THE INFLUENCE OF INTERVIEWERS' CONTACT BEHAVIOR ON THE CONTACT AND COOPERATION RATE IN FACE-TO-FACE HOUSEHOLD SURVEYS

Michael Blohm, Joop Hox and Achim Koch

In surveys, interviewers serve as the agents of data collection. Their task includes contacting the target persons, gaining their cooperation, and conducting the interviews according to the rules of standardized interviewing.

Interviewers are not equally successful at doing their job. They differ both in the quality of the data collected and in the response rate they achieve (Biemer & Lyberg, 2003, pp. 110–11, p. 156ff.). It is often difficult to distinguish to what extent these differences arise from differences among interviewers or from differences between the areas (and the target persons, living in these areas) assigned to the interviewers. Research using interpenetrated sample designs, however, has shown that interviewer effects can remain strong even when area effects are controlled (Campanelli & O'Muircheartaigh, 1999).

The reasons for the different response rates between interviewers are still not totally clear. In their summary of the research literature on this topic, Groves and Couper (1998, p. 191) conclude that demographic interviewer characteristics like gender and age play a minor role in gaining cooperation. The same holds for stable interviewer personality characteristics such as extraversion. What do seem to be important factors are interviewer experience, expectations and attitudes, and the behavior of the interviewers in their interaction with the respondents. However, in an international comparative study Hox and De Leeuw (2002) found that interviewer attributes, attitudes, and avowed behaviors explained only a small part of the variation in nonresponse rates between countries.

There are two important components of nonresponse: Noncontact and refusal (De Leeuw & De Heer, 2002). Research on reducing nonresponse has mainly focused on reducing refusals, for instance, through incentives (Singer, 2002), interviewer behavior (Snijkers, Hox, & De Leeuw, 1999; Morton-Williams, 1993), and tailoring the interviewer approach to the initial refuser (Groves & Couper, 1998, pp. 37–42; Stoop, 2004). So far research on calling strategies has mainly concentrated on the effects of number and timing of calls (day of week and time of day) on probability of contact (e.g. Purdon, Campanelli, & Sturgis, 1999). Survey mode can have an impact on noncontact and refusal too. Goyder (1987) and De Leeuw and Van der Zouwen (1988) report meta-analyses that show amongst others that telephone surveys tend to have a lower response rate than face-to-face surveys. In a more detailed meta-analysis, Hox and De

Leeuw (1994) show that there is, on average, a higher response rate to face-to-face surveys than either telephone or mail surveys, but there is considerable variation among the response rates in individual surveys. This is corroborated by Bretschneider and Schumacher (1996).

REASONS WHY MODE OF CONTACT SHOULD MATTER

In the present paper our primary focus is on a specific aspect of the contact behavior of interviewers in face-to-face surveys. We investigate whether the mode of the first contact to a respondent makes a difference: Do interviewers who show up on the doorstep unannounced achieve better or worse results than interviewers who telephone ahead to set up an appointment?

Why should the mode of the first contact make a difference? There are several reasons why an unexpected first contact via a personal visit may have a positive effect especially on the cooperation rate:

1. The first contact with a respondent is typically a very brief interaction. On the telephone, this interaction is even shorter than in a face-to-face contact (Groves & Couper, 1998, p. 219). In addition, in a telephone contact the interviewer has to rely on auditive cues (De Leeuw, 1992, pp. 16–17), s/he cannot see the respondent (his/her dress, facial expression, etc.), nor observe, for example, characteristics of the housing unit the respondent is living in. This produces more difficulties for the interviewer to communicate his/her purpose and to tailor his/her approach to the specific respondent and the respondent's situation and concerns (Groves & Couper, 1998, p. 301).
2. The central question for the respondent during the first contact with an interviewer is: What does this person want of me? As sales calls are predominantly done by phone, the interviewer might be more easily mistaken for a salesperson. This, all the more, since the interviewer cannot display identification nor written materials that convey the legitimacy of the survey request on the telephone (De Leeuw & Hox, 2004, p. 465).
3. Telephone communications between strangers are guided by different norms than face-to-face communications (Groves & Couper 1998, p. 301; De Leeuw, 1992, pp. 14–15). Even when a respondent has adequately perceived the purpose of the interviewer's call, but does not want to consent to the survey request, it is easier for him/her to hang up the phone than to close the door while the interviewer is physically present.

In contrast, there are also reasons why a first contact via a telephone may have a positive effect on both the contact rate and the eventual cooperation rate.

1. The failure to contact prospective respondents is a major cause for low response rates. Increasing the number of contacts has implications for the survey costs (cf. Groves, 1989). Using the phone for the first contact makes it possible to increase the number of contact attempts dramatically and still keep the costs low. It also makes it possible to schedule the contact attempts at times when respondents can

be reached and interviewers are not willing to roam the streets, such as early evenings. Also the telephone makes it possible to contact persons in areas that interviewers try to avoid for safety reasons, and in buildings that are difficult to enter due to security restrictions (e.g. flats with intercoms, condominiums with closed video and security cameras). Thus, using a telephone for the first contact facilitates follow up contacts, which increases the chance that ultimately a contact is made, and facilitates contacts in areas that otherwise are not canvassed.

2. From a psychological point of view, a telephone conversation can also act as a foot-in-the-door. When a telephone contact is made, the respondent has already consented in talking to the interviewer on the telephone. Social psychological findings (cf. Groves, Cialdini, & Couper, 1992) suggest that, after consenting to a small request, subjects are more likely to consent to a larger request as well. So, the foot-in-the-door provided by the first telephone contact may result in higher response rates.

To summarize: Our main question is whether the mode of first contact has an effect on the response rate. To investigate the process more fully, we distinguish between contact rate (the probability that an eligible respondent will be contacted) and cooperation rate (the probability that a respondent, once contacted, will cooperate). In addition, we investigate the effect of a number of available respondent and interviewer characteristics on the probability of response.

METHOD

SAMPLING DESIGN

We use data from the German ALLBUS 2000 survey. The ALLBUS General Social Survey is a bi-annual survey, fielded every second year since 1980. Its target is the long-term monitoring of attitudes, behavior, and social structure. Sampling and fieldwork are commissioned to a commercial survey organization. In 2000 Infratest Social Research (Munich) was responsible for data collection.

ALLBUS uses a national area probability sample of non-institutionalized adults in Western and Eastern Germany, with some oversampling of Eastern Germany. In 2000 the sample was drawn in two stages. In the first stage 151 communities (including 162 Primary Sampling Units, PSUs) were selected. In the second stage 40 addresses of individuals were randomly selected from the communities lists of residents for every PSU. For (nearly) every selected address information on sex, age, (non-)German citizenship was available from the registers.

ALLBUS 2000 was fielded as a CAPI survey. The average length of the interview was about 60 minutes, including a supplementary questionnaire. The fieldwork started in January and ended in July. In total 3,138 interviews were realized, 2,036 in Western Germany and 1,102 in Eastern Germany. The response rate was 49.1 percent (response rate RR5 as defined by the AAPOR *Standard Definitions*, AAPOR, 2004). In Eastern Germany the interviewers were slightly more successful than in Western Germany (response rate 53.7 vs. 46.9 percent). For the present analyses we will restrict ourselves to

the 2,109 interviews realized in the main fielding period, because in this period no re-issuing of cases to another interviewer took place.

The eligible sample in the main fielding period comprised 6,286 addresses. The response rate totaled to 33.6 percent.

OUTCOME VARIABLES AND RESPONDENT VARIABLES

Since the causes and correlates of the two main reasons of nonresponse, noncontact and refusals, are quite distinct, we analyze two separate outcome variables: The contact rate and the cooperation rate. The contact rate gives an indication of how successful the attempts at contacting were. The cooperation rate indicates how successful an interviewer is in persuading a potential respondent after the contact has been made. The contact rate was calculated as CON3 and the cooperation rate as COOP3 as defined by the AAPOR *Standard Definitions* (AAPOR, 2004).

The *outcome variables* in the analysis were: (1) Contact: a dichotomous variable indicating if a contact was made (contact rate 86.9 percent), and (2) Cooperation: a dichotomous variable indicating if a contacted respondent provided an interview (cooperation rate 41.1 percent). The following *respondent variables* were available for each targeted respondent (i.e. for both respondents and nonrespondents): age, sex, nationality, city size, housing with entrance intercom, and dilapidated housing. Detailed information on question wording and coding of these variables is given in Appendix A.

INTERVIEWERS AND INTERVIEWER VARIABLES

In 2000 a total of 296 interviewers had been working on ALLBUS in the main fielding period. The interviewers mainly interviewed persons living in the same general geographical area in which they lived. There was no advance letter sent out centrally to the target persons. However, all interviewers received a letter they could hand over to the respondents once in contact with them. The interviewers had to make at least four calls to the target persons, spread over different days of the week and different times of the day. The interviewers were allowed to decide whether they contacted the subjects first by telephone or in person. Where possible, the survey organization provided the telephone number of the target persons to the interviewers. The telephone contact could only be used to set up an appointment, the interview itself had to be conducted face-to-face.

On average, the interviewers needed 3.2 calls to realize an interview (1.9 personal visits + 1.3 telephone calls). Outcomes coded as noncontacts had on average 4.2 calls (2.6 personal visits + 1.6 telephone calls), refusals had on average 2.4 calls (1.2 personal visits + 1.2 telephone calls).

Approximately one year after the data collection for ALLBUS, a questionnaire was administered to all interviewers who had been working on the survey. The questionnaire included questions on demographics, on the job as an interviewer, on contact behaviors, and on the experiences with interviewing for ALLBUS 2000. In total 232 interviewers returned the questionnaire (78.4 percent). Of the remainder, 33 interviewers (11.1 percent) did not return the questionnaire, and 31 interviewers (10.5 percent) had left the interviewing staff.

The interviewer variables are divided into four broad categories, each including a number of related variables. The first group consists of the basic demographic variables: age, sex, education, employment status, and degree of interviewer experience. The second group relates to interviewer contact behavior: Telephone ahead for appointment, show up unannounced for interview, show up unannounced for appointment. The third group describes two types of interviewer doorstep behavior: Stressing the importance of the survey to target person and society, and presenting examples. The fourth group consists of job (interviewing)-related variables: How often an interviewer was deployed for refusal conversion, intrinsic job motivation, satisfaction with interviewing job, self-confidence in ability to persuade, preference for specific subgroups, two different interviewer attitudes (willingness to accept refusal and doubting data quality if subject is coerced), and interviewing habits (works on weekdays, works on weekends, works in the afternoon, and works other hours). Further information on the exact wording and coding is in Appendix B.

ANALYSIS MODEL AND ANALYSIS STRATEGY

In ALLBUS 2000 the PSUs and the sample persons were assigned to the interviewers on a geographical basis, that is no interpenetrated design was used. There is therefore some potential confounding of interviewer and target persons characteristics. We use a multilevel logistic regression model with respondents nested within interviewers to control for differences between target persons (Hox, De Leeuw & Kreft, 1991; Campanelli & O'Muirchartaigh, 1999). Interviewer effects are modeled conditional on the target persons' characteristics (for details, see Appendix C). Using multilevel modeling also effectively incorporates the clustering in the sample caused by having respondents nested within interviewers (Hox, 2002).

When we add interviewer variables to the multilevel logistic regression model, we face a serious problem concerning incomplete data. Not all interviewers filled in the interviewer questionnaire, and some interviewers incidentally skipped questions. At the interviewer level, missing values are handled by listwise deletion, and in the multilevel model this means that all information from the target persons related to deleted interviewers is also deleted. If a large number of interviewer variables is included in the model, this results in a serious loss of cases. A good solution is to restrict the number of explanatory variables at the interviewer level to the best predictors. Therefore, before the interviewer variables were added to the model, a selection procedure is followed. In a preliminary multilevel analysis, all interviewer variables are added blockwise, and in each block variables with associated p-values $> .25$ are removed. Explanatory variables with associated p-values $\leq .25$ for either contact rate or cooperation rate are retained for the final multilevel analysis.

RESULTS

First we present the results of multilevel analyses on both contact rate and cooperation rate, using only explanatory variables at the respondent level. This allows us to assess the impact of target persons' variables using all available respondent information. Subsequently, we

add the preliminary selected interviewer variables and again present the results of multi-level analyses. As described above, this involves a certain loss of data due to missing data at the interviewer level. To assess the potential bias due to deleting cases with incomplete interviewer data, we repeat the analyses on the smaller sample using only respondent-level variables. By comparing the results of these analyses with the results from the former (based on the total sample of target persons), we can determine the amount of this bias.

TARGET PERSONS' VARIABLES

For each outcome variable, contact and cooperation, two models are presented in Table 1. The first model is the baseline model, without explanatory variables. This model shows how much variance is located at the interviewer level. The second model adds the respondent variables (fixed effects) and significant effects for respondent variables which vary across interviewers (random effects) with varying slopes if these are significant. The sample sizes for this analysis are 280 interviewers and 6,015 target persons for *contact*, and 278 interviewers and 5,198 target persons for *cooperation*.

The baseline model for *contact* shows a significant variance component at the interviewer level. In the model that includes the respondent-level variables, all variables except nationality (non-German, $p = .99$) have a significant effect on the probability of contact. Elderly people are more likely to be contacted than younger people. Women are easier to contact than men. It is easier to contact people in rural than in urban areas. If

TABLE 1 **Parameter estimates for target person variables, for** *contact* **and** *cooperation*

Predictor variables	Outcome: Contact		Outcome: Cooperation	
	Baseline	Target persons	Baseline	Target persons
Fixed				
Target person level				
Intercept	2.22 (.09)	2.33 (.09)	−0.65 (.07)	−0.63 (.07)
Age		0.18 (.03)		−0.05 (.02)
Sex		0.22 (.08)		−0.10 (.06)^{ns}
Non-German		−0.003 (.19)^{ns}		0.75 (.16)
City size		−0.18 (.04)		−0.11 (.04)
Entrance intercom		−0.15 (.05)		−0.21 (.03)
Dilapidated housing		−0.12 (.04)		−0.20 (.03)
Random				
Intercept	1.45 (.20)	1.30 (.19)	0.95 (.12)	0.94 (.13)^{ns}
Entrance intercom	—	—	—	0.08 (.04)
Deviance	15352.69	15231.58	15014.43	14909.78
N	6015	6015	5198	5198

Note: Standard deviation in brackets. ns: Coefficient not significant.
Source: ALLBUS 2000, survey of interviewers.

there is an entrance intercom the contact rate is lower. The more neglected the housing is, the lower the contact rate. After including the target person variables, there is still significant between-interviewer variance left.

The baseline model for *cooperation* shows a significant variance component at the interviewer level. In the target person-variables model, all variables with the exception of sex ($p = .13$) have a significant effect on the probability of cooperation. Age has a negative effect on likelihood of cooperation. Non-Germans are more likely to cooperate than Germans. The cooperation rate is lower in large cities, if there is an entrance intercom, and if the housing is in a neglected condition. After including the target person variables, there is still significant between-interviewer variance left. There is one explanatory variable that shows significant between-interviewer slope variance: The presence of an intercom at the entrance of the building or housing unit.

INTERVIEWER VARIABLES

The selection procedure established the following interviewer variables for the final analysis: basic demographics (age, fully employed, experience); typical contact behavior (arrive unannounced for appointment, arrive unannounced for interview);[1] typical doorstep behavior (stress importance, give examples); and job (interviewing)-related variables (works weekdays, works weekend, works afternoon, works other hours). Given these interviewer variables, the sample sizes for the final analyses are 204 interviewers and 4,434 target persons for *contact*, and 203 interviewers and 3,892 target persons for *cooperation*.

The first model for both outcomes is a model that includes only respondent variables. This model allows us to check whether the target persons who were dropped because of missing interviewer data can be considered a random subsample. The second model adds the interviewer variables. Table 2 presents the parameter estimates.

We found no noticeable differences for the target persons' characteristics between the models including all target persons (Table 1) and the models with the reduced number of target persons due to missing data at the interviewer level (Table 2). The largest difference is for the intercept in both models, but it is not significant ($p > .22$). The interviewer variance also appears somewhat lower in the reduced sample, but again this difference is not significant ($p > .23$). Therefore we interpret our final sample as a random subsample of all targeted respondents.

For both outcome variables (contact and cooperation) the effects of the target persons' variables remain the same after including the interviewer variables in the model (cf. Table 1).

In the model for *contact* we find no effect of the reported mode of first contact. From the interviewer variables only three achieve a significant effect. The older the interviewer, the higher the contact rate s/he achieves. Interviewers working primarily in the afternoon (between 3 p.m. and 8 p.m.) are more successful in contacting target persons, whereas interviewers working on the weekend are less successful.

[1] The variable 'Telephone ahead for appointment' is left out in the further analyses, because it is highly correlated ($-.62$; $p < .001$) with the variable 'Show up unannounced for interview.' In the following we interpret the variable 'Show up unannounced for interview' as an continuum with the end points 'Always show up unannounced' to 'always phone ahead.'

TABLE 2 **Parameter estimates for contact and cooperation, respondent-level and final model**

Predictor variables	Outcome: Contact		Outcome: Cooperation	
	Target persons	Interviewer	Target persons	Interviewer
Fixed				
Target person level				
Intercept	2.45 (.10)	2.43 (.09)	−0.50 (.08)	−0.52 (.08)
Age	0.16 (.03)	0.16 (.03)	−0.06 (.02)	−0.06 (.02)
Sex	0.23 (.09)	0.23 (.10)	−0.05 (.07)	−0.04 (.07)^{ns}
Non-German	0.00 (.21)^{ns}	0.02 (.22)^{ns}	0.76 (.18)	0.76 (.18)
City size	−0.19 (.05)	−0.20 (.06)	−0.12 (.04)	−0.10 (.04)
Entrance intercom	−0.19 (.06)	−0.20 (.06)	−0.19 (.04)	−0.19 (.04)
Dilapidated housing	−0.15 (.05)	−0.16 (.05)	−0.20 (.04)	−0.20 (.04)
Interviewer level				
Experience		−0.02 (.13)^{ns}		—
Fulltime job		—		−0.58 (.18)
Age		0.20 (.09)		0.07 (.08)^{ns}
Unann. appointment		—		−0.12 (.08)^{ns}
Unann. interview		—		0.22 (.07)
Stress importance		—		−0.15 (.07)
Give examples		—		—
Works weekdays		—		—
Works weekend		−0.73 (.30)		—
Works afternoon		0.79 (.40)		—
Works other hours		—		—
Random				
Intercept	0.96 (.20)	0.80 (.19)	0.83 (.13)	0.65 (.11)
Entrance intercom	—	—	0.08 (.04)	0.08 (.04)
Deviance	11096.15	11078.58	11259.62	11228.21
N	4434	4434	3892	3892

Note: Standard deviation in brackets. ns: Coefficient not significant, but $p < .10$. Dash indicates $p > .10$
Source: ALLBUS 2000, survey of interviewers.

In contrast, in the model for *cooperation*, the mode of the first contact does have a significant effect on the cooperation rate. Interviewers who report that they normally show up unannounced and try to conduct the interview achieve better results than interviewers who telephone ahead to set up an appointment. With respect to the deviance reduction this is the strongest effect among the interviewer level variables. In addition, there is also a strong effect for the variable 'fulltime job': Interviewers who work fulltime in addition to their interviewing job achieve worse cooperation rates than interviewers who do not work fulltime in addition to their job as an interviewer. Contrary to what we may expect from persuasion theory (Groves et al., 1992), interviewers who stress the importance of the survey are less successful in persuading target persons. There is no significant effect of the degree of interviewer experience on the cooperation rate in our final model.

DISCUSSION

The present analysis focused on the question: Does the mode of the first contact to a respondent—either by a phone call or by a personal visit—make a difference for survey participation in face-to-face household surveys? We found that interviewers who report that they normally show up unannounced to conduct an interview achieve higher cooperation rates. Interestingly enough, the mode of first contact had no effect on the contact rate. A possible explanation is the limited number of call attempts made by our interviewer, who did not use the telephone as a means to pursue potential respondents.

Our results must be interpreted with some caution. First, we could not investigate the interviewers' contact behavior in a controlled experiment. Although we control statistically for selectivity effects, there might be some left. Second, we analyzed the effect on the cooperation rate, but it is unclear if a higher cooperation rate also results in a lower response bias. Further research has to analyze whether contacting in person results in a higher cooperation/response rate but in a lower total survey error as well. Third, the most important limitation of the present analysis lies in the fact that we rely on reports from the interviewers on what they usually do in a survey. The question is how well these reports match actual behavior. Groves and Couper (1998, p. 73ff) caution that comparisons of such reports with self-reports at the contact level show some over-reporting. However, analyzing actual contact behavior is more complex than it appears, because in most cases there is a sequence of contact attempts, which is ended by a decision (refusal, cooperation) or because a maximum number of contact attempts is exceeded. In the present case, some limited contact information is available, and the central independent variable 'contact behavior' shows a close relationship with the aggregate number of telephone calls and personals visits derived from the contact forms ($p < .001$).

The analysis of the interviewer and respondent characteristics showed results that are also found elsewhere in the literature. The exception is interviewer experience, which did not show the expected result. An interesting result is the effect of interviewers who have a fulltime job in addition to their job as an interviewer. Those interviewers have less success in persuading target persons to cooperate. The most likely explanation is that these interviewers (can) spend only a limited amount of time in attempts to persuade reluctant interviewees.

A key conclusion from this study is that the choice of using a telephone to contact potential respondents should not be left to the individual interviewers. Rather, using the telephone should be an explicit option when specified conditions are met, such as buildings with physical obstacles (e.g. monitors), risky neighborhoods, or unusual time slots. That means that the approach to the target persons should only be adapted to their needs, especially their reachability, and not to the needs of the interviewers.

APPENDIX A: OUTCOME AND RESPONDENT VARIABLES

OUTCOME VARIABLES

Contact rate: Contacts excepting frame errors: $N_{refused} + N_{not\,able} + N_{interviews} / N_{not\,contacted} + N_{refused} + N_{not\,able} + N_{interviews}$

Cooperation rate: Cooperation excepting frame errors, noncontacts, and unable to respond: $N_{interviews} / N_{refused} + N_{interviews}$

Subject Variables

Age: Originally in years, for reasons of scale divided by 10. *Sex*: male = 1, female = 2. *Nationality*: German = 1, Non-German = 2. *City size* (inhabitants): 0 ≤ 2,000 = 1, 2,000 ≤ 5,000 = 2, 5,000 ≤ 20,000 = 3, 20,000 ≤ 50,000 = 4, 50,000 ≤ 100,000 = 5, 100,000 ≤ 500,000 = 6, > 500,000 = 7. *Entrance intercom*: (mostly) Presence of phone system, composite Z-score. *Dilapidated housing*: (mostly) Large, run-down housing, composite Z-score.

APPENDIX B: INTERVIEWER VARIABLES

Basic Demographics

Age: Originally in years, for reasons of scale divided by 10. *Sex*: male = 1, female = 2. *Education*: 'What's your highest school certificate?' 1 = Lowest formal qualification of Germany's tripartite secondary school system, after 8 or 9 years of schooling; 2 = Intermediary secondary qualification, after 10 years of schooling; 3 = Certificate fulfilling entrance requirements to study at a polytechnical college or a university; 4 = polytechnical college degree or university degree.

Employment status: 'Do you do anything apart from your work as an interviewer?' 1 = fulltime employed, 0 = part time employed, pupil/student, retired, unemployed, housewife/houseman, other. *Interviewer experience* (in years): 'How long in total have you been working as an interviewer?'

Contact Behavior

'When you conduct interviews with a sample of named addresses, without a central advance letter, how do you go about it?' Items: I phone ahead for an appointment. I show up unannounced for an interview. I show up unannounced for an appointment. Answer categories for each item: Always = 5, often = 4, occasionally = 3, seldom = 2, never = 1.

Doorstep Behavior

'How often do you do or say any of the following? It is not important whether you have used the exact words we use here.' Items: see Table B1. Categories: Always = 5, often = 4, occasionally = 3, seldom = 2, never = 1.

We use as variables for analysis two factor scores based on the nine avowed behavior questions in Table B1. We call the factors: (1) Stress importance to subject and society and (2) Present examples.

TABLE B1 **Interviewer behavior question items and factor loadings on 'Stress importance' and 'Present examples'**

	Stress importance	Present examples
Mention that this is the chance for them to give their opinion	.63	−.02
Explain the general public will benefit from survey results	.75	.08
Mention that you don't want to sell anything	.24	.11
Say that the topic of the survey will interest them	.49	.33
Explain the topics of the survey	.01	.69
Give the target person some examples of what the survey is about	.13	.79
Understate the duration of the interview	.01	−.01
Explain that for politics, national economy surveys are an important basis for decisions	.49	.09
Mention that it is not a test of knowledge and that there are no right or wrong answers	.36	.08

JOB (INTERVIEWING)-RELATED VARIABLES

Deployed for refusal conversion (by the survey organization): 'How often are you deployed for refusal conversion/difficult respondents?' Always = 5, often = 4, occasionally = 3, seldom = 2, never = 1.

Intrinsic job motivation: 'How important is the following reason for you to work as an interviewer?' Item: Job is interesting. Very important = 4, quite important = 3, less important = 2, not at all important = 1.

Satisfaction with interviewing job: 'Generally speaking how satisfied or dissatisfied are you with your work as an interviewer?' Very satisfied = 5, fairly satisfied = 4, neither satisfied nor dissatisfied = 3, fairly dissatisfied = 2, very dissatisfied = 1.

Self-confidence in ability to persuade: 'If you assess yourself as an interviewer, what do you think are your strengths and weaknesses?' Fairly strong = 1 to strong = 5.

Preference for specific subgroups: Interviewers were asked how much they like to hold interviews with people from specific subgroups (men, women, elderly, middle aged, younger, people with lowest formal educational qualification, people with intermediary educational qualification, people with high educational qualification/university degree). Categories: Like very much = 5, neither like nor dislike = 3, dislike very much = 1. We used the standard deviation of the preferences for the eight specific subgroups as variable 'preference for specific subgroups'.

Interviewer attitudes: 'To what extent do you agree or disagree—based on your experience—with the following statements?' Items: see Table B2. Categories: 4 = strongly agree, 3 = agree, 2 = disagree, 1 = strongly disagree. We use as variables for analysis two factor

TABLE B2 **Interviewer attitude question items and factor loadings on 'Willingness to accept refusal' and 'Doubting data quality if subject coerced'**

	Willingness to accept refusal	Doubting data quality if subject coerced
One should always emphasize the voluntary nature of participation	.21	.14
Reluctant respondents should always be persuaded to participate	.05	−.37
If a respondent is reluctant, a refusal should be accepted	.32	.44
With enough effort, even the most reluctant respondent can be persuaded to participate	−.60	−.02
It does not make sense to contact reluctant target persons repeatedly	.45	.05
An interviewer should respect the privacy of the respondent	.52	.02
If you just catch them at the right time, most people will agree to participate	−.42	.19
Respondents persuaded after great effort do not provide reliable answers	−.20	.70

scores based on the eight interviewer attitude questions in Table B2. We call the factors: (1) Willingness to accept refusal; and (2) Doubting data quality if subject is coerced.

Interviewer scheduling habits: 'On which days do you typically try to conduct your interviews?' 'At which times of the day do you typically try to conduct your interviews?' Multiple responses were possible in both questions. Variables were constructed as follows: *Works weekdays* = number of choices for weekdays divided by 5. *Works weekend* = number of choices for Saturdays or Sundays divided by 2. *Works afternoon* = number of choices between 3 p.m. and 8 p.m. divided by 4. *Works other hours* = number of choices until 3 p.m. and after 8 p.m. divided by 3.

APPENDIX C: DETAILS ON THE ANALYSIS MODEL

The basic multilevel model analyses the outcome variable at the individual (respondent) level, as follows:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \tag{1}$$

In Equation (1), Y_{ij} is the outcome variable (contact or cooperation) for respondent i approached by interviewer j, X_{ij} is a respondent-level explanatory variable, for example the respondents' age or sex, and e_{ij} is the residual error term. The intercept β_{0j} is allowed to vary across interviewers, which reflects the interviewers' varying ability to contact or persuade respondents. In the multilevel model, interviewer-level explanatory variables

predict overall differences between interviewers by predicting differences in the interviewers' intercepts, as follows:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \tag{2}$$

In equation (2), Z_j is an interviewer-level explanatory variable, for example the interviewer's age or sex.

The regression slopes β_{*j} for the respondent-level explanatory variables may also vary across interviewers. If a respondent-level explanatory variable shows significant variation across interviewers, this means that the effect of this variable differs across interviewers, which indicates a possible interaction of respondent and interviewer characteristics. In the multilevel model the regression slopes β_{*j} are predicted by interviewer-level explanatory variables, for example as follows:

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \tag{3}$$

By substituting equations (2) and (3) into equation (1) we obtain the single-equation version:

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}Z_j + \gamma_{11}X_{ij}Z_j + u_{1j}X_{ij} + u_{0j} + e_{ij} \tag{4}$$

The regression coefficients gamma for the intercept and the regression slopes of the respondent and interviewer variables can be interpreted as in a standard multiple regression analysis. By combining respondent-level and interviewer-level explanatory variables in one analysis, the effects of interviewer variables are estimated conditional on the respondent variables.

In our study, the dependent variable Y is dichotomous (contact vs. noncontact or cooperation vs. non-cooperation, both coded 0/1). This leads to a multilevel logistic regression model, where the individual-level errors are fixed. The variance of the u_{0j} residuals reflects the variance of the intercepts and the variance(s) of the u_{*j} residuals reflects the variance of the regression slopes. The models in this study were estimated with HLM (Raudenbush, Bryk, Cheong, & Congdon, 2000), using numerical integration, which gives more accurate results for dichotomous outcomes than the Taylor series approximation used in most other software (Hox, 2002). The significance of regression coefficients is assessed with a t-test, and the significance of the interviewer-level variances is assessed with a chi-square test, as described in Raudenbush and Bryk (2002).

REFERENCES

AAPOR (2004). *AAPOR standard definitions. Final dispositions of cases and outcome rates for surveys.* Retrieved June 7, 2004, from http://www.aapor.org

Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality.* New York: Wiley.

Bretschneider, M., & Schumacher, J. (1996). DEMOS—Eine Datenbank zum Nachweis Kommunaler Umfragen auf dem Weg zum Analyseinstrument. *ZA-Information, 38*, 59–75.

Campanelli, P., & O'Muircheartaigh, C. (1999). Interviewers, interviewer continuity, and panel survey nonresponse. *Quality and Quantity, 33*, 59–76.

De Leeuw, E. D. (1992). *Data quality in mail, telephone and face-to face surveys*. Amsterdam: TT-Publikaties.

De Leeuw, E. D., & De Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 41–54). New York: Wiley.

De Leeuw, E. D., & Hox, J. J. (2004). I am not selling anything: 29 experiments in telephone introductions. *International Journal of Public Opinion Research, 16*, 464–473.

De Leeuw, E.D., & Van der Zouwen, J. (1988). Data quality in telephone and face-to-face surveys: A comparative analysis. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls III, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 283–299). New York: Wiley.

Goyder, J. (1987). *The silent minority: Nonrespondents on sample surveys*. Cambridge: Polity Press/Westview Press.

Groves, R. M. (1989). *Survey errors and survey costs*. New York: Wiley.

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

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

Hox, J. J. (2002). *Multilevel analysis. Techniques and applications*. Mahwah, NJ: Erlbaum.

Hox, J. J., & De Leeuw, E. D. (1994). A comparison of nonresponse in mail, telephone, and face-to-face surveys. *Quality & Quantity, 28*, 329–344.

Hox, J. J., & De Leeuw, E. D. (2002). The influence of interviewers' attitude and behavior on household survey nonresponse: An international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 103–120). New York: Wiley.

Hox, J. J., De Leeuw, E. D., & Kreft, I. G. G. (1991). The effect of interviewer and respondent characteristics on the quality of survey data: A multilevel model. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 439–463). New York: Wiley.

Morton-Williams, J. (1993). *Interviewer approaches*. Hanover NH: Darthmouth.

Purdon, S., Campanelli, P., & Sturgis, P. (1999). Interviewers' calling strategies on face-to-face interview surveys. *Journal of Official Statistics, 15*, 199–216.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear modeling*. Thousand Oaks, CA: Sage.

Raudenbush, S., Bryk, A., Cheong, Y. F., & Congdon, R. (2000). *HLM 5. Hierarchical linear and nonlinear modeling*. Chicago: Scientific Software International.

Singer, E. (2002). The use of incentives to reduce nonresponse in household surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 163–177). New York: Wiley.

Snijkers, G., Hox, J. J., & De Leeuw, E. D. (1999). Interviewers' tactics for fighting survey nonresponse. *Journal of Official Statistics, 15*, 185–198.

Stoop, I. A. L. (2004). Surveying nonrespondents. *Field Methods, 16*(1), 23–54.

BIOGRAPHICAL NOTES

Michael Blohm is Researcher at the European Center for Cross-Cultural Surveys at ZUMA, Center for Survey Research and Methodology in Mannheim, Germany, where he works in the ALLBUS Department which is responsible for the research program and the overall design of the German General Social Survey. He is mainly interested in survey methodology, especially nonresponse and interviewer behavior.

Joop J. Hox is professor of social science methodology at the Department of Methodology and Statistics of the Faculty of Social Sciences at Utrecht University. His research interests focus on two lines of work: data quality in social surveys, and multilevel modeling. In survey methodology, he has published on nonresponse problems and the effects of data collection mode and interviewers on various aspects of data quality. In multilevel modeling, he has written numerous articles, chapters and a monograph.

Achim Koch is a Senior Researcher at the European Center for Cross-Cultural Surveys at ZUMA, Center for Survey Research and Methodology in Mannheim, Germany. His major research interests are survey methodology and social monitoring.

Address correspondence to Michael Blohm, ALLBUS Dept., Center for Survey Research and Methodology (ZUMA), Postfach 122155, 68072 Mannheim, Germany, e-mail: blohm@zuma-mannheim.de