infer about the intent of the survey from observable interviewer traits (e.g., gender, age, race/ethnicity, accent). In addition, information explicitly conveyed about the survey provides the sample person further cues about the request. Hence, although predispositions of the sample persons guide their decisions, the stimuli that are prominent in the brief exposure to the survey request can be highly variable over recruitment episodes.

In interviewer-assisted surveys, this information varies across interviewers, based on what they choose to emphasize. In self-administered questionnaires, on the other hand, advance mailings, postcards, cover letters, and the questionnaire itself can vary in emphasizing the sponsorship of the survey, the data collection organization, the topic of the survey, the uses of the data, the confidentiality provisions, and so forth. When what is made salient (e.g., a reputable

sponsor) stimulates positive predispositions of the sample person, then response propensities are increased. When what is made salient (e.g., a threatening or embarrassing topic) stimulates negative predispositions, then response propensities are decreased. This perspective has been labeled “leverage-salience” theory (Groves, Singer, and Corning 2000) because the effect of any particular stimulus on a person is a joint function of its centrality to the person (its leverage) and its salience relative to the survey introduction’s other stimuli.

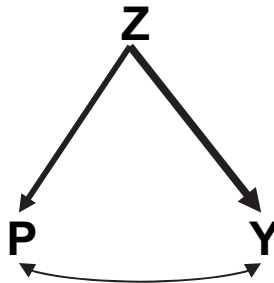
Nonresponse bias in a survey estimate arises when the set of mechanisms that influence the participation decision is related to the variables involved in the estimate. As Bethlehem (2002) observed, the nonresponse bias in the unadjusted respondent mean on the variable  $y$  is:

$$\frac{\sigma_{yp}}{\bar{p}}$$

where the numerator is the covariance between the variable,  $y$ , and the response propensity,  $p$ , and the denominator is the mean response propensity in the sample.

Discovering when nonresponse affects error requires finding conditions where this covariance departs from zero. One such condition, brought about when both the response propensity,  $p$ , and the survey variable,  $y$ , are influenced by the same attribute—in essence, a “shared cause”—is shown in figure 1.

In this model, some  $Z$  variable (or vector of variables) influences both the propensity to respond,  $P$ , and the survey variable in question,  $Y$ . The covariance between  $Y$  and  $P$  arises because of that shared influence; further, the diagram describes the case where the covariance of  $Y$  and  $P$ , controlling on values of  $z$ ,



**Figure 1.** The “common cause” model of a  $Z$  variable causing both response propensity,  $P$ , and the survey variable,  $Y$ .

is zero. This corresponds to the missing at random notion of Little and Rubin (2002) (in contrast to “missing completely at random”). Groves (2006) shows a variety of other causal models involving  $P$  and  $Y$  that have very different implications.

An example of a  $Z$  variable that, when controlled, reduces the covariance between  $Y$  and  $P$  to zero is interest in the survey topic.  $Y$  variables related to topic interest would include knowledge, attitudes, and behaviors related to the topic. For example, Baumgartner et al. (1998) show that participants in an electric utility’s time-of-day pricing program had higher response rates to an attitude survey about such programs than nonparticipants did. This implies that estimates of the general public’s knowledge of the program would be affected by the higher participation of the more interested. Likewise, Groves, Singer, and Corning (2000) show that those involved in their community responded at higher rates to a survey on an important community issue than did others. This produces higher estimates of community involvement among respondents than in the full sample. However, Groves, Presser, and Dipko (2004) show that the mechanisms underlying these effects are not consistent over situations that should produce them, leaving unanswered questions about when topic interest follows the common cause model and when it does not.

#### KEY QUESTIONS

This article addresses two unanswered questions about the nature of the mechanisms producing nonresponse bias:

1. Under what circumstances do influences arising from topic interest affect nonresponse bias?
2. How do monetary incentives affect the link between topic interest and nonresponse bias?

### Research Design

To address these questions we designed a randomized experiment that systematically varies the survey topic across samples from populations believed to have different levels of interest in the topic. We identified three populations judged a priori to have different levels of interest in two different survey topics. Sampling frames were acquired for each of the three populations, and probability samples were drawn from each. Two experimental factors were crossed in the design, through random assignment to each of the three populations: the topic of the survey and the level of a prepaid cash incentive (\$2 included in an advance letter versus an advance letter only). Random subsamples were sent one of two different questionnaires. The experimental design is shown in figure 2.

Incentive Treatment	Population 1: Likely High Interest in Topic A		Population 2: Likely Medium Interest in Topic A		Population 3: Likely Low Interest in Topic A	
	Survey Topic Treatment		Survey Topic Treatment		Survey Topic Treatment	
	Topic A	Topic B	Topic A	Topic B	Topic A	Topic B
No Incentive						
\$2 Prepaid Incentive						

**Figure 2.** Illustration of two-factor experimental design (topic by incentive) for three different populations.

We used mail questionnaires, partly for reasons of cost and partly to eliminate uncontrolled influence from interviewers on response propensities. Both the focal topic questionnaire (Topic A) and the contrast topic questionnaire (Topic B) included questions about Topic A, such as:

How often do you find yourself thinking about TOPIC A?

- Frequently
- Sometimes
- Occasionally
- Never

Compared to most people you know, how much do you know about TOPIC A?

- Less
- About the same
- More

Other interest and behavior questions were tailored to specific topics.<sup>1</sup>

This design allows the following questions to be addressed:

1. Does the presumed variation in topic interest across the three populations lead to different cooperation rates across the two questionnaires?
2. If so, are the differences diminished when incentives are introduced?
3. If there are differences in cooperation rates as hypothesized in question 1, do *respondents* from those populations with lower overall topic interest report higher interest in the focal topic questionnaire than in the contrast?

1. Questionnaires for all experiments in this article are available from the senior author on request.

In short, the research design requires three steps to test the hypothesized relationships:

1. Demonstration that a group a priori high on topic interest responds to a focal questionnaire about that topic at higher rates than it does to a contrast questionnaire on another topic. (This establishes the basic mechanisms of topic interest driving participation.)
2. Demonstration that respondents from populations with lower overall topic interest report higher interest in the focal topic questionnaire than in the contrast topic questionnaire. (This demonstrates the impact on survey estimates of the phenomenon described in [1] immediately above.)
3. Demonstration that both these effects are diminished when incentives are used. (This demonstrates the tendency for incentives to bring into the respondent pool those less interested in the survey topic.)

We present three sets of experiments. The first two failed to achieve the expected results. The second failure led to a reconsideration of the mechanisms affecting participation in the context of mail questionnaires. The third experiment, designed on the basis of the reconsideration, exhibited most of the hypothesized effects.

## Results

### INITIAL FAILED EFFORTS

Experiment 1 used three samples from members of a military veterans' health services system. The three populations studied were diabetics under treatment, persons at risk for diabetes, and all others. The diabetics did not respond at substantially higher rates to a "Survey of Americans about Diabetes" than to a "Survey of Americans' Quality of Life" (63 percent to 58 percent,  $p < .53$ ). On substantive variables (e.g., perceived importance of diabetes as a national problem) there were some (largely nonsignificant) tendencies for the nondiabetic respondents to report that it was a larger problem on the diabetes questionnaire than on the quality of life questionnaire.

This prompted a careful self-critique. We wondered whether the findings of Groves, Presser, and Dipko (2004) about the effects of topic were limited to the telephone mode; and we wondered whether the salience of topic and sponsor differed between the diabetes survey and the earlier telephone survey, which focused on a variety of topics.

Hence, we first attempted to replicate the Groves, Presser, and Dipko (2004) findings using a mail mode. Groves and his colleagues had shown that a sample of teachers responded at a much higher rate to a telephone survey described as being on "Education and the Schools" than to one described as being on "Issues Facing the Nation" (56 percent versus 39 percent,  $p < .005$ ).

Both surveys had been described as being conducted by the University of Maryland Survey Research Center; the dependent variable was whether the respondent agreed to the interview on first exposure to the topic.

We repeated that experiment using a mailed questionnaire (experiment 2a) but failed to obtain the original result (as in the diabetes experiment above). As shown in figure 4 below, the “issues facing the nation” questionnaire actually yielded a higher response rate among teachers. This caused us to reflect on the differences between the data collection modes:

1. The mail questionnaire, which was limited to 6 pages and 22 questions, had fewer questions than the telephone mode, which had about 20 minutes of questions.
2. The telephone mode respondents were exposed mainly to the introductory descriptions given by interviewers (primarily name, affiliation, and survey purpose); the mail mode provided an advance letter (mentioning the sponsor and purpose) and a questionnaire package (again mentioning the sponsor and purpose, as well as revealing the questions to be answered).
3. The relative salience of the affiliation of the requestor seemed greater in the mail mode. The advance letter and cover letter for the questionnaire package used University of Michigan Survey Research Center stationery and the signature of the project director; envelopes carried the seal of the University of Michigan; and the return envelope for the questionnaire carried a university address. In contrast, the affiliation of the requestor was mentioned in a single phrase in the telephone survey introduction.
4. The mail questionnaire recipient could examine every question in the instrument prior to deciding to respond. However, the actual questions (which were rather general questions about schools and education) may not have been as appealing to a supporter of education and the schools as was the title of the survey itself.

On the basis of these observations, we conducted two follow-up mail-mode experiments that attempted to mimic the participation decision context of the telephone survey. The first examined whether the telephone results could be replicated if a mail sample were given merely the survey title, not the full questionnaire. The second addressed whether the salience of the sponsor overwhelmed other factors manipulated in the design. These experiments are described in the next sections.

#### THE POSTCARD RESPONSE EXPERIMENTS

In the postcard experiments, we used the same sample frames and design as Groves, Presser, and Dipko (2004). We used a stamped return postcard to indicate whether the sample person wanted to receive a questionnaire on a



**Survey on Education and the Schools**

We would like to send you a questionnaire about education and the schools. Please let us know if you would be willing to participate in the survey by completing this postcard and mailing it back to our office.

Sincerely,

Dr. Robert M. Groves  
Director, Survey on Education and the Schools

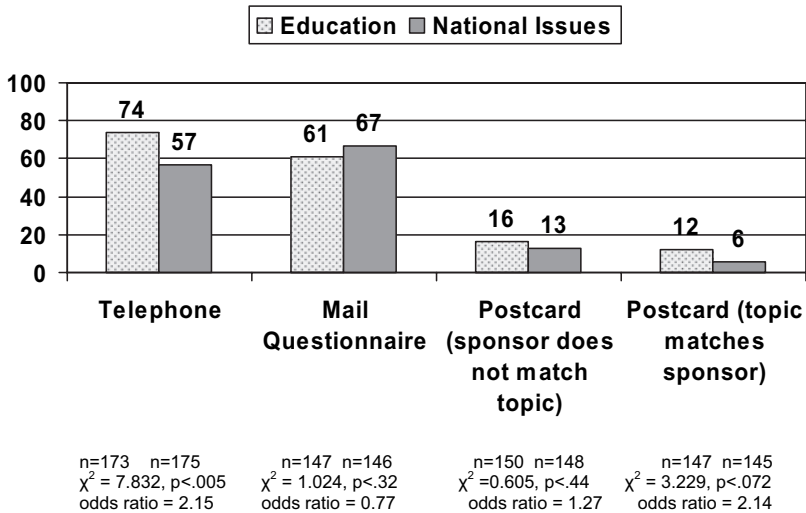
**Yes, I would like to participate. Please mail me a questionnaire about *Education and the Schools*.**

**No, I do not want to participate. Please do not mail me a questionnaire about *Education and the Schools*.**

**Figure 3.** Stamped return postcard for experiment 2c.

given topic (see figure 3). The protocol involved a single mailing of a letter containing the return postcard (with or without an incentive). In the first postcard experiment (experiment 2b), we kept the sponsor as the University of Michigan Survey Research Center (SRC). In the second postcard experiment (experiment 2c), we changed the survey sponsor to be identical to the topic. That is, instead of the request coming from the University of Michigan Survey Research Center, the request came from the “Research Study on Education and the Schools” or the “Research Study on National Issues.” Stationery, envelopes, and return addresses consistently offered that message.

As expected with a single-mailing protocol, the return rate was low, but our purpose was to examine treatment group differences in returns. As shown in figure 4, the first postcard experiment (with SRC sponsorship) obtained a 16 percent “Yes” return for the education and schools topic and a 13 percent “Yes” for the national issues topic ( $p < .44$ ). When the sponsor was the same as the topic (experiment 2c), the education and schools topic generated a 12 percent “Yes” return and the national issues survey generated a 6 percent “Yes” return ( $p < .07$ ). While the base return rates of the postcards are much lower than those of the questionnaires, experiment 2c achieves an odds ratio of cooperation due to topic similar to that achieved in the telephone survey (odds ratio of 2.14 versus 2.15 in the Groves, Presser, Dipko [2004] telephone survey).



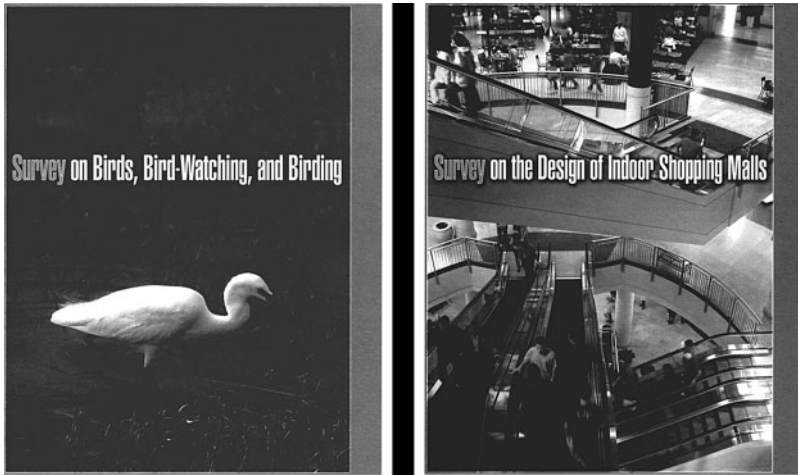
**Figure 4.** Cooperation rate for Groves, Presser, and Dipko (2004) and response rates for mail experiments in current study, by recruitment protocol of mail experiment and topic of survey, teacher sample.

These two experiments move the results in the direction of those found on the earlier telephone survey and have potentially offered us lessons in how the participation decision varies across modes. For example, an important question for a future meta-analysis is whether sponsorship effects are larger in mail questionnaire surveys than in interviewer-assisted surveys. The fact that the direction and magnitude of the topic effects appeared to be a function of the sponsor effects reminds us that both influences are at work when the sample person is exposed to both.

### EXPERIMENT 3

Experiment 3 incorporated some of our knowledge from the failed diabetics experiment. First, we chose a survey topic that a priori would be positively valued by the target population. Second, we sought input on other likes and dislikes of the target population. We also incorporated what we had learned from the postcard experiment by varying the sponsorship of the survey.

The population chosen for study was members of the American Birding Association (ABA), a nonprofit membership organization devoted to the promotion of birders (i.e., bird watchers) and bird habitats. We held a focus group of staff and members of the ABA and learned that passive time use and indoor



**Figure 5.** Covers for two questionnaires used in experiment 3.

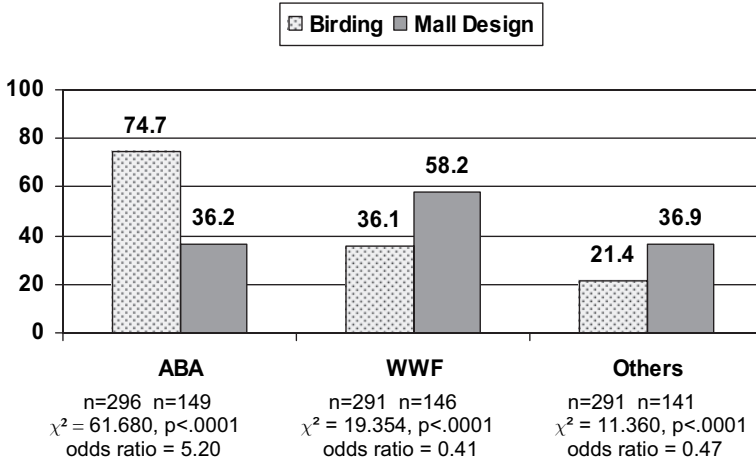
activities were of little interest to ABA members. One focus group member volunteered that indoor shopping malls probably epitomized the kind of topic that birders would find uninteresting.

We designed two questionnaires: the “Survey on Birds, Bird-Watching, and Birding” and the “Survey on the Design of Indoor Shopping Malls.” Three populations were sampled: ABA members; donors to the World Wildlife Fund (WWF) (thought to be interested in environmental issues but not specifically birds and birding); and other adults (from a commercial list vendor). Half the cases were offered a \$2 incentive in an advance mailing. Half the birding questionnaires were portrayed as being conducted by the Michigan Survey Research Center; half, by the ABA.

The protocol involved (a) an advance mailing, with or without an incentive, (b) the questionnaire mailing, including a cover letter, a questionnaire, and a stamped return envelope, (c) a reminder postcard, and (d) if needed, a follow-up second questionnaire.

The questionnaire covers appear in figure 5.

*Response Rate Results of Experiment 3.* First, we examine the response rate results of the experiment, for the groups without incentives. Figure 6 shows much higher relative response rates for the birding questionnaire versus the shopping mall design questionnaire among ABA members (28.5 percentage point increase). The other two groups have *lower* response rates for the birding questionnaire than for the mall design questionnaire. A logistic regression model for the no-incentive condition shows that the interaction effect of topic



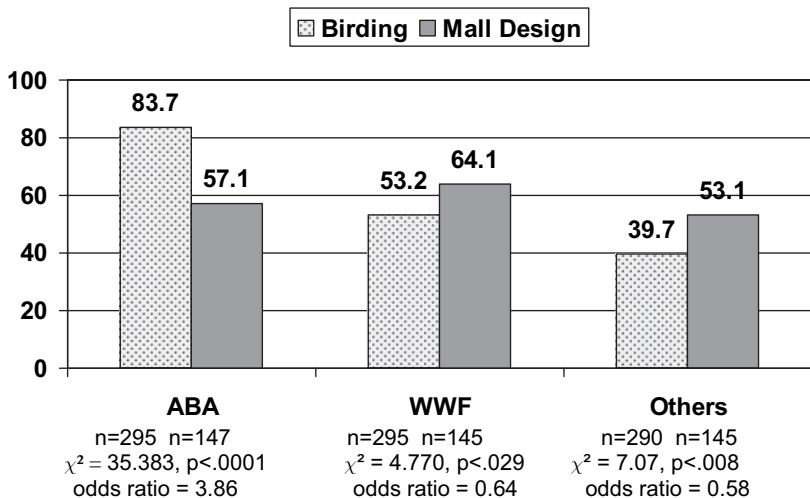
**Figure 6.** Experiment 3 response rates for no-incentive condition, by topic by population.

and population is significant (the ABA members show a distinctively large preference for the birding questionnaire,  $p < .0001$ ; the WWF and the “Other” groups are not distinguishable, given the sample size); the population that we expected to be interested in birding and uninterested in malls returned the birding questionnaire at a much higher rate than the mall questionnaire compared with the other two populations.

We hypothesized that the effect of topic interest would be diminished in surveys in which monetary incentives were offered to all sample persons. Figure 7 shows that all three groups’ odds ratios shrink toward 1.0; topic effects on response rates tend to be smaller with prepaid incentives for all three groups. For the ABA population, the interaction of topic and incentive on response rate achieves a  $p < .35$ ; for the WWF,  $p < .13$ , and for “Others,”  $p < .47$ . Thus, the dampening effects of incentives do not achieve conventional levels of statistical significance.

*Impact of Response Participation Decisions on Survey Variables.* The two questionnaires contained a battery of identical questions. We can use these questions to test the hypothesis that estimates of interest in birding will be higher for respondents who chose to answer the birding questionnaire than for respondents to the shopping mall design questionnaire.

A key item asked whether the respondent engaged in birding or bird-watching activities in the last 12 months. The marked tendencies for the ABA sample members to prefer the birding questionnaire to the shopping mall design questionnaire confirmed that those interested in birding were, as expected, less interested in the shopping mall questionnaire. Thus, we expected the



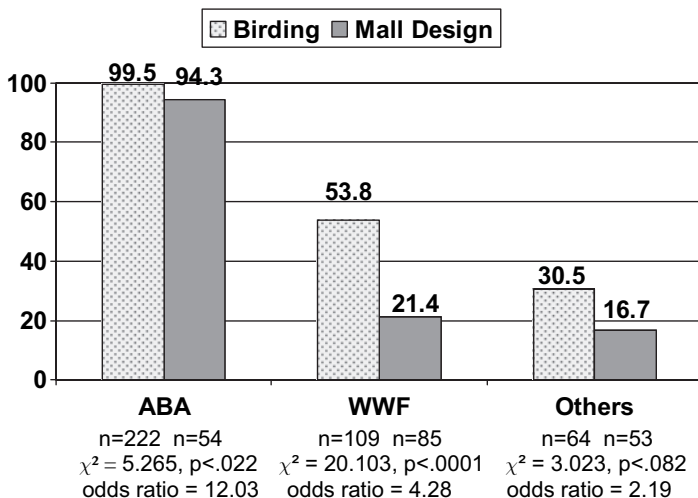
**Figure 7.** Experiment 3 response rates for incentive condition, by topic by population.

WWF and “Other” respondents to the birding questionnaire to display more interest in birding than the corresponding respondents to the shopping mall questionnaire.

Figure 8 shows that the estimated percentages of people participating in birding activities are greater from the birding survey than the shopping mall design survey in the no-incentive condition. The differences exist for all three groups beyond the levels expected from sampling error alone. The ABA respondents show very high levels of birding prevalence, 99 percent with the birding questionnaire and 94 percent with the shopping mall design questionnaire. Larger percentage differences occur in the WWF and the “Other” respondent groups, 32 percentage points and 13 percentage points, respectively. The ABA and WWF birding prevalence estimates differ at rates beyond traditional significance levels ( $p < .02$  and  $p < .0001$ , respectively). The chi-square test for the “Other” population yields a  $p < .08$ .

To summarize thus far, those persons expected a priori to be interested in birding respond to a birding topic questionnaire at higher rates than to a shopping mall design questionnaire. Consequently, prevalence estimates of birding are much higher in the birding questionnaire than in the shopping mall design questionnaire for the two target populations whose members vary in their interest in birding.

We hypothesized that the use of a prepaid incentive would bring into the respondent pool persons who were uninterested in the stated topic of the

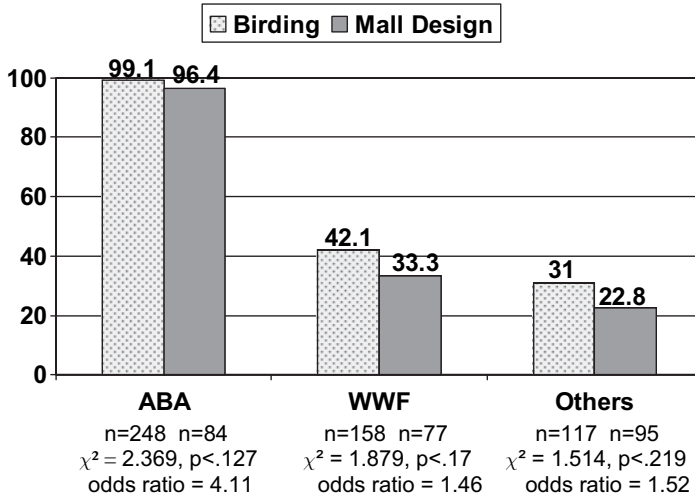


**Figure 8.** Percentage reporting participating in birding/bird-watching in past 12 months, experiment 3 respondents, by questionnaire by population (no-incentive condition).

questionnaire. In this experiment, we expected that those interested in birding—whether they were ABA members or not—would be more likely to respond to the birding questionnaire than to the shopping mall questionnaire. Hence, the odds ratio of the response rates between the two topics should decline when incentives are prepaid. That is, with monetary incentives more of those uninterested in birding would respond to the birding questionnaire and more of the sample uninterested in shopping malls (among whom are those interested in birding) would respond to the mall questionnaire.

Figure 9 shows the same statistics as those in figure 8 for the treatment groups that were given an incentive. As expected, all three of the odds ratios shrink toward 1.0 in figure 9 relative to figure 8. All of the chi-square tests of differences fall above levels expected from sampling alone. That is, there are no measurable differences in birding prevalence estimates by questionnaire topic when incentives are prepaid to the sample. Incentives bring into the respondent pool persons who counteract the greater tendency for those interested in the topic to cooperate. However, when formal tests are conducted to measure the change in odds ratios due to the use of incentives, only the WWF population shows a reliable decline ( $p < .014$ ).

We can summarize our findings using table 1, which presents the odds ratios of topic effects on two dependent variables—the response rate and



**Figure 9.** Percentage reporting participating in birding/bird-watching in past 12 months, experiment 3 respondents, by questionnaire by population (incentive condition).

**Table 1.** Odds Ratios for Topic Effects on Response Propensity and Prevalence of Birding by Population by Use of Prepaid Incentives

Population	Odds Ratios for Topic Effect					
	Response Rate			Prevalence of Birding		
	Without Incentives	With Incentives	$p^a$	Without Incentives	With Incentives	$p^a$
ABA						
Members	5.20	3.86	.35	12.03	4.11	.49
WWF						
Donors	0.41	0.64	.13	4.28	1.46	.01
Others	0.47	0.58	.47	2.19	1.52	.53

<sup>a</sup> *P*-value on test of differences of odds-ratios, using logistic regression

reported birding activities in the past 12 months. The odds ratios moving toward 1.0 when incentives are used imply that the topic effects on the tendency to respond and on birding prevalence statistics are diminished with incentives. The *p*-value columns show that the decline in topic effects due to

incentives is not large enough to be detected reliably at these sample sizes, with one exception.

In short, the experiment yielded both the hypothesized effect of topic interest on response rates and the hypothesized increase in the estimated popularity of activities related to the topic. The incentives do not, however, dampen the effects of topic interest at a level that can be reliably detected given the sample sizes of the experiment.

## Summary and Conclusions

We began this research hypothesizing that sample persons may use the topic of the survey as one factor in their decision about participation—those more interested in the topic should give more positive attention to the request. This mechanism should lead to higher estimates of knowledge, favorable attitudes, and activities relevant to the topic.

We further hypothesized that this effect could be suppressed by other features of the recruitment protocol. Specifically, those uninterested in the topic of the survey could be induced to participate if they were given pre-paid incentives.

From comparisons with our prior experiments (Groves, Presser, and Dipko 2004), we learned that topic effects on mail questionnaires differed from those produced by telephone surveys. We hypothesized that the relative salience of the survey topic might be diminished in mail questionnaires because of repeated communication of the survey's sponsorship in written materials. The chance to look at the questions themselves might also alter the impact of the survey topic on the participation decision. Some of these speculations have not yet been tested experimentally but are consistent with the follow-up experiments on the study of education and the schools.

A further refinement flowed from the failure to support some of the key hypotheses in the diabetes experiment. Because diabetic patients' response rate to the diabetes survey was insignificantly higher than their response rate to the quality of life survey, we believe that the hypothesized role of topic interest needs to be refined. Participation in the survey is apparently not triggered by topic interest or relevance to the self-image alone, but by the likelihood that thinking about the topic will be rewarding to the respondent. These rewards may be pleasant memories, psychic benefits of demonstrating knowledge in an area one considers important, or the gratification of knowing that the survey may increase society's attention to an issue related to key self-interests. When the topic of the survey is relevant to the sample person but generates negative thoughts, unpleasant memories, or reminders of embarrassing personal failings, then the topic may suppress participation despite its personal relevance. This was one post hoc interpretation of the diabetes experiment results.



When the mechanisms perform as expected (as in the birding experiment), they lead to exaggerated population estimates of the prevalence of interests and activities related to the survey topic. Incentives appear to diminish these biasing tendencies somewhat.

There is one caution about these interpretations of our findings. We have interpreted the higher estimated prevalence of birding activity from a birding survey versus a shopping mall survey as evidence of nonresponse bias. However, it is also possible that the higher estimates are produced by differences in measurement error. That is, respondents to the birding questionnaire may have overreported their interest in birding in order to conform to the perceived intent of the survey. These possible measurement error mechanisms are in the same direction as those of the nonresponse bias effects we have hypothesized. The results of the experiments thus confound the impact of nonresponse and measurement error.

The practical implications of these experimental results may be large. When survey introductions and survey materials emphasize the topic of the survey, they risk stimulating participation among persons whose self-interests can be served by responding and depressing participation among those who perceive no such interests. When the survey variables are correlated with those interests, differential nonresponse bias can result, varying by the stated topic of the survey. Such risks of nonresponse bias are ubiquitous. For example, many labor force surveys around the world describe the survey as a crucial tool to measure unemployment; election survey introductions describe their importance in informing society about political attitudes and actions. When the topic is made very salient in the recruitment protocol, the effects we have demonstrated experimentally may occur in practice. Two repairs seem possible. In interviewer-assisted surveys the topic of the survey can be made less salient in the introduction; and in all modes, incentives can be offered to heighten the extrinsic benefits of participating to those uninterested in the survey topic itself.

## References

- Baumgartner, Robert, Patricia Rathbun, Keven Boyle, Michael Welsh, and Drew Laughland. 1998. "The Effect of Prepaid Monetary Incentives on Mail Survey Response Rates and Response Quality." Paper presented at the annual meeting of the American Association of Public Opinion Research, St. Louis, MO.
- Bethlehem, Jelke. 2002. "Weighting Nonresponse Adjustments Based on Auxiliary Information." In *Survey Nonresponse*, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, pp. 275–88. New York: Wiley.
- Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. "The Effects of Response Rate Changes on the Index of Consumer Sentiment." *Public Opinion Quarterly* 64:413–28.
- de Leeuw, Edith, and Wim de Heer. 2002. "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison." In *Survey Nonresponse*, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, pp. 41–54. New York: Wiley.
- Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Error in Household Surveys." *Public Opinion Quarterly* 70:646–75.

- Groves, Robert, Stanley Presser, and Sarah Dipko. 2004. "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly* 68:2–31.
- Groves, Robert M., Eleanor Singer, and Amy Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and an Illustration." *Public Opinion Quarterly* 64:288–308.
- Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves, and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64:125–48.
- Little, Roderick, and Donald Rubin. 2002. *Statistical Analysis with Missing Data*. 2d ed. New York: Wiley.
- Merkle, Daniel, and Murray Edelman. 2002. "Nonresponse in Exit Polls: A Comprehensive Analysis." In *Survey Nonresponse*, ed. Robert M. Groves, Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, pp. 243–58. New York: Wiley.