Pp. 103-120 in R.M. Groves, D.A. Dillman, J.L. Eltinge & R.J.A. Little (Eds) (2002). *Survey Nonresponse*. New York: Wiley.

CHAPTER 7

# The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison

***Joop Hox,*** *Utrecht University, The Netherlands*
***Edith de Leeuw,*** *MethodikA, The Netherlands*

**ACKNOWLEDGMENTS**

This chapter was written in collaboration with Mick Couper and Bob Groves (ISR), Wim de Heer (Statistics Netherlands), Vesa Kuusela (Statistics Finland), Risto Lehtonen (Statistics Finland), Geert Loosveldt (KU Leuven), Peter Lundqvist and Lilli Japec (Statistics Sweden), Jean Martin and Roeland Beertens (ONS), Sylvie Michaud and Tamara Knighton (Statistics Canada), Peter Mohler (ZUMA), Patrick Sturgis and Pamela Campanelli (LSE), Vasja Vehovar (University of Ljubljana), and Metka Zaletel and Eva Belak (Statistical Office of the Republic of Slovenia).

This project was possible thanks to the enthusiasm and work of researchers all over the world. The coordinators of the data collection in each country, who contributed and commented on this text, are acknowledged as coauthors. We also sincerely thank all others who in some stage of this project contributed with comments, support, and data. We especially thank Bonnie Brandreth, Pamela Campanelli, Bob Groves, Elizabeth Martin, and Jean Martin for their helpful suggestions. The usual disclaimers of responsibility apply.

We kindly acknowledge the permission of the health care services in Finland for using their data, and especially the members of the research team: Sirkka-Sisko Arinen (Social Insurance Institute of Finland), Unto Häkkinen (Research and Development Centre of Welfare and Health), Timo Klaukka (Social Insurance Institute of Finland), and Jan Klavus (Research and Development Centre of Welfare and Health).

Pp. 103-120 in R.M. Groves, D.A. Dillman, J.L. Eltinge & R.J.A. Little (Eds) (2002). *Survey Nonresponse*. New York: Wiley.

CHAPTER 7

# The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison

***Joop Hox,*** *Utrecht University, The Netherlands*
***Edith de Leeuw,*** *MethodikA, The Netherlands*

## 7.1 INTRODUCTION

Several studies have addressed the role of the interviewer in nonresponse. There is little evidence that interviewer attributes, such as age and sex, influence response rates. In addition there is hardly any consistent evidence that personality traits or characteristics play any role (cf. Groves and Couper, 1998), but there is some evidence that social skills do play a role (Morton-Williams, 1993). Social skills are associated with knowing and applying rules of accepted behavior and communication, and they can be can be trained (Argyle, 1969; Morton-Williams, 1993). Nonetheless, interviewer training devotes little time to skills or tactics for fighting nonresponse; an exception is Statistics Sweden, which allocates 35% of its training to nonresponse (Luiten and de Heer, 1994). To attain satisfactory response rates, interviewers must learn how to convince reluctant respondents in practice or by informal exchanges with other interviewers. It is therefore not surprising that interviewer experience positively influences response (Durbin and Stuart, 1951; Groves and Fultz, 1985; Couper and Groves, 1992; de Leeuw and Hox, 1996; Singer et al., 1983). What makes these experienced interviewers achieve higher response rates?

Morton-Williams (1993) analyzed tape recordings of survey introductions and identified successful strategies for obtaining respondent cooperation. Important factors were: appear trustworthy (e.g., always identify yourself immediately), appear friendly (e.g., smile, make a compliment), adapt to the situation at the doorstep, and

react to the respondent. Successful interviewers use their cultural knowledge and their local knowledge of the sample neighborhood (Groves and Couper, 1998) to optimize their approach and behavior.

Interviewer-respondent interaction is a central concept in the theoretical work of Groves et al. (1992). Using a completely different research method, Snijkers et al. (1999) were able to replicate the main conclusions of Morton-Williams (1993) and of Groves and Couper (1992, 1996): professional competence, social skills, tailoring of the introduction, and maintaining the interaction were all named as good tactics by the more successful interviewers.

A different perspective was introduced by Lehtonen (1996), who concentrated on interviewers' attitude towards persuasion strategies and the role of the interviewer. Those interviewers who have a strong belief in the importance of the voluntary nature of participation and feel negative towards strong persuasion strategies also had a higher probability of nonresponse in face-to-face interviews. The early work of Singer et al. (1983) found that interviewers' stated expectation about the ease of persuading respondents to agree to an interview correlated with interviewers' response rate in a telephone survey. Groves and Couper (1998) find a positive relationship between interviewers' confidence ("can convince almost anyone to respond") and response rate.

Building on the two perspectives—attitude and behavior—de Leeuw et al. (1997) investigated the influence of the interviewer on survey response in a face-to-face interview. They replicated Lehtonen (1996) and showed that interviewer attitude and response rate were correlated in face-to-face interviews. Interviewers with a positive attitude towards persuasion strategies attain a higher response rate. No significant differences between interviewers were found regarding self-reported doorstep behavior.

The questionnaire used by de Leeuw et al. (1997) was partly based on questionnaires used by Campanelli et al. (1997) in the United Kingdom and Couper and Groves (1992) in the United States. To broaden the scope of the study, an international research project was started in 1996 at the international workshop on household survey nonresponse in Mannheim, Germany. Three research questions are central in this international comparison: (1) "Do interviewers in different countries have different attitudes towards the interviewer role?"; (2) "Does interviewer attitude predict interviewer response rate within and across different countries?"; and (3) Does (self-reported) interviewer behavior add to the predictive power for interviewer response rate within and across countries?

## 7.2 METHOD

### 7.2.1 Data Collection Procedure

Ideally, a comparative study would use exactly the same survey topics and data collection procedures in all of the collaborating countries. Since this is impossible to achieve, we asked the contributors to provide data from several different sur-

veys, with the intention that the final collection of data sets shows sufficient variation on key characteristics so that wider generalization is warranted. The required data files were on interviewer level. The main dependent variable is the response on the interviewer level for a particular survey; the main independent variables are interviewer attitude and behavior (interviewer questionnaire); background variables are interviewer age, sex, and experience. Most participants provided response data from several surveys. In a few cases, interviewers participated in more than one survey. For interviewer level comparisons, the double interviewer records were removed. When analyzing response rates, they were included in the analysis.

### 7.2.2 Data Sources

The core of the data set were the data collected by de Leeuw et al. (1997), and Campanelli et al. (1997), who used comparable interviewer questionnaires. Couper and Groves provided data that omitted some attitude questions; Lehtonen provided his original data set (Lehtonen 1996), which did not contain questions on interviewer behavior. New data were collected in the following countries: Belgium, Canada, Finland, Germany, Sweden, Slovenia, and the United Kingdom. In all these cases, the standardized version of the interviewer questionnaire was used. (For a concise description of the available data sets, see the Appendix.) In sum, data was available from nine different countries and 32 surveys. The data came from both official statistics and research institutes, and both face-to-face and telephone surveys were included. Overall, 3064 interviewers approached 32,1947 potential respondents. Country, agency, and interview mode were added as background variables. The response rates were cooperation rates, corrected for noncontacts.

For most studies (except for the United States), we had also data on interviewer background. The majority of the interviewers (87%) were female. There were not many differences between countries regarding interviewer sex, with the exception of the United Kingdom and Germany, which had relatively more male interviewers. The average interviewer age was 47.6 with a standard deviation of 1.8 year. The United Kingdom and Germany had relatively older interviewers, whereas Slovenia had relatively young interviewers. Finland and Holland had the interviewers with most years worked (respectively 12 and 11 years on average at agency), whereas Slovenia had the youngest and least experienced interviewers (on average 3 years at agency).

### 7.2.3 Instrumentation

We started with the construction of indices for interviewer attitude and for avowed doorstep behavior. To avoid capitalization on chance, we randomly split the total data file into a file for exploration and a file for cross-validation. Step 1 was an exploratory factor analysis using standard techniques and Varimax rotation. The exploratory factor analysis was followed by a confirmatory factor analysis on the same data file. In step 2, the final confirmatory factor analysis model of step 1 was

tested on the data of the validation file. In all cases, the final model had a satisfactory fit in the cross-validation. In the third and last step, the complete file was used to calculate factor scores for each interviewer. Some data sets do not contain all the questions, which led to missing data. These missing data were assumed to be missing at random, and they were handled by direct estimation using maximum likelihood (Arbuckle, 1996). The variables were originally scored on a five point scale (1 = always to 5 = never). The final factor scores were calculated in such a way that a high value indicates positive agreement with or frequent behavior on the specific factor.

Exploratory analysis on the ten available attitude questions indicated three distinct factors or groups of related questions. One factor indicates the relevance of "persuasion" (e.g., reluctant respondents can be persuaded), the second factor emphasizes the "voluntary" nature of the interview (e.g., accept refusal of a reluctant respondent), and the last factor stresses the usefulness of sending another interviewer (e.g., when a respondent has refused, it is better to send a different interviewer). One question (more important to gain interest than seek a quick decision), did not have significant loadings on any factor and was dropped from the model. In his original study, Lehtonen (1996) distinguished the same two factors "voluntary" and "persuasion," which was replicated on a much larger sample of interviewers in our study.

The exploratory factor analysis was followed by a confirmatory factor analysis (Bollen, 1989). Since the sample size was large, the assessment of model fit was based on two goodness-of-fit indices that are less sensitive to sample size. The following fit indices for the model were used: the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and a test of close fit ($p_{close}$), which is a chi-square test of the hypothesis that the RMSEA is not larger than 0.05 (Arbuckle, 1996). For the validation file, the three-factor model also described the data well. Therefore, based on this three-factor model, we calculated factor scores for each interviewer (step 3). A high score on the factor "persuasion" indicates that an interviewer is persuasion-oriented. This factor includes questions that Groves and Couper interpret as interviewer confidence, which we view as a necessary component of persuasion orientation. A high score on "voluntariness" means that the interviewer is more oriented towards the acceptance of refusals and emphasizes the voluntary nature of participation. A high score on "send other" indicates that the interviewer does not prefer to try again, but thinks it best to send another interviewer. This can be seen as a form of "stepping back" (cf. Groves and Couper, 1998). It should be noted that the correlation between the three factors is extremely low; the highest correlation was between "voluntary" and "send other" (–0.14), which is not significant. The attitudes are essentially uncorrelated. The factor matrix for the total sample and the fit indices are depicted in Table 7.1. Factor loadings not present in the table have been fixed to zero.

The same three-step procedure was used for the ten questions on avowed doorstep behavior. Exploratory analysis indicated three distinct factors or groups of related questions. The first factor indicates the avowed use of "social validation" arguments (e.g., mention that most people participate, and emphasize the positive as-

**Table 7.1. Standardized factor loadings, attitude questions**

| | I Persuasion | II Voluntary | III Send other | |
|---|---|---|---|---|
| C1 | — | — | — | Gain interest, not quick decision |
| C2 | — | — | 0.52 | Send other interviewer if no time |
| C3 | — | — | 0.99 | Send other interviewer if no cooperation |
| C4 | 0.45 | — | — | Always persuade reluctant respondents |
| C5 | 0.53 | — | — | Reluctant respondents can be persuaded |
| C6 | — | 0.41 | — | Respect privacy of respondent |
| C7 | — | 0.64 | — | Accept refusals of reluctant respondents |
| C8 | 0.22 | 0.44 | — | Emphasize voluntary nature of participation |
| C9 | — | — | 0.10 | If reluctant, withdraw to return later |
| C10 | 0.44 | — | — | Caught at right time, most people respond |
| Exploration | | Validation | | Total file |
| CFI = 0.88 | | CFI = 0.86 | | CFI = 0.87 |
| RMSEA = 0.04 | | RMSEA = 0.05 | | RMSEA = 0.04 |
| ($p_{close}$ = 0.95) | | ($p_{close}$ = 0.82) | | ($p_{close}$ = 0.95) |

pects of participation). The second factor stresses the importance of obtaining a representative sample and uses "scarcity" arguments to persuade (e.g., this is the chance to give your opinion). The last factor concerns a "foot-in-the-door" tactic (e.g., ask to enter the home). The factor matrix for the total sample and the fit indices are depicted in Table 7.2.

The three-factor model gave a satisfactory fit on the validation sample. Again, factor scores were calculated for each interviewer on the total file. A high score on the factor "social validation" indicates that interviewers report that they often use arguments regarding social validation in their persuasion attempts. The factor scores on "scarcity" and "foot-in-the-door" can be interpreted in a similar way. There were substantial positive correlations between the factors. The correlation between social validation and scarcity was 0.41, between social validation and foot-in-door was 0.33, and between scarcity and foot-in-door was 0.54.

### 7.2.4 Analysis

We started our analyses with a comparison of interviewer attitude and avowed "doorstep" behavior across countries, using analysis of covariance. As there were some differences between countries in interviewers' age, sex, and experience, we used these variables as covariates in the analysis (The data sets from Groves and Couper and Lehtonen lacked some items. If there was at least one question for a factor, the factor score was estimated using direct maximum likelihood estimation. If all questions for a factor were missing, the factor score was set to missing and not used in the comparison. Since the multilevel procedures presented later use the factor scores as predictors, for these analyses the factor scores were also estimated if

**Table 7.2. Standardized factor loadings, avowed doorstep behavior**

| | I Social validation | II Scarcity | III Foot-in-door | |
|---|---|---|---|---|
| B3A | 0.65 | — | — | Say topic should interest them |
| B3B | 0.28 | — | — | Say you are not salesperson |
| B3C | 0.74 | — | — | Say most people enjoy interview |
| B3D | 0.74 | — | — | Say most people participate |
| B3E | 0.45 | 0.23 | — | Say this is the chance to give opinion |
| B3F | — | 0.45 | 0.12 | Explain how household was selected |
| B3G | — | 0.98 | — | Mention they represent other people |
| B3H | — | — | 0.53 | Make a compliment (only face-to-face) |
| B3I | — | — | 0.47 | Ask to go into the home (only face-to-face) |
| B3J | — | — | 0.38 | Begin asking a question |
| Exploration | | Validation | | Total file |
| CFI = 0.97 | | CFI = 0.95 | | CFI = 0.96 |
| RMSEA = 0.04 | | RMSEA = 0.05 | | RMSEA = 0.04 |
| ($p_{close}$ = 0.98) | | ($p_{close}$ = 0.68) | | ($p_{close}$ = 0.96) |

all questions for a specific factor were missing. This plug-in estimate is very close to overall mean substitution.)

In the next analysis step, we concentrated on the second and third research questions: which interviewer variables (attitude and behavior) predict interviewer-level response rate. As surveys were nested within countries (see the Appendix), we used a multilevel logistic regression model (cf. Goldstein, 1995). The separate levels in the analysis were interviewers (lowest level), surveys (second level), and countries (third level). In total, we had data from 3064 interviewers, employed in 32 surveys in nine countries. The logistic regression model analyzes response rates as proportions of the interviewer workload; differences among interviewers in workload are automatically included in the model as a weight factor. The estimation method employed was restricted maximum likelihood using second-order Taylor linearization and penalized maximum likelihood estimation (Goldstein, 1995).

## 7.3 RESULTS

### 7.3.1 Interviewer Attitudes in Different Countries

Three distinct factors were found that describe interviewer attitudes toward the interviewer role in response or nonresponse. The first is persuasion-oriented, the second reflects the interviewers belief in privacy and voluntariness, and the third reflects the specific feeling that it is better to send a different interviewer than have the same return (see Section 7.2.3).

When we compare interviewer attitudes over the nine countries on which data are available, we see some striking differences. First of all countries do differ on "persuasion" with The Netherlands as the extreme low case ($F = 36.2$, $df = 8$, $p = 0.00$). Ranking countries from high to low on persuasion orientation gives the following order: Germany, Slovenia, United States, United Kingdom, Canada, Finland, Sweden, Belgium, and The Netherlands. Correcting for interviewer characteristics (age, sex, and experience) and for survey organization (government versus nongovernment), the same conclusions hold. The only significant covariate is experience ($p = 0.01$); experienced interviewers tend to be more persuasion-oriented than inexperienced interviewers.

A different picture emerges when we look at "voluntariness." Countries do differ ($F = 73.2$, $df = 7$; $p = 0.00$), but now the most extreme case is Slovenia. Interviewers in Slovenia are strongly persuasion-oriented, and they also value voluntariness. When we rank the countries from high to low on voluntariness, we find the following order: Slovenia, Germany, Sweden, Belgium, The Netherlands, Finland, United Kingdom, and Canada. Correcting for interviewer variables does not change the conclusions. The only significant covariate is again experience ($p = 0.01$). Experienced interviewers tend to be less voluntary-oriented than inexperienced interviewers.

As Table 7.1 shows, the factor "voluntariness" is based on respecting respondents' privacy, accepting a refusal, and emphasizing voluntary nature of participation. Privacy and data protection has been the topic of much discussion in several European countries (e.g., Germany and Sweden); also, refusal conversion is still not generally accepted in Europe. The extreme position of Slovenia favoring privacy and voluntariness can be attributed to the fact that Slovenia is one of the emerging new countries, formerly belonging to Eastern Europe. Under the old communist regime in Eastern Europe, there was not much room for privacy. Therefore, it seems reasonable to assume that at present voluntariness and privacy are highly valued in former Eastern European countries both by interviewers and respondents. It would be interesting to investigate whether other former Eastern European countries show the same trend of valuing privacy and voluntariness.

Finally, we focussed on "send other." Interviewers who score high on send other, think it is a good thing to let another interviewer contact a reluctant respondent (no time, no interest) the second time. Again, there are significant differences between countries ($F = 60.6$, $df = 7$; $p = 0.00$). No significant effects were found for experience. The only significant covariates are age and government. Elder interviewers and interviewers working for official (government) statistical offices are more inclined to favor a renewed try by another interviewer. The rank order of countries is Sweden, Canada, Slovenia, United Kingdom, Belgium, Germany, Finland, and The Netherlands. Interviewers in Sweden are extremely favorable towards a recontact by another interviewer.

Figure 7.1 shows the distribution of interviewer attitudes across countries. Of course, there is also a large variation between interviewers within countries. For, instance an individual Canadian interviewer can score high on voluntariness and an individual Swedish interviewer high on persuasion.

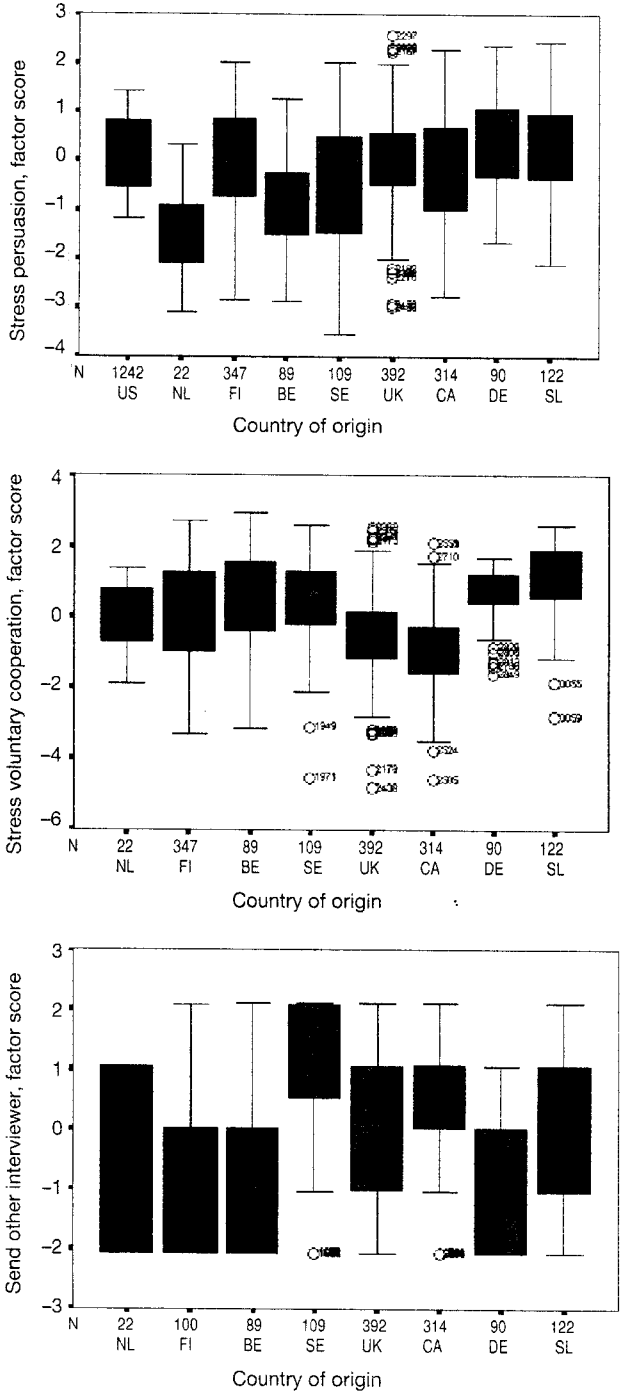



**Figure 7.1** Distribution of interviewer attitudes across different countries.

### 7.3.2 Avowed Interviewer Behavior in Different Countries

What interviewers do or say in their first contact with potential respondents will reflect organizational differences and individual differences. Available are self-reports of avowed behavior on average, not actual observations. But, both methods indicate the same trends, although the intermethod reliability is rather low (Campanelli et al., 1997). We found three distinct factors describing avowed interviewer behavior: social validation, scarcity, and foot-in-door (see Section 7.2.3).

Although there is an overall significant difference between countries on "social validation" ($F = 8.7$, $df = 8$; $p = 0.00$), this is mainly caused by the relatively high value of Canada and the low value of The Netherlands. Interviewers in Canada report that they often use arguments such as, "most people enjoy the interview," "most people participate," "the topic would interest you." When we correct for differences in interviewer characteristics (age, sex, and experience) and organization (government versus nongovernment) the differences between countries become somewhat greater. The only significant covariate is survey organization ($p = 0.00$); interviewers working in official statistics use fewer social validation arguments than interviewers in nongovernmental agencies. Ranking countries from high to low on avowed use of social validation arguments gives the following order: Canada, United Kingdom, Sweden, Finland, United States, Slovenia, Germany, Belgium, The Netherlands.

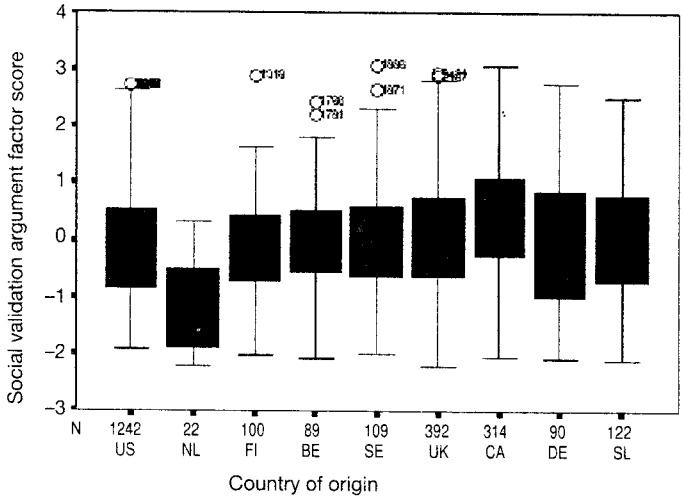
Looking at the use of "scarcity" arguments (mentioning they represent other people, this is *the* chance to give opinion, explain household selection), we again find a significant difference between countries ($F = 21.5$, $df = 8$, $p = 0.00$). In Canada, these arguments are relatively often used and in The Netherlands, Belgium, and Slovenia, relatively seldom. The only significant covariate is interviewers' sex ($p = 0.00$); female interviewers reported that they use the scarcity arguments more. Ordering the countries from high to low on the use of scarcity arguments gives: Canada, Sweden, Germany, United Kingdom, Finland, United States, Slovenia, Belgium, The Netherlands.

The use of foot-in-the-door or consistency arguments (e.g., begin asking questions) also differs between countries ($F = 24.3$, $df = 8$, $p = 0.00$). There are no significant covariates. The rank order of the countries from high (often use arguments) to low is Canada, Slovenia, Sweden, United States, Germany, Finland, United Kingdom, Belgium, and The Netherlands.

Figure 7.2 shows the distribution of avowed interviewer behavior across the countries. It clearly shows that although countries do differ in avowed interviewer behavior, there is also in some cases a large variation within countries.

### 7.3.3 Response Rates

The response rates differ considerably across interviewers. The average response rate is 0.82, with a standard deviation of 0.18. The country with the highest average response rate is the United States, at 0.91, and the country with the lowest average



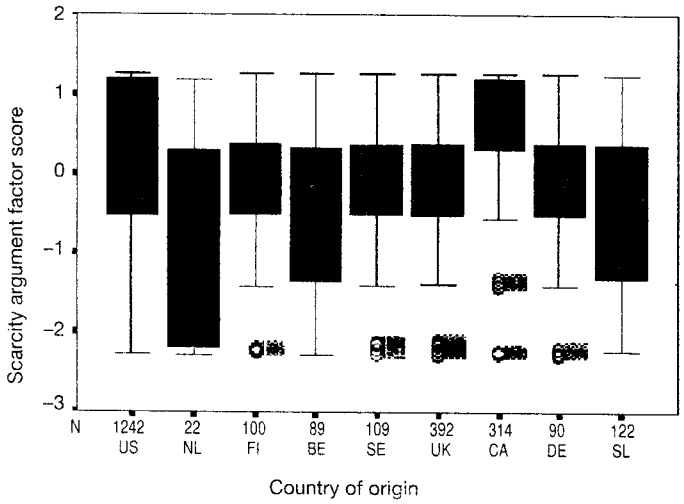




**Figure 7.2** Distribution of interviewer behaviors across different countries.

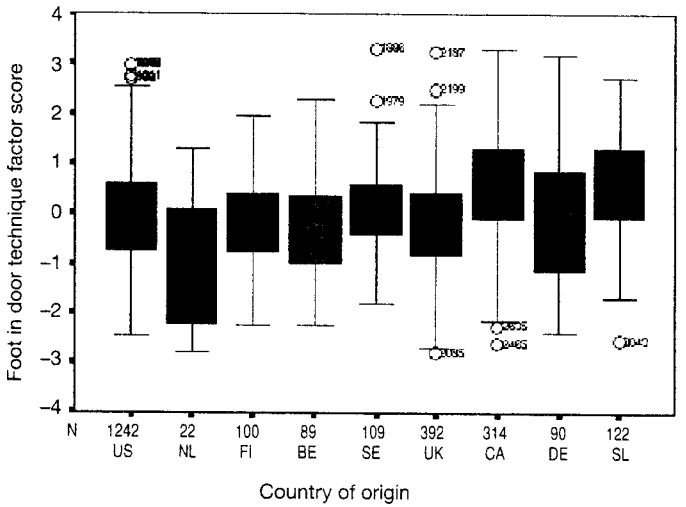



**Figure 7.2** Distribution of interviewer behaviors across different countries *(continued)*.

response rate is Germany at 0.52. The highest scoring individual study is the US Census health study, and the lowest is the face-to-face condition in the German social-economic study. However, these figures are not directly comparable because the studies differ in various characteristics. Therefore, we used multilevel logistic regression models to analyze the response rates, including available interviewer characteristics as explanatory variables and study characteristics as covariates.

### 7.3.4 Predicting Nonresponse Within and Across Countries

For predicting nonresponse, three sets of interviewer-related variables are available. The first set is individual interviewer attributes: age, sex, and amount of interviewing experience. The second set is avowed interviewer behavior: "social validation," "scarcity," and "foot-in-the-door." The third set consists of interviewer attitudes: "persuasion," "voluntariness," and "send other interviewer." Table 7.3 presents the results of four separate models. Each model uses one of the variable sets as predictors of nonresponse. The first model is the null model, which serves as a baseline for model comparison.

In logistic regression modeling, the variance at the lowest level (the interviewer level) is scaled to the variance of the standard logistic distribution, which is 3.29. The null model decomposes the total variance in response into three components. The proportion variance at the country level is 0.14, the proportion variance at the survey level is 0.10, and the proportion variance at the interviewer level is 0.76. Although the country level variance appears to be relatively small, it is not negligible. A country that is one standard deviation below the average has an expected response rate that is 15.9 percentage points lower than average.

**Table 7.3. Multilevel logistic regressions on interviewer response rates**

| Model/ predictor | Null model | Interviewer attributes | Interviewer behavior | Interviewer attitude |
|---|---|---|---|---|
| Constant | 1.25 (0.30) | 0.79 (.30) | 1.26 (.29) | 1.29 (.29) |
| Age (years) | | 0.01 (0.001) | | |
| Sex (1 = female) | | 0.03 (0.015) | | |
| Experience (years) | | 0.01 (0.001) | | |
| Factor Scores | | | | |
| Social validation | | | −0.02 (0.01) | |
| Scarcity | | | 0.003 (0.01)$^{ns}$ | |
| Foot-in-door | | | 0.03 (0.01) | |
| Persuasion | | | | 0.10 (0.01) |
| Voluntariness | | | | −0.02 (0.01) |
| Send other | | | | −0.02 (0.01) |
| $\sigma^2_{country}$ | 0.59 (0.37) | 0.62 (0.38) | 0.58 (0.37)$^{ns}$ | 0.55 (0.35)$^{ns}$ |
| $\sigma^2_{survey}$ | 0.41 (0.13) | 0.40 (0.12) | 0.41 (0.12) | 0.41 (0.12) |
| Deviance | 557.8 | 443.0 | 547.0 | 225.6 |

*Note:* ns = not significant.

If we consider the deviance of each model, which is a measure of misfit, we see that the model with the interviewer attitude variables has the lowest deviance. Thus, interviewer attitudes are the best predictors of response. Interviewer attributes are the second important set of predictors. The effect of avowed interviewer behavior is significant, but small.

Table 7.4 presents a multilevel regression model that contains as predictors all variables that have a significant contribution in Table 7.3. The regression coefficients are defined on the logistic scale. To facilitate interpretation, the last column in Table 7.4 indicates how many percentage points the average response changes if the predictor variable changes one unit.

The deviance of the final model indicates a good fit. All three attitudes are significant, but persuasion makes the largest contribution. Interviewers who are persuasion oriented achieve higher response rates. If interviewers go from one standard deviation below the average to one standard deviation above the average, their response increases on average by $2 \times 1.8 = 3.6$ percentage points.

Interviewer attributes do not appear to be very important, but age and experience are counted in years. Older interviewers have a somewhat higher response rate: a difference of 10 years results in 2 percentage points higher response rates. Experience counts for less: a difference of 10 years results in a predicted increase of 1 percentage point. Age and experience are, of course, correlated, but not perfectly, so the simultaneous effect of age and experience would be slightly higher. Sex does not have a strong influence: women have on average a 0.8 percentage point higher rate.

**Table 7.4. Final multilevel model for interviewer response rates**

| Model/ predictor | Null model | Final model | Effect in percentage points |
|---|---|---|---|
| Constant | 1.25 (0.30) | .80 (0.40) | |
| Age (years) | | 0.01 (0.001) | 0.2 |
| Sex (1 = female) | | 0.05 (0.02) | 0.8 |
| Experience (years) | | 0.01 (0.001) | 0.1 |
| Factor (Z-) Scores | | | |
| Social value | | –0.02 (0.01) | –0.3 |
| Foot-in-door | | 0.01 (0.01)$^{ns}$ | 0.1 |
| Persuasion | | 0.10 (0.01) | 1.8 |
| Voluntariness | | –0.02 (0.01) | –0.4 |
| Send other | | –0.01 (0.005) | –0.2 |
| $\sigma^2_{country}$ | 0.59 (0.37) | 0.58 (0.36) | |
| $\sigma^2_{survey}$ | 0.41 (0.13) | 0.39 (0.12) | |
| Deviance | 557.8 | 81.9 | |

*Note:* ns = not significant.

Avowed doorstep behavior hardly has any effect. It is interesting to see that the use of social-validation arguments even accounts for a slightly lower response. We will come back to this in the discussion.

The country level variance in the null model is 0.59. Translated back to percentages, this means that a country that is 1 standard deviation below the average in response rate achieves a 15.9 percentage points below average response. Given the large differences between countries, in our data as well as in the data analyzed by de Leeuw and de Heer (Chapter 3, this volume), it is interesting to compare the country variance to the effect of our strongest interviewer variables: interviewer age, experience, and persuasion. For this, the group interviewers is split on the median value for these variables into groups of interviewers that are young versus old, not experienced versus experienced, and low on persuasion versus high on persuasion. On average, the older interviewers have a response rate that is 2.8 percentage points higher than the younger interviewers. The more experienced interviewers have a response rate that is 0.2 percentage point higher than the less experienced interviewers. Finally, the more persuasion-oriented interviewers have a response rate that is 3.6 percentage points higher than the less persuasion-oriented interviewers. Although these differences are not totally negligible, it is clear that they cannot explain the much larger differences between the nine countries.

Finally, there are no significant variances for the regression slopes of the various predictor variables. This implies that there are no large differences in the effectiveness of the predictors across countries.

## 7.4 SUMMARY AND CONCLUSIONS

Two main questions in this study are whether interviewers in different countries have different attitudes and behaviors, and if such differences can explain differences in response rates within and between countries. Our results show that there are large differences between interviewers in all three attitude dimensions (persuasion orientation, stress voluntariness, and send other interviewer) and in all three doorstep behavior dimensions (social validation argument, scarcity argument, foot-in-door technique). If we inspect the patterns of the differences closely, there are some consistent results. In their attitudes, the United Kingdom and Canada cluster together, including the United States for the one dimension on which the U.S. interviewers are scored. Germany and Slovenia are also close to each other in their attitude scores. Extreme scores were found for The Netherlands, which is very low on persuasion, Belgium, which is simultaneously relatively low on persuasion and relatively high on voluntariness, and Sweden, which is extremely high in sending a replacement interviewer. In the avowed behaviors, most countries are very similar, except for The Netherlands, which scores on average below all other countries for all three dimensions, and Belgium and Slovenia, which score below average in using scarcity arguments. These differences are significant, and remain so if we control for available interviewer characteristics (i.e., age, sex, and experience) and for survey organization (i.e., government versus nongovernment).

We find that the interviewer attributes, attitudes, and avowed behaviors explain only a small part of the variation among countries. Part of the problem may be that there are differences between the studies, which confound the interviewer results. To assess the importance of this problem, we have done follow-up analyses of the response rates. In these follow-up studies, we included the interviewer characteristics reported in Table 7.4 as explanatory variables, with selected study characteristics added to the model as covariates. In general, these follow-up analyses corroborate our results. That is, response differences between countries continue to be important, and the available interviewer characteristics explain only part of them.

We also performed a sensitivity analysis to investigate the importance of the size of the contributed data set. There are large differences between the contributions, with as extremes the United States, with usable data from 1242 interviewers, and The Netherlands with 22 interviewers. We have reanalyzed the model reported in Table 7.4, with weights to compensate for the different sizes of the contributed data sets. The results of the follow-up and sensitivity analyses give us much confidence in the conclusions of our study. We conclude that the generalizability of our results is not impaired by our sample of countries.

The effect sizes we find for the interviewer variables are comparable to the interviewer effects found in other studies. For instance, Hox et al. (1991) report a small interviewer effect (intraclass correlation 0.02) on the response rate in telephone and face-to-face interviews. The small differences between interviewers were not related to interviewer characteristics, which included five personality

measures. Groves and Couper (1998, Chapter 7) discuss several studies that relate interviewer characteristics to the interviewer-level response rates. They conclude that there is no strong evidence for a relation between interviewer-level response rates and personality factors, but that interviewer experience and attitudes do have an effect. The effects we find are similar in magnitude, which corroborates their validity. The small but negative effect of using social validity arguments is contrary to expectance (Groves et al., 1992). However, Dijkstra and Smith (Chapter 8, this volume) also find a small negative effect for using social validation arguments. Social validation arguments may remind people too much of a sales pitch and may invoke the wrong ("oh no, they want to sell something") cognitive script. This association will lower the trust in the legitimacy of the interviewer and the survey and may result in the opposite of what is intended: a refusal (cf. van Leeuwen and de Leeuw, 1999).

A limitation of our study is that the studies differ in a number of important characteristics, such as the survey organization, the topic (or mix of topics in an omnibus survey), fieldwork conditions, and so on. This confounding inflates the between-country variance, and at the same time makes it more difficult to find strong interviewer effects. Thus, we suspect that country explains less variance than our analyses suggest, and that the interviewer effects should probably be somewhat larger than we report here. However, weighting the countries differently has hardly an effect on the regression coefficients. Also, none of the regression coefficients shows a significant variation across the nine countries, and at the country level we find no large residuals. The lack of variation in the regression coefficients across countries is reassuring because it indicates that the cultural setting (cf. Johnson et al., Chapter 4, this volume) does not have a strong effect on the efficacy of the interviewer characteristics in our analysis. So, the effect of interviewer experience and persuasion orientation is similar in the different countries.

## ACKNOWLEDGMENTS

This chapter was written in collaboration with Mick Couper and Bob Groves (ISR), Wim de Heer (Statistics Netherlands), Vesa Kuusela (Statistics Finland), Risto Lehtonen (Statistics Finland), Geert Loosveldt (KU Leuven), Peter Lundqvist and Lilli Japec (Statistics Sweden), Jean Martin and Roeland Beertens (ONS), Sylvie Michaud and Tamara Knighton (Statistics Canada), Peter Mohler (ZUMA), Patrick Sturgis and Pamela Campanelli (LSE), Vasja Vehovar (University of Ljubljana), and Metka Zaletel and Eva Belak (Statistical Office of the Republic of Slovenia).

This project was possible thanks to the enthusiasm and work of researchers all over the world. The coordinators of the data collection in each country, who contributed and commented on this text, are acknowledged as coauthors. We also sincerely thank all others who in some stage of this project contributed with comments, support, and data. We especially thank Bonnie Brandreth, Pamela

Campanelli, Bob Groves, Elizabeth Martin, and Jean Martin for their helpful suggestions. The usual disclaimers of responsibility apply.

We kindly acknowledge the permission of the health care services in Finland for using their data, and especially the members of the research team: Sirkka-Sisko Arinen (Social Insurance Institute of Finland), Unto Häkkinen (Research and Development Centre of Welfare and Health), Timo Klaukka (Social Insurance Institute of Finland), and Jan Klavus (Research and Development Centre of Welfare and Health).

## APPENDIX

**Overview of Surveys In Data File**

| Contact | Country | Study topic | Organization | Main mode | Reference | Number of interviewers | Number of responses | Year |
|---|---|---|---|---|---|---|---|---|
| Couper | United States | Expenditure (CE) | Census | Face-to-face | Groves and Couper (1998) | 206 | 10,634 | 1990 |
| Couper | United States | Health (HIS) | Census | Face-to-face | Groves and Couper (1998) | 139 | 18,322 | 1990 |
| Couper | United States | Crime (NCS) | Census | Face-to-face | Groves and Couper (1998) | 366 | 47,979 | 1990 |
| Couper | United States | Census particip. | NORC | Face-to-face | Groves and Couper (1998) | 106 | 2483 | 1990 |
| Couper | United States | Drug abuse | RTI | Face-to-face | Groves and Couper (1998) | 273 | 11,073 | 1990 |
| Couper | United States | Comorbidity | Michigan SRC | Face-to-face | | 152 | 5642 | 1990 |
| De Heer | The Netherlands | Living conditions (POLS) | Stat. Netherl | CAPI | de Leeuw et al. (1997) | 22 | 13,066 | 1996 |
| Kuusela | Finland | Health elderly | Stat. Finland | CAPI | | 118 | 1446 | 1998 |
| Kuusela | Finland | Labor force (LFS) | Stat. Finland | CATI | | 129 | 8462 | 1998 |
| Lehtonen | Finland | Health security | Stat. Finland | CAPI | Lehtonen (1996) | 122 | 1699 | 1995 |
| Lehtonen | Finland | Health security | Soc. Insurance | Face-to-face | Lehtonen (1996) | 95 | 534 | 1995 |
| Loosveldt | Belgium | Tourism/recreation | ISPO | Face-to-face | | 88 | 2364 | 1997 |
| Loosveldt | Belgium | Religion/moral | ISPO | Face-to-face | | 74 | 1445 | 1998 |
| Lundqvist | Sweden | Labor force (LFS) | Stat. Sweden | CATI | Japec et al. (1998) | 98 | 24,397 | 1997 |
| Lundqvist | Sweden | Living cond.(SLC) | Stat. Sweden | CATI | Japec et al. (1998) | 92 | 4667 | 1997 |
| Martin | United Kingdom | Labour force (LFS) | ONS | CAPI | | 122 | 40,483 | 1998 |
| Martin | United Kingdom | Housing (SEH) | ONS | CAPI | | 228 | 17,355 | 1998 |

*(continued)*

**Overview of Surveys In Data File** *(continued)*

| Contact | Country | Study topic | Organization | Main mode | Reference | Number of interviewers | Number of responses | Year |
|---|---|---|---|---|---|---|---|---|
| Michaud | Canada | Children (NLSCY) | Stat. Canada | CATI | | 314 | 39,427 | 1999 |
| Mohler | Germany | Social/economic | Infratest | CAPI (panel 1st wave) | | 87 | 1475 | 1998 |
| Mohler | Germany | Social/economic | Infratest | Face-to-face (panel 1st wave) | | 6 | 84 | 1998 |
| Sturgis | United Kingdom | Political tracking | NOP-research | Face-to-face | Campanelli et al. (1997) | 16 | 510 | 1996 |
| Sturgis | United Kingdom | Family resources | NatCen | CAPI | Campanelli et al. (1997) | 16 | 345 | 1996 |
| Vehovar | Slovenia | Nat. media (BGP) | Mediana | Face-to-face | | 22 | 609 | 1998 |
| Vehovar | Slovenia | Adult literacy (IALS) | Mediana | Face-to-face | | 56 | 4053 | 1998 |
| Vehovar | Slovenia | TV advertising (ODMEV) | Graliteo | Telephone | | 27 | 19,749 | 1998 |
| Vehovar | Slovenia | Brands (UA) | Graliteo | Telephone | | 25 | 36,903 | 1998 |
| Zaletel | Slovenia | Labor Force | Official Stat. | CAPI | | 14 | 1303 | 1997 |
| Zaletel | Slovenia | Labor Force | Official Stat. | Face-to-face | | 16 | 1292 | 1997 |
| Zaletel | Slovenia | Labor Force | Official Stat. | CATI | | 8 | 1453 | 1997 |
| Zaletel | Slovenia | Household budget | Official Stat. | CAPI | | 10 | 605 | 1997 |
| Zaletel | Slovenia | Household budget | Official Stat. | Face-to-face | | 8 | 539 | 1997 |
| Zaletel | Slovenia | Consumer | Official Stat. | CATI | | 9 | 1555 | 1997 |