

Contact-Level Influences on Cooperation in Face-to-Face Surveys

Robert M. Groves¹ and Mick P. Couper¹

Many postsurvey adjustment procedures for unit nonresponse in surveys are based on implicit models of survey participation, attempting to estimate the response propensity of sample persons. The statistical properties of adjusted survey estimates depend on the specification of such models. To motivate such specification, a theory of face-to-face survey participation is reviewed and empirical analyses presented for one level of influences on response propensity: those influences on the decision to participate that arise during the interaction between a survey interviewer and a householder. The data suggest that traditional socio-demographic correlates of survey response rates do not predict well the outcome of individual contacts between interviewers and householders, except for the first contact. Observable behaviors of the householders, documented by interviewers after the interaction, however, are informative about the likely behavior in future contacts and about the final disposition of the sample case. Interviewers who change the content of their interactions with householders, over successive contacts, tend to achieve somewhat higher cooperation rates.

Key words: Survey nonresponse; response propensity models; householder-interviewer interaction.

1. Introduction

Despite ingenious and vigorous efforts at persuading sample householders to grant survey requests, interviewers still fail to obtain universal participation. Reluctant respondents lead to increased survey costs and errors simultaneously, and researchers committed to high quality surveys must choose between repeated conversion efforts bordering on harassment or attempting to compensate for nonresponse error through postsurvey adjustment.

Many postsurvey adjustment procedures for unit nonresponse are based on weighting schemes that identify groups with similar characteristics on key survey variables and with similar likelihoods of inclusion in the respondent database (see Kalton 1983; Lessler and Kalsbeek 1992; Little 1986). Most of the literature on postsurvey adjustment examines statistical properties of estimators given some adjustment model. However, little of the literature focuses on how to build the adjustment model: What is the appropriate specification of a response propensity equation; what are

¹ Survey Research Center, University of Michigan, P.O. Box 1248, Ann Arbor, MI, 48106, U.S.A.

Acknowledgments: An earlier version of this article was presented at the 1994 meetings of the American Association for Public Opinion Research. Data collection for this study was supported in part by grant #MH46376 from the U.S. Alcohol, Drug Abuse, and Mental Health Administration (Ronald Kessler, Principal Investigator), and the Survey Research Center, University of Michigan.

the causes of survey participation; why do people respond to surveys? As with any model-based statistical procedure, however, the desirable features of the estimation are dependent on proper specification of the model.

We believe that correct specification of an adjustment model must begin with a systematic theory of survey participation. We have been conceptualizing and testing such a model for face-to-face surveys (see Groves, Cialdini, and Couper 1992; Groves and Couper 1993a; Groves and Couper 1994). It posits that judgments by persons about participating in a survey are largely made during their brief interactions with survey interviewers. These judgments, although quickly made, are not based solely on the arguments, cues, and information presented in the brief interactions. They are based, in addition, on a large set of memories and prior experiences, knowledge, beliefs, and attitudes that are evoked during the interaction. Some of these may be directly relevant to the topic of the survey, the sponsoring organization of the survey, or the stated goals of the survey information (and thus are potential sources of nonignorable nonresponse). Others may be strong positive or negative reactions to very peripheral attributes of the survey request (e.g., interpretations of an interviewer's smile).

Our theoretical perspective includes various sets of influences that form the foundation of the householder's reaction to a survey request (see Groves, Cialdini, and Couper 1992; Groves and Couper 1995). These include factors such as the social environment or household(er) characteristics that are largely out of the control of the research design, and factors that are determined in large part by the research design (survey design characteristics and interviewer attributes).

Most postsurvey adjustment models utilize indicators from the social environmental level (e.g., urban/rural status, type of neighborhood) or household level (e.g., household size, type). These are used not because of theoretical assertions that they are causal factors for survey participation, but because they are generally the only measures available both on respondents and nonrespondents. Few have utilized attributes of the interviewer or characteristics of the interaction, yet the theoretical perspective we take asserts that these influences can be powerful. This article focuses on the interaction between the householder and the survey interviewer prior to the householder making a decision to participate in the survey.

1.1. The initial interaction from the householder's perspective

During the initial moments of the interaction, we believe the householder is actively engaged in an effort to comprehend the intent of the interviewer's visit (see Figure 1). The concepts inherent in cognitive script theory are useful here (see Abelson 1981). These would assert that the words, behavior and physical appearance of the interviewer will be used to identify possible "stories" or scripts that are potential explanations of the interviewer's purpose. These scripts are used as heuristic devices that quickly help the householder anticipate next steps in the interaction and the ultimate goals of the visitor. Past experience with similar situations will dictate behavior (e.g., if this is a sales encounter, then be reluctant or uninterested).

On rare occasions no likely script will be evoked by the interviewer's call – the householder has no clue about the interviewer's intent. Then the opportunity costs

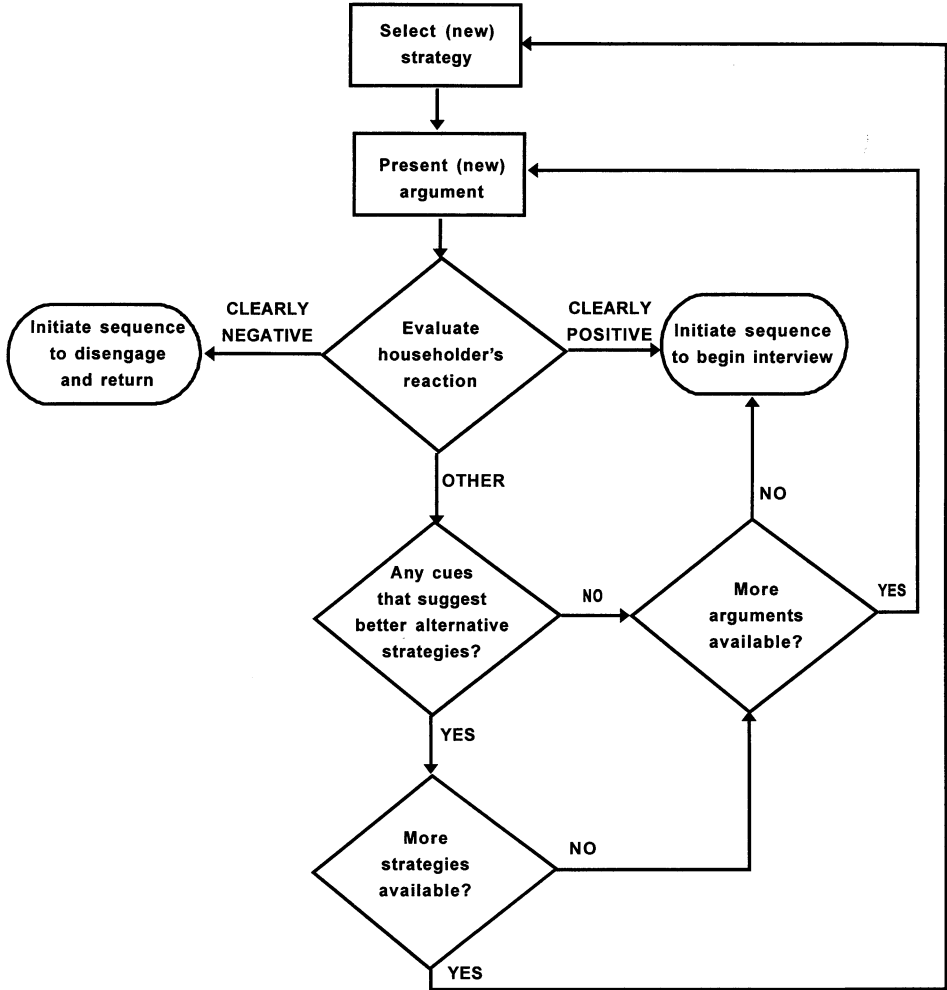


Fig. 1. Interviewer strategies of tailoring and maintaining interaction

of the interaction to the householder are important predictors of the householder's response. If there are few desirable alternative uses of time, the householder will attempt to clarify the purpose of the visit. If not, the householder will probably attempt to terminate the interaction.

Sometimes words of the interviewer evoke multiple scripts at once (e.g., "from the University of Michigan," "a study for the U.S. Bureau of Labor Statistics"). Because of the cognitive burdens of assessing alternative scripts, we think the householder places a premium on rapid identification of which script is appropriate. How detailed an assessment is conducted is probably a function of what else the householder had to do at the time. If the opportunity costs for continuing the conversation are low, the householder will question the interviewer about alternative scripts in order to eliminate inappropriate ones; for example, "What are you selling?" Such attempts might be especially appropriate for candidate scripts that lead to clear negative behavioral

outcomes. In the above example, once it is known what is being sold, the decision to continue or disengage can be made quickly.

When seeking more information about what script is appropriate is judged too burdensome itself, the householder merely chooses a script and delivers an answer based on it (e.g., “I don’t want any,” or “Sure, come on in”). A specific example of such a high burden circumstance is when the householder is engaged in some other activity at the time of the contact. Here the cognitive burdens of determining the intent of the interviewer are exacerbated by the opportunity costs of doing so, the inability to continue the prior activities. Sometimes this generates responses such as “I’m too busy” or “Could you call later?” or a host of other similar responses. We believe, however, that the same statements are also convenient masks for refusals to participate.

When key phrases of the interviewer point only to unattractive scripts, the householder quickly attempts to disengage. Sometimes all potentially applicable scripts lead to the same behavioral guideline; for example “all purposes of a male stranger calling on the household must be harmful to the householder.” This type of reaction might also be characterized as a rule of behavior adopted by the householder that is uniformly exercised. Such rules come closest to the survey researchers’ notions of “hard core” refusers.

Earlier analyses of the householder-interviewer interaction support the script theoretical view of householder behavior illustrated in Figure 1. In past analyses (Groves and Couper 1994), we have learned that sometimes householders ask about the purpose of the survey, about whether the interviewer is selling anything, about how much time the interviewer needed, or about sample selection. The most common householder comment is to say that they are “too busy,” “not interested,” or “don’t know about the survey topic.” Finally, householders ask many more questions in the first contact with interviewers than in later contacts on the same survey.

It appears that householders approach these encounters as they would those of other strangers who may approach their doors. In urban areas, for example, for units with some barriers to public access (e.g., bars on the windows, “no solicitors” signs), this leads to more negative reactions being voiced. Until and unless the interviewer communicates successfully the nature of the survey request, this general “unwanted stranger script” may dominate the householder’s interpretation of the encounter.

1.2. Interaction from the interviewer’s perspective

In order to construct our theoretical perspective from the interviewer’s viewpoint, we held discussions with expert panels – focus groups of survey interviewers. These included groups of interviewers and groups of supervisors from the University of Michigan Survey Research Center and the U.S. Census Bureau, run over a three-year period. They shaped our viewpoint that the interviewers are in a heightened state of alert as they approach the sample unit, the door answerer, and finally, the chosen respondent. We use two concepts to describe this process – tailoring and maintaining interaction.

Experienced interviewers often reported that they adapt their approach to characteristics of the sample unit. In some sense, expert interviewers have access to a large repertoire of cues, phrases, or descriptors corresponding to the survey request. Which

statement they use to begin the conversation is the result of observations about the housing unit, the neighborhood, and immediate reactions upon first contact with the person who answers the door. The reaction of the householder to the first statement dictates the choice of the interviewer's second statement. With this perspective, all features of the communication are relevant – not only the words used by the interviewer, but the inflection, volume, pacing (see Oksenberg, Coleman, and Cannell 1986), as well as physical movements and demeanor of the interviewer.

From the focus groups we found that some interviewers are aware of their “tailoring” behavior: “I give the introduction and listen to whay they say. I then respond to them on an individual basis, according to their responses. Almost all responses are a little different, and you need an ability to intuitively understand what they are saying.” Or, “I use different techniques depending on the age of the respondent, my initial impression of him or her, the neighborhood, etc.” Or, “From all past interviewing experience, I have found that sizing up a respondent immediately and being able to adjust just as quickly to the situation never fails to get their cooperation, in short being able to put yourself at their level be it intellectual or street wise is a must in this business. . . .”

Tailoring sometimes cannot occur within a single contact. Many contacts are very brief and give the interviewer little opportunity to respond to cues obtained from the potential respondent. Tailoring may take place over a number of contacts with a single household, with the interviewer using the knowledge he/she has gained in each successive visit to that household. Tailoring may also occur across sample households. The more an interviewer learns about what is effective and what is not with various types of potential respondents encountered, the more effectively requests for participation can be directed at similar others. This implies that interviewer tailoring evolves with experience. Not only have experienced interviewers acquired a wider repertoire of persuasion techniques, but they are also better able to select the most appropriate approach for each situation.

1.3. Maintaining interaction

Another finding of the focus groups was the premium placed by interviewers on avoiding a termination of contact with the householder during a visit. The introductory contact of the interviewer and householder is a small conversation. It begins with the self-identification of the interviewer, contains some descriptive matter about the survey request, and ends with the initiation of the questioning, a delay decision, or the denial of permission to continue.

The implicit aim of survey administration in probability samples is maximizing the probability of obtaining an interview from each sample unit. The number of contacts required to obtain cooperation on each case is of secondary concern. Given this, interviewers often apply the tailoring over several turns in the contact conversation. How to tailor the appeal to the householder is increasingly revealed as the conversation continues. Hence, the odds of success are increased with the continuation of the conversation. Thus, the interviewer does not maximize the likelihood of obtaining a “yes” answer in any given contact, but minimizes the likelihood of a “no” answer over repeated turntaking in the contact.

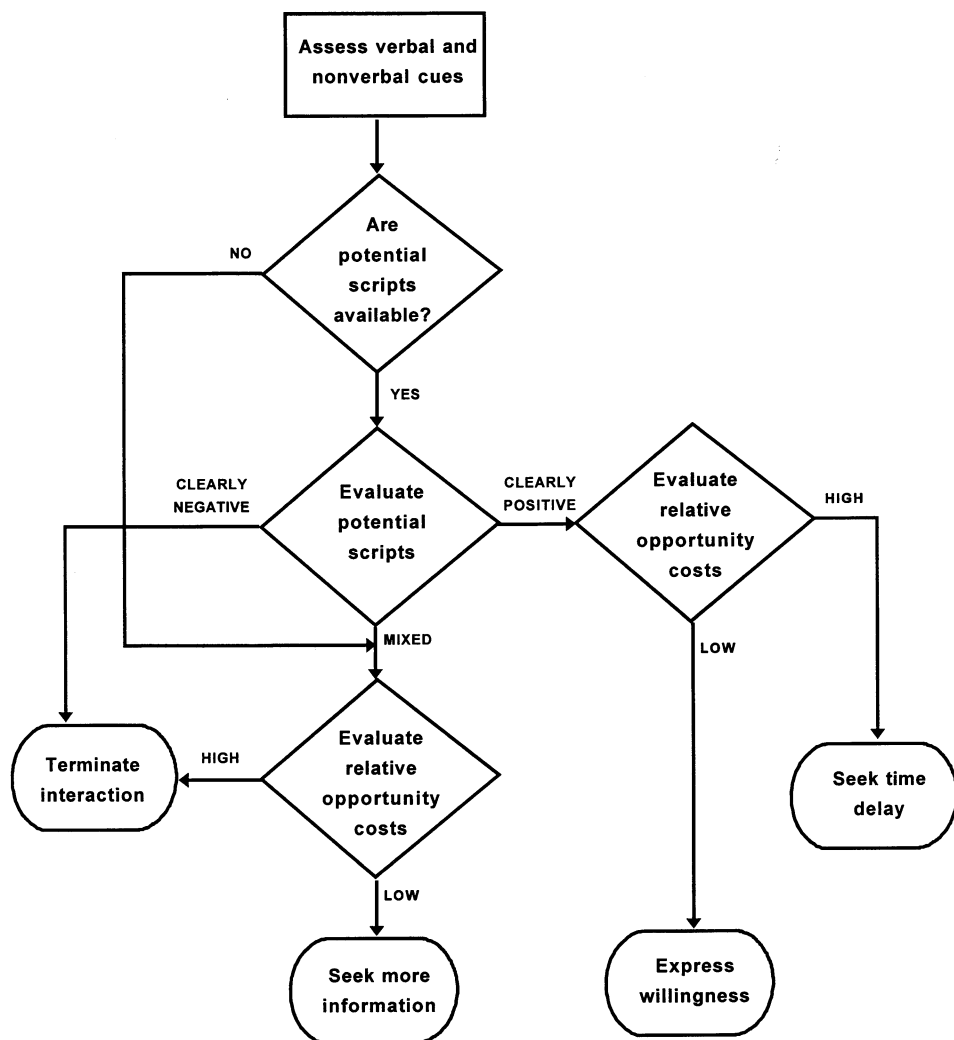


Fig. 2. Householder strategies for evaluating survey request

We believe the techniques of tailoring and maintaining interaction are used in combination. Maintaining interaction is the means to achieve maximum benefits from tailoring, for the longer the conversation is in progress, the more cues the interviewer will be able to obtain from the householder. However, maintaining interaction is also a compliance-promoting technique in itself, invoking the commitment principle (Groves, Cialdini, and Couper 1992) as well as more general norms of social interaction. That is, as the length of the interaction grows, it becomes more difficult for one actor to summarily dismiss the other.

Figure 2 is an illustration of these two interviewer strategies at work. We distinguish between the use of a general compliance-gaining strategy (e.g., utilizing the principle of authority) and a number of different (verbal and nonverbal) arguments or tactics within each strategy (e.g., displaying the ID badge prominently, emphasizing the sponsor of the survey, etc.). The successful application of tailoring depends on the ability of the

interviewer to evaluate the reaction of the householder to his/her presence, and the effectiveness of the arguments presented. Note that the interviewer's initial goal is to maintain interaction (avoiding pushing for the interview) as long as the potential respondent's reaction remains neutral or noncommittal. An interviewer will continue to present different arguments until either the householder is clearly receptive to an interview request, or there are no more arguments to present. For inexperienced interviewers the latter may occur before the former, forcing the interviewer to (prematurely in some cases) initiate the interview request.

Successful tailoring thus requires that a number of conditions be met. First, the interviewer must have a wide variety of techniques or strategies at his/her disposal. Expert interviewers tend to have access to a large repertoire of cues, phrases, or descriptors corresponding to the survey request. Only then can the appropriate statement for a particular situation be selected. Second, the interviewer must be a good reader of verbal and nonverbal cues from the householder and surroundings. Third, the interviewer must apply the appropriate strategy/technique in response to the cues received. Interviewers must be able to make fast and accurate judgments about the particular script reflected in the householder's initial response to the request, and react accordingly.

At times, whether through low motivation to seek alternative scripts, or because of a misjudgement regarding the intent of the interviewer, a script is evoked that is judged inappropriate (from the interviewer's perspective) to the survey request. In such cases, interviewers are often quick to redress the situation by distinguishing their requests from other similar situations. More often, though, a script may be used whose meaning is ambiguous to the interviewer. For example, "I'm not interested" may mean "I think you're selling something and I'm not interested in buying," or "I know this is a survey request, but I'm not interested in participating." To be successful in tailoring, an interviewer must be a good reader of such cues, or must find ways of eliciting additional cues that reveal the intentions of the householder.

Finally, the interaction (whether within a single contact or across multiple contacts) must be of sufficient duration for tailoring to be applied. There is some support from training procedures that the "maintaining interaction" model is correct. First, interviewers are typically warned against unintentionally leading the householder into a quick refusal. If the person appears rushed, preoccupied by some activity in the household (e.g., fighting among children), the interviewer should seek another time to contact the unit. A common complaint concerning inexperienced interviewers is that they create many "soft-refusals" (i.e., cases easily converted by an experienced interviewer) by pressing the householder into a decision prematurely. Unfortunately, only rarely do interviewer recruits receive training in the multi-turn repartee inherent in maximizing the odds of a "yes" over all contacts. Instead, they are trained in stock descriptors of the survey leading to an interview request rather quickly.

2. Plan of this Article

This article presents analyses of interviewer observations about what occurred when they contacted sample households. These observations were taken both on responding and nonresponding households. Although the initial analyses of these

data (Groves and Couper 1994) yielded results that are consistent with the theoretical perspective, the real focus of the theory has not yet been tested – do these data reveal systematic influences on the likelihood of participation in the survey? From a purely predictive perspective, does the nature of the interaction provide new information about the likelihood of consent to the interview, separate from the socio-demographic variables commonly used or the results of intermediate contacts? From a theoretical perspective, is there evidence of support for the notion that tailoring of interviewer behavior to householder characteristics is related to higher response propensities?

We address questions like these in two steps. First, we examine the results of individual contacts on sample households and the influence of householder and interactional characteristics on contact-level outcomes. In that portion of the article we dissect the decision process into successive contacts on sample households, to examine whether the theoretical perspective above is informative about the likelihood of a refusal on a specific contact. Second, we go to the level of the sample case and examine whether the nature of the interaction between householder and interviewer, as reflected in these observations, is informative about the likelihood that the household will eventually yield an interview.

3. Data Resources

The data examined come from the National Survey of Health and Stress (NSHS; also called the National Comorbidity Survey), a 1991 face-to-face survey of noninstitutionalized adults age 15–54 in the coterminous 48 United States (see Kessler et al. 1994), based on a self-weighting multistage area probability sample design (Heeringa 1992), conducted by the University of Michigan Survey Research Center.

Within each sample housing unit, the interviewer obtained a complete listing of household members. If no eligible adults age 15–54 were present in the housing unit, no interview was requested (these cases are omitted from the following analyses). If one or more eligible adults were present, the interviewer proceeded to select a random respondent for the interview, which lasted between one and two hours on average.

Attempts were made to conduct face-to-face interviews with all respondents; however, some contact attempts were made by telephone to set appointments, and follow-up of some hard-to-reach respondents was undertaken from a centralized telephone facility. In this article we excluded all contacts made from the centralized facility, as our primary interest was in the behavior of field interviewers. However, including the telephone facility contacts in the analyses produces essentially the same results.

About 12,000 housing units were sampled for this study, of which 9,863 yielded households with one or more eligible persons. The final disposition of these 9,863 eligible sample units is as follows:

	<i>Number</i>	<i>Percentage (%)</i>
Interview	8,451	85.7
Refusal	1,026	10.4
Noncontact	152	1.5
Other noninterview	234	2.4
Total	9,863	100.0

As our analytic interest in this article is on the nature of interaction between interviewers and householders, we dropped those cases that were never contacted during the course of the field period. Also excluded from these analyses are those not interviewed due to language problems, illness, death or other inability to be interviewed. Further pursuit of these cases after the initial determination is not normally undertaken. This leaves us with 9,477 sample households who were contacted and were capable of being interviewed. The cooperation rate (defined as interviews / (interviews + refusals)) for this group is 89.2%. Additional cases are lost in the analyses described here, either through the exclusion of telephone facility cases, or through deletion of item missing data.

As part of the NSHS a variety of data collection efforts were undertaken to gain further understanding of the influences on survey participation. These included the following:

- a. Contact description questions, described in more detail below;
- b. Neighborhood observations, one-time interviewer reports of the quality of the housing unit and surrounding area, including potential barriers to contact (e.g., bars on windows, locked lobby, “no trespassing” sign), indicators of neighborhood cohesion, and crime potential;
- c. Call records, information on day and time of each call, mode of contact, and disposition of the call; and
- d. Interviewer questionnaires, measuring interviewer attitudes, expectations, behavior and experience regarding survey participation. The interviewer confidence indicator (measured as agreement with the statement “With enough effort I can convince almost anyone to participate”) is from this questionnaire.

The data from these sources were merged with other coversheet and administrative data to form a file with information on interviewers, cases (sample households), contacts and calls. The final merged file contains records for 54,586 calls and 26,888 contacts for the 9,863 cases. Standard error estimates and statistical tests presented here are calculated using Taylor Series approximation, reflecting stratification and clustering of the survey design (using SUDAAN; see Shah, Barnwell, Hunt, and LaVange 1993). Since the majority of interviewers were assigned cases in only one primary sampling unit, some portion of the between-interviewer component of variance is also reflected in the resulting variance estimates.

3.1. Contact description questions

To document specific aspects of the conversations between interviewers and householders, the coversheet booklet for NSHS contained a series of questions to be answered by interviewers after each contact was completed. With the knowledge that only a limited number of items could be included in the contact description questions, we focused on testing theoretically-based propositions rather than attempting to cover all possible behaviors in the introductory conversation. Focus groups with interviewers from various organizations were used to develop and refine the measures used. The contact description questions were also thoroughly pretested as part of the

main study pretests. Pretest interviewers were debriefed and further changes made to the instrument. Interviewers on the NSHS were trained in the use of the contact description questions prior to production interviewing.

Interviewers were instructed to complete these items for *each* contact with members of a sample household with whom they had a conversation regarding the survey. The introductory conversation refers to the time between when the interviewer introduced him/herself to the householder and either the time they proceeded with the interview or, failing that, the time the interviewer left the housing unit. Characteristics of the householder refer to the person with whom the interviewer spoke primarily about the survey during that contact.

Interviewers were instructed to complete the contact description instrument as soon as possible after leaving each sample household. Sufficient copies of the instrument were provided for interviewers to record information on all contacts with householders.

3.2. *Limitations of the data*

The ideal data set for our analytic purposes would be measured at the level of each turn of the conversation between interviewers and householders. This would enable observation of how response propensities change based on what interviewers say in response to each householder behavior. Key to the tailoring hypothesis is the temporal order of householder and interviewer behavior. The contact description data we employ do not reveal the temporal order of utterances by the two actors.

Another attribute of the ideal data set would be an interpenetrated assignment of interviewers to sample cases. This was not part of the NSHS design. In the models to be presented, we attempt controls on those attributes of the sample households that may be related to the dependent variables, prior to examining interviewer-level effects.

The third limitation of the contact description data is the restriction to a specific set of behaviors that were identified prior to the study. Only this set, derived from the theoretical interests of the investigators, is documented. Some important behaviors by either householders or interviewers may have been missed.

In a coordinated study with European collaborators, McCrossan (1993) mounted a study that had trainers (supervisors) observing interviewers complete the same contact form as did interviewers. Supervisors completed the form during the interaction of the interviewer with the householder; interviewers completed the form, independently of the supervisor, after the contact was completed. Because of some differences in training, measures of agreement between trainers and interviewers might be viewed as lower limits of accuracy of the NSHS data.

We expect that individual statements by householders may produce classification problems. For example, "I'm not interested because I don't know anything about (the survey topic)," can be classified both as "I'm not interested" excuse and "I don't know anything about the survey topic." The analysis that follows focuses most attention on a collapsing of individual statements into thematic groups. We categorized householder actions into three groups:

- a. Negative statements include "I don't know anything about the survey topic," "I'm not interested," "Surveys are a waste of taxpayers' money";

- b. Time-delay statements include "I'm too busy," "Let me think about it"; and
- c. Questions.

For negative statements, the observers and interviewers were in agreement 93% of the time. That is, 93% of all contacts had interviewers and trainers agreeing on whether one or more negative statements were made. For time delay statements, they were in agreement 85% of the time. For questions, the agreement rate was 77%.

The mismatch between interviewers and trainers is not merely a result of interviewers recording fewer householder behaviors than trainers. Trainers record negative comments in 68% of the contacts that interviewers record negative comments, agree with interviewer notes of time delay statements 62% of the time and, with questions 64% of the time. Overall, interviewers record the same number of negative comments as trainers, fewer time delay statements than trainers, but more questions than trainers.

A second evaluative step was comparison to counts of behaviors by householders coded from audio tapes of 149 first contact doorstep introductions in Britain by Morton-Williams (1993). In both the NSHS and the British data, the most frequent first householder behavior is the question, "What is the survey about?" (Morton-Williams, table 5.8, p. 90). The NSHS data, however, show higher frequencies of questions about the length of the survey (NSHS was a longer interview (at least one hour) than the British (25 minutes)) and questions about how the household was chosen. NSHS shows lower percentages of the "I'm not interested" comment (note that the NSHS had a higher response rate, 86%, than the British survey cases, 79%). There are some differences in the classification of behaviors that make more detailed comparisons impossible, but the impression formed from the British data is not unlike that from the NSHS data.

4. Testing the Influence of Interviewer-Householder Interaction on Contact-Level Outcomes

We believe that full understanding of the process of unit nonresponse must acknowledge that sample cases often require several contacts by the interviewer, prior to obtaining an interview or a final refusal. Thus, we first examine how the nature of the interaction between householder and interviewer affects the outcome of individual contacts. Then we examine (in Section 5) whether the character of those interactions affects the final disposition of the case.

Customized appeals for householder participation requires the interviewer to maintain interaction with the householder over time in order to observe the unique concerns of the householder. "Success" by the interviewer on any individual contact, *under the principle of maintaining interaction*, is either an interview or some other result that permits future interaction with the sample household. We choose, therefore, to use the contact as the initial unit of analysis and to examine the outcome of individual contacts for each use.

In this section we examine the likelihood that an interview, an appointment, or other outcome occurs relative to a refusal. We do this separately for first, second, third, and fourth contacts with sample households. We have thus dissected the survey

activities pertaining to a sample household into smaller pieces, looking for systematic influences on householder reluctance or cooperation, within those pieces.

4.1. *Socio-demographic correlates of contact-level outcomes*

Because we are interested in placing our findings in the context of the past literature on socio-demographic correlates of survey participation, Table 1 presents the results of logistic regressions predicting the outcome of the first four contacts on cases, using socio-demographic indicators and interviewer-level variables. The predictors in the models are those used in Groves and Couper (1993a, 1993b), the result of the theoretical motivation above. All of these predictors were found to be influential on the response propensity of householders in six major U.S. Federal government surveys on topics ranging from consumer expenditures to drug usage.

The reader will note that the case bases for the regressions decline with the contact number. Approximately $9,101 - 6,977 = 2,124$ cases obtained their final disposition in the first contact; 3,330 in the second, etc. Although some households required more than four contacts to obtain a final disposition, their numbers are too small for useful statistical analysis.

Table 1 shows that the demographic variables tend to be predictive of the likelihood of a positive result on contact 1, but are less so for contacts 2–4. The important predictors from the past literature on final disposition are indeed important also for the *first* contact – refusals are higher among units with barriers to access, males, and those who live alone. On *later* contacts these variables are less important, with uniformly smaller coefficients (accompanied by higher standard errors because of the smaller case bases).

For none of the contacts do the urbanicity variables have measurable effects. Urbanicity is one of the most frequently cited correlates of response propensity (see Couper and Groves 1996), but has no statistically significant effect on the outcome of individual contacts. (In separate analyses, we find significant urbanicity differences on final dispositions in NSHS.) The smaller effects on the contact level could arise from interviewers' retreating, whenever possible, in the face of some resistance, regardless of the residential environment.

We interpret these results as implying that, with no indication of householder behavior, socio-demographic correlates commonly found in the past literature *do* separate householders by likelihood of cooperation. However, among those requiring more than one contact to obtain a final disposition, the socio-demographic variables lose their predictive power. This occurs for two reasons: (a) some of those demographic categories with higher response propensities fall out of the active set of sample householders because they give interviews on the first contact, and (b) for later contacts behavioral attributes of prior contacts are important indicators of the likelihood of cooperation.

4.2. *Interviewer-level effects*

Our theoretical perspective led us to expect that, controlling on householder attributes, experienced, confident interviewers would tend to obtain positive outcomes at the individual contact level. There is some evidence for this on the first contact

Table 1. Logit model coefficients for contact-level outcomes, using householder and interviewer attributes as predictors

	Contact 1	Contact 2	Contact 3	Contact 4
Intercept	2.61** (0.28)	2.08** (0.24)	1.93** (0.35)	2.50** (0.31)
<i>Householder attributes</i>				
Central cities of CMSA's	-0.28 (0.19)	0.018 (0.18)	0.34 (0.23)	0.20 (0.24)
Balance of CMSA	-0.14 (0.17)	0.066 (0.16)	0.20 (0.19)	0.17 (0.18)
Other MSA	0.13 (0.18)	0.26 (0.15)	0.28 (0.17)	0.11 (0.17)
(Non-MSA)				
Barriers to access for unit	-0.44** (0.11)	-0.27** (0.099)	-0.21 (0.14)	-0.15 (0.25)
Male householder	-0.25* (0.091)	-0.18* (0.084)	0.040 (0.15)	-0.25 (0.17)
30 or younger	-0.24 (0.17)	-0.15 (0.12)	0.05 (0.18)	0.26 (0.25)
31-40 years old	0.12 (0.19)	0.12 (0.14)	0.43* (0.18)	0.62 (0.30)
41-50 years old	-0.11 (0.21)	0.016 (0.14)	0.19 (0.19)	0.079 (0.27)
(51 or older)				
Single person HH	-0.31* (0.12)	-0.065 (0.15)	-0.27 (0.20)	0.16 (0.32)
(Multiple person HH)				
<i>Interviewer variables</i>				
Log (years worked)	0.12 (0.07)	0.042 (0.085)	-0.043 (0.084)	0.14 (0.098)
Confidence	0.15* (0.06)	0.095 (0.068)	0.064 (0.083)	-0.28** (0.10)
<i>Goodness of fit</i>				
Satterthwaite adjusted <i>F</i>	6.06	1.85	1.09	2.15
<i>p</i> -value for adjusted <i>F</i>	0.000	0.092	0.389	0.050
(<i>n</i>)	(8,837)	(6,769)	(3,503)	(1,687)

* $p < .05$ ** $p < .01$

Standard errors in parentheses.

(see Table 1), but these relationships too diminish in size and statistical significance over later contacts.

We are surprised by the negative effect of interviewer confidence in contact 4, but note the highly selective nature of these cases. Further, in an earlier analysis using these data (Groves and Couper 1994), we found that interviewers with the least experience (less than six months) had high levels of confidence relative to more

experienced interviewers. It may be that the confidence measure does not reflect the self-assurance that comes from command of persuasive skills, but from some spirit of enthusiasm that is not useful for refusal conversions.

In interpreting this finding it is also important to note that the data do not arise from a design with interpenetrated interviewer assignments; that is, interviewers have not been randomized to sample households. The assignment procedures can affect these results; Groves and Couper (1993b) found that experienced interviewers are often asked to persuade householders to cooperate after they have already refused another interviewer. In that analysis, when statistical controls were added for such reassigned cases, more experienced interviewers were found to increase the likelihood of cooperation over less experienced interviewers. Restricting the present analyses to contacts by field interviewers only, the number of interviewer switches on adjacent contacts is extremely small (about 3% on average across pairs of contacts) for the NSHS data. Adding an indicator for such switches, however, adds little effect to the interviewer experience measure.

Since we have shown that householder demographic variables lose their predictive power for cases requiring multiple contacts, the next logical analytic step is to examine whether the character of the interaction between interviewer and householder predicts the outcome of these contacts.

4.3. How the interaction between householder and interviewer affects contact-level outcome

We examine in two separate steps whether the contact description form captured behaviors that affect contact level outcomes:

- a. We test whether, controlling on the recorded outcome of the prior contact, the documented householder behavior informs us about the likelihood of a success (measured as avoidance of a refusal) on the current contact. This is important to the entire effort to measure interaction level attributes. If interviewers can routinely make observations about the contact-level interaction that are informative about the likelihood of later cooperation, then there is hope both for field administrative use of these data (adjusting field follow-up efforts based on the information) and for use in postsurvey nonresponse adjustment.
- b. We test whether changes in what the interviewer does from one contact to the next are related to the likelihood of success on a contact. This is a weak indicator of between-contact tailoring. For example, if interviewers are tailoring their behavior in the second contact to the householder cues in the first contact, we should expect more changes between the two contacts in what they do, relative to those who are not tailoring.

Table 2 presents the results of logistic regressions by contact number, using the same dependent variables as Table 1, but using predictors reflecting the nature of the interaction. For example, the first column uses the characteristics of the interaction in contact 1 and contact 2 to predict the outcome of contact 2. The dependent variable is coded as a binary variable with positive outcome (interview or appointment) modeled relative to a negative outcome (refusal). In all models we control

Table 2. Logit models predicting contact-level outcomes, using characteristics of prior contact and tailoring indicator

	Contact 2 outcome	Contact 3 outcome	Contact 4 outcome
Intercept	−0.14 (0.23)	0.74* (0.27)	0.59* (0.26)
<i>Outcome of prior contact</i> (Listing, refusal)			
Listing, no refusal	3.08** (0.19)	2.44** (0.24)	2.36** (0.27)
No listing, no refusal	2.53** (0.23)	1.76** (0.34)	1.70** (0.41)
No listing, refusal	0.28 (0.21)	−0.34 (0.28)	−0.15 (0.54)
<i>Householder behavior</i> <i>in prior contact</i>			
Negative statements	−1.12** (0.14)	−1.21** (0.18)	−1.20** (0.26)
Time-delay statements	−0.49** (0.13)	−0.37 (0.19)	−0.46* (0.22)
Questions	0.18 (0.10)	0.37* (0.15)	0.34 (0.28)
Indicator of tailoring (change of behavior)	0.28 (0.15)	−0.10 (0.18)	−0.15 (0.18)
<i>Significance tests (p values)</i>			
Prior contact outcome	0.000	0.000	0.000
Householder behavior	0.000	0.000	0.002
<i>Goodness of Fit</i>			
Satterthwaite adjusted <i>F</i>	136.69	72.86	30.56
<i>p</i> -value for adjusted <i>F</i>	0.000	0.000	0.000
(<i>n</i>)	(7,046)	(3,674)	(1,774)

**p* < .05

***p* < .01

Standard errors in parentheses.

for the outcome of the prior contact, modeled as three dummy variables representing the 2 × 2 table of refusal/no refusal by whether a household listing had already been obtained. Earlier analyses revealed the importance of treating these two variables as interactive rather than main effects.

One of the key analytic questions generated by the conceptual framework above is: “Does the character of the interaction in contact 1 color the outcome of contact 2, beyond the effects of the coded outcome of contact 1?” This question is interesting both theoretically and practically.

Theoretically the recorded disposition of contact 1 reflects the announced decision of the householder at the time of contact 1. Thus, if there are marginal effects of the first contact interaction, we have evidence for the impermanence or inadequacy of the householder’s stated initial reaction to the survey request. Practically, response rates

might be explained by whether interviewers attend to the nature of the prior interaction in preparing for the next contact with the sample household. This has implications for the use of call notes by interviewers to guide their actions on later contacts.

Table 2 shows that, controlling for the outcome of the first contact, householder behavior during the prior contact is informative about the chances of success of the current contact. Specifically, the coefficient for negative statements is negative and larger than that for time-delay statements, suggesting that the former has a greater negative effect on the probability of gaining cooperation in the current contact. Householders who ask questions in the prior contact, controlling for the outcome of that contact, are no less or more likely to produce a positive outcome in the current contact.

The strength of the householder behaviors as predictors does not radically change over successive contacts. (The larger standard errors for coefficients reflect diminished case bases.) There is some evidence that householder questions in later contacts become even more powerful predictors of later cooperation (a coefficient moving from .18 to .34), but replication is needed to have much assurance in this result. This offers a stark contrast to the declining importance of householder socio-demographic attributes over successive contacts.

In short, Table 2 answers the first question posed by finding that the likelihood of a positive outcome in the current contact is informed both by the character of the earlier interaction and the stated outcome of the contact. This means that the outcome codes provide summary guidance about the likelihood of future behaviors at the sample household, but that the nature of the interaction provides independently informative information. Prediction of the outcome of the current contact from merely knowing the outcome of the prior contact is less powerful.

The second analytic question we pose is whether the concept of tailoring finds empirical support in the data. The concept of tailoring notes that successful interviewers alter their behavior in reaction to cues provided by the householder. This alteration is aimed at addressing specific concerns of the householders, as suggested by their verbal and nonverbal behavior. Unfortunately, the data provide only a weak indicator of tailoring. It is not a measure of altered behavior in direct response to householder behavior, but instead, a measure of how much interviewers change what they do over successive contacts. This is expressed as $1 - (I^c \cap I^{c-1}) / (I^c \cup I^{c-1})$, where I^c is the set of different interviewer actions performed in the c th contact. This indicator of tailoring is also grossly measured in terms of broad categories of interviewer actions, and does not capture the nuances of variation in arguments within a particular strategy.

We expect that interviewers would observe smaller numbers of new cues from the householders over successive contacts (found in Groves and Couper 1994). Interviewers are most active in contact 1 and contact 2, providing a variety of information and arguments for cooperation to the householder. The frequency of such behavior in later contacts is lower. Given this, we speculated that *changes* in interviewer behavior were less likely to be measurable positive influences on later contact-level outcomes.

Table 2 shows that, controlling for characteristics of the prior contact, the tailoring indicator has effects in the right direction, but does not attain traditional levels of statistical

significance ($p = 0.066$) for the outcome of the second contact. Alternative versions of the tailoring measure (e.g., collapsing the measure into binary indicators or using counts of behaviors rather than ratios) *do* reach traditional levels of significance ($p < 0.05$).

We interpret this finding to mean that change in interviewer behavior between contacts may influence the likelihood of a positive outcome in the next contact. Emphasizing different aspects of the survey avoids refusals in the second contact, regardless of the outcome of the first contact.

No effects of tailoring are evident in models for later contacts in Table 2. This is consistent with the expectation that the greatest opportunities for tailoring arise in early contacts and with the fact that the power of tests is much reduced by fewer observations.

Thus, to summarize the models on contact-level outcomes, the typical socio-demographic variables partially explain the outcome of the initial contact, but are not useful in predicting the outcome of subsequent contacts. Further, simple interviewer observations of what the householder does in one contact are useful in predicting the outcome of the following contact. Finally, tailoring might increase the likelihood of success in early contacts with a sample household.

5. Effects of Interviewer-Householder Interaction on the Final Disposition of Sample Households

This section turns our focus to that more typical of the literature on survey non-response – the likelihood of obtaining an interview among those sample cases contacted. Of key interest is whether the measures of interviewer and householder interaction are informative about the likelihood of ultimate cooperation, controlling on socio-demographic attributes of the householder.

To accomplish this, we use a household-level, not a contact-level, analysis, limit our analysis to cases with more than one contact, and jointly examine the effects of socio-demographic attributes of the householder, householder behavior, interviewer attributes, and tailoring. To avoid any ambiguity of causal order, the householder behavior variables reflect behavior in contacts prior to the one yielding the final disposition. For example, the “negative statements” variable equals 1 for cases where a householder made any negative statements in contacts prior to the last one; 0, otherwise. The tailoring indicator measures changes in interviewer behavior only between the first and second contact. (Using some other summary indicator was rejected because it would fail to be a simple analogue of the indicators used on the contact-level models). The base model is Model 1 of Table 3, including only the socio-demographic predictors. Model 2 tests the marginal effect of householder behaviors. Model 3 tests the further marginal effect of interviewer tailoring.

Model 1 shows the expected result of lower cooperation among male and older householders, and among those with barriers preventing easy contact with the household. Although the urbanicity variables show decreased cooperation in urban areas, as expected, the effects are not statistically significant. The hypothesized lower cooperation of persons living alone also is not found. Finally, the marginal effects of interviewer tenure and self-confidence are negligible. In short, the base model shows

Table 3. Logit models predicting final disposition, using socio-demographic and interaction predictors, multi-contact cases only

	Model 1	Model 2	Model 3
Intercept	1.85** (0.27)	2.76** (0.29)	2.68** (0.29)
<i>Householder attributes</i>			
Central cities of CMSA's	-0.28 (0.15)	-0.25 (0.16)	-0.27 (0.16)
Balance of CMSA	-0.22 (0.15)	-0.047 (0.16)	0.010 (0.16)
Other MSA (Non-MSA)	0.11 (0.14)	0.22 (0.15)	0.20 (0.16)
Barriers to access for unit	-0.25* (0.11)	-0.19 (0.13)	-0.20 (0.14)
Male householder	-0.28** (0.80)	-0.20* (0.10)	-0.20* (0.10)
30 or younger	0.51** (0.16)	0.20 (0.19)	0.23 (0.19)
31-40 years old	0.68** (0.17)	0.44* (0.18)	0.48** (0.18)
41-50 years old (51 or older)	0.35* (0.17)	0.17 (0.18)	0.24 (0.18)
Single person HH (Multiple person HH)	-0.033 (0.14)	0.0045 (0.16)	0.0044 (0.16)
<i>Interviewer variables</i>			
Log (years worked)	-0.049 (0.071)	-0.14 (0.086)	-0.14 (0.084)
Confidence	0.088 (0.067)	0.043 (0.078)	0.036 (0.077)
<i>Householder behavior</i>			
Negative statements		-2.45** (0.13)	-2.47** (0.13)
Time-delay statements		-0.67** (0.099)	-0.68** (0.10)
Questions		0.77** (0.098)	0.76** (0.097)
<i>Tailoring between contact 1 and 2</i>			
			0.19 (0.15)
<i>Goodness of Fit</i>			
Satterthwaite adjusted <i>F</i>	4.62	52.53	49.62
<i>p</i> -value for adjusted <i>F</i>	0.002	0.000	0.000
(<i>n</i>)	(6,835)	(6,625)	(6,477)

p* < .05*p* < .01

Standard errors in parentheses.

most of the expected effects of householder socio-demographic attributes, but not those of interviewer attributes.

Model 2 shows that controls on the householder behavior diminish the effect of their socio-demographic attributes, and improve the overall fit of the model relative to Model 1. For example, although householder age was informative of the likelihood of an interview in Model 1, it is less so in the presence of controls on householder behavior. In other words, among persons who provide negative statements, time-delay statements, or ask questions at the same rate, there are few differences in cooperation across age groups. Similar statements can be made for different urbanicity groups, housing units with some barrier to access, male householders, and single person households. All of these results are evidence that the interviewer observations about what happens during contacts are independently informative about the ultimate likelihood of cooperation.

Model 3 adds to these variables another central tenet of the theory – interviewer tailoring, as measured by changes in interviewer behavior between the first and second contact. This indicator has weaker predictive power for the final disposition than it had for the result of the second contact (first column of Table 2), and it fails to achieve statistical significance at traditional levels. Its effect is in the right direction; changes in interviewer behavior between the first and second contacts improve the likelihood that the household will eventually grant an interview.

In short, the models in Table 3 show that differences in householder behaviors explain some of the traditional effects of socio-demographic attributes on survey cooperation. It is gratifying to observe that even rather crude measures of such householder behaviors can demonstrate these effects. Such crude measurement is not sufficient to reflect interviewer tailoring, however, and the data do not themselves support the marginal effects of tailoring, controlling on householder behavior. We believe, however, that tailoring remains an influence on final disposition and assert that a purified measure of it, reflecting interviewer effectiveness in adapting their behavior to respondent concerns, would provide empirical support for the theory.

6. Conclusions

We have found results largely consistent with the theoretical propositions outlined above. Socio-demographic and psychological predispositions of householders are useful proxy indicators to their likely behavior in the first contact with an interviewer. These attributes are associated with behavioral tendencies that are consistent over time. The indicators, however, are fallible; they are correlates, not causes of the survey participatory behavior we study.

Once the first contact is completed, the behaviors manifested in that contact are much more powerful indicators of the likelihood of survey participation in any specific later contact with the household as well as of the likelihood of participation occurring over all future contacts.

Further, the nature of the householder behavior provides systematic indications of response propensity. Negative statements at any prior contact bode poorly for ultimate participation. But negative statements are relatively rare (see Groves and

Couper 1994), and they are by no means perfect predictors of negative outcomes. Interviewer behavior in subsequent contacts can eliminate their effects.

Questions, on the other hand, are indications of householder cognitive engagement in the decision to participate; they bode well for participation in future contacts. Time delay statements fall in-between the two other behaviors. These may contain a mix of genuine time constraints and polite proxies for refusals.

Interviewers can act to change the base tendency of a householder to participate, even after they have exhibited one of the behaviors above. Altering the kind of information they provide to the respondent, emphasizing one aspect of the design or intent of the survey over others, can influence householders to participate. We believe that the most effective changes in interviewer behavior are those shaped by real concerns revealed by householders. This notion of tailoring has the power to change the calculus of the decision-making on the part of the householder.

We believe these results, measured relatively crudely across contacts here, would extend to behavior within a given contact. We expect that within-contact tailoring should be a more powerful influence on the decision to participate, as it immediately addresses the concerns and objections raised by the householder.

As we noted earlier, the decision to cooperate or refuse is shaped in large part by the interactions between householders and interviewers. The data at our disposal, albeit providing weak indicators of the richness of these interactions, nevertheless support the notion that such interactions are informative of the process of reaching a decision about participation in a survey.

This article began with commentary on the nature of postsurvey adjustments for survey nonresponse and criticized the past literature for ignoring model specification issues. We do not claim the models presented in this article are ideal specifications. However, we do hope the article makes the case that attention at the survey design stage to measurement of indicators of response propensity can lead to better fitting adjustment models. Even the rather crude initial indicators used in this study are explanatory of survey cooperation. Further development of such measures, guided by theoretical propositions regarding survey cooperation, deserves attention.

7. References

- Abelson, R. (1981). The Psychological Status of the Script Concept. *American Psychologist*, 36, 715–729.
- Couper, M.P. and Groves, R. M. (1996). Social Environmental Impacts on Survey Cooperation. *Quality and Quantity*, 30, 173–188.
- Groves, R.M., Cialdini, R.B., and Couper, M.P. (1992). Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly*, 56, 475–495.
- Groves, R.M. and Couper, M.P. (1993a). Unit Nonresponse in Demographic Surveys. *Proceedings of the U.S. Bureau of the Census Annual Research Conference*, 593–619.
- Groves, R.M. and Couper, M.P. (1993b). Multivariate Analysis of Nonresponse in Personal Visit Surveys. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 514–519.

- Groves, R.M. and Couper, M.P. (1994). *Householders and Interviewers: The Anatomy of Pre-Interview Interactions*. Survey Research Center, University of Michigan: SMP Working Paper No. 11.
- Groves, R.M. and Couper, M.P. (1995). Theoretical Motivation for Post-Survey Nonresponse Adjustment in Household Surveys. *Journal of Official Statistics*, 11, 93–106.
- Heeringa, S. (1992). *National Survey of Americans' Mental Health, Study Sample Design*, Survey Research Center, University of Michigan: unpublished report.
- Kalton, G. (1983). *Compensating for Missing Survey Data*. Ann Arbor: Institute for Social Research.
- Kessler, R., McGonagle, K., Zhao, S., Nelson, C., Hughes, M., Eshelman, S., Wittchen, H., and Kendler, K. (1994). Lifetime and 12 Month Prevalence of DSM-III-R Psychological Disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry*, 51, 8–19.
- Lessler, J.T. and Kalsbeek, W.D. (1992). *Nonsampling Error in Surveys*. New York: John Wiley and Sons.
- Little, R.J.A. (1986). Survey Nonresponse Adjustments. *International Statistical Review*, 54, 139–157.
- McCrossan, L. (1993). *Respondent-Interviewer Interactions in Survey Introductions*. Paper presented at the Fourth International Workshop on Household Survey Non-response, Bath, U.K., September.
- Morton-Williams, J. (1993). *Interviewer Approaches*. Aldershot: Dartmouth.
- Oksenberg, L., Coleman, L., and Cannell, C. (1986). Interviewer Voices and Refusal Rates in Telephone Surveys. *Public Opinion Quarterly*, 50, 97–111.
- Shah, B.V., Barnwell, B.G., Hunt, P.N., and LaVange, L.M. (1993). *SUDAAN User's Manual, Release 6.34*. Research Triangle Park, N.C.: Research Triangle Institute.

Received September 1994

Revised January 1996