Journal of Official Statistics
Vol. 5. No. 3, 1989, pp. 265–280

Interviewers and Data Quality in a Less Developed Setting[1]

William G. Axinn[2]

Abstract: The problem of non-sampling error in applications of survey methods to developing countries has been raised often. This paper examines interviewer characteristics as one possible source of non-sampling error. Intensive field supervision and ethnographic cross-checks were used in Nepal to generate data on the technical quality of data from interviews. Interviewers were assigned to interactions with respondents randomly and the study incurred zero non-response. Variations in the amount of technical errors, "don't know" responses and false information gathered during the interview are analyzed as indicators of data quality. The paper examines three hypotheses. First, that interviewers are more careful in irregular interactions. Second, that respondents provide better information when interviewed by someone with similar characteristics. Third, that respondents provide better information when interviewed by females. Same-gender, cross-gender, same–ethnicity and cross–ethnicity interviews are examined controlling for a variety of interviewer characteristics to test these hypotheses. The evidence provides some support for the conclusion that interviewers are more careful in irregular interactions, but no support for the idea that matching interviewers and respondents by characteristics improves data quality. When the sex of the interviewer has an influence, female interviewers produce higher quality data than male interviewers.

Key words: Interviewer effects; data quality; Nepal; surveys in less developed countries.

[1] An earlier version of this paper was presented at the 1988 meetings of the American Association for Public Opinion Research.

[2] Population Studies Center, Institute for Social Research, and Department of Sociology, University of Michigan, Ann Arbor, MI 48106-1248, U.S.A.

Acknowledgments: The Tamang Family Research Project is supported by grants from The National Institute of Child Health and Human Development (grant # RO1 HD22543-01) and The National Science Foundation (grant # SES-8607288). I wish to thank the Center for Family and Demography of the Institute for Social Research and the Population Studies Center of the University of Michigan for their support during the preparation of this manuscript. The paper benefited from comments made by Mary Jackman, Arland Thornton, Thomas Fricke, Robert Jackman, Robert Groves, Charles Cannell, Howard Schuman, Barbara Anderson, Brian Silver, Rachel Caspar, Cynthia Housh, and anonymous reviewers. I also wish to thank Thomas Fricke for his assistance in collecting the data for this paper and the interviewing staff of the Tamang Family Research Project for their willingness to cooperate through random assignment. While grateful for the assistance of all those mentioned, I alone am responsible for any errors contained in this paper.

1. Introduction

Applications of survey research in less developed settings require that survey methods be adjusted to suit the particular physical and cultural circumstances. Researchers working in less developed settings may make adjustments they believe appropriate without considering the implications of those changes for the quality of the resulting data. Often researchers adjust the pool of interviewers. Usually interviewers are selected to meet some standard of education and reading and writing ability, and researchers may also choose interviewers on the basis of other characteristics, such as gender. For example, under some conditions female interviewers may not be recruited because, "the arduous nature of field work . . . virtually preclude(s) the employment of females," (Nepal Fertility Survey (1976, p. 21)). By adjusting the pool of interviewers in this fashion, researchers may introduce non-sampling error into their results.

Interviewer characteristics such as age, sex, race, education, social status, and interviewing experience are considered an important source of non-sampling error in survey data (Sudman and Bradburn (1974)). For example, a recent investigation of gender effects on survey data collected in the United States showed interview responses to be partially dependent on the interviewer's gender (Groves and Fultz (1985)). Other current studies have shown that the characteristics of interviewers can bias survey data enough to alter our understanding of the underlying phenomena (Anderson, Silver, and Abram son (1987, 1988)). Furthermore, when addressing the special problem of adapting the survey technique to less developed countries, extra attention must be paid to the role of the interviewers in obtaining the data (Back and Stycos (1959)).

One of the largest applications of survey research in less developed countries, the World Fertility Survey (WFS) has been strongly criticized because, among other reasons, "surveys do not reach anything like their full potential because of less than perfect interviewing," (Caldwell, in Cleland and Hobcraft (1985, p. 46)). Stone and Campbell (1984) claim that the Nepal Fertility Survey (part of the WFS) generated a high degree of non-sampling error. They are concerned with biases produced by problems with linguistic comprehensibility, reinterpretation by respondents, the structure of the interaction and the specific context of the interview. Stone and Campbell particularly point to discrepancies between survey data and their anthropological cross check on inquiries to female respondents about their knowledge of contraception and abortion. In the cross check, female respondents claimed to be embarrassed to discuss such items, especially if the interviewers were male or others were present at the time of the interview. The effects of the presence of others on respondents' answers to sensitive questions have been carefully examined using data from the World Fertility Survey (Casterline and Chidambaram (1984)). However, little evidence is available concerning the effects of the interviewer's sex on data quality in surveys carried out in developing countries.

The study reported here recruited interviewers with varying characteristics including gender, even though the setting was the rugged, rural Hill and Mountain regions of the Nepalese Himalayas. The variability in interviewer characteristics, combined with random assignment of interviewers to interviews, allows us to examine explicitly the influence of characteristics of the interviewers on the quality of data.

2. Hypotheses

This analysis will examine variations in technical errors made by interviewers with particular characteristics who have all been exposed to the same training and field supervision. As Kahn and Cannell (1957, pp. 193–195) propose, the characteristics of interviewers affect the perceptions, attitudes, expectations, motivations and thus the behavior of both the interviewers and respondents, which in turn affects the resulting data. Note that it is the behavior of interviewers and respondents in reaction to certain characteristics, not the characteristics themselves which purportedly causes variations in the technical quality of the data.

The effect of interviewer characteristics on data quality depends partly on the norms of the society in which the interviews take place. Two norms related to Nepalese social interaction are likely to have an effect. First, as in most societies, males and females do not enjoy social equality. The Hindu Kingdom of Nepal is strongly influenced by the male dominated Hindu religion, and unrelated men and women do not generally interact socially (Bennett (1983)). The strength of this norm, however, varies among ethnic groups, and the respondent population examined here (Tamang) is not Hindu but instead Buddhist (Fürer-Haimendorf (1956)). The Buddhist Tamang have greater sexual equality than most ethnic groups in Nepal (Acharya and Bennett (1981)), but this does not mean they are completely egalitarian. Though weaker among the non-Hindu Tamang, social norms still impose constraints on cross-gender social interaction and are likely to affect interactions with interviewers.

Second, Nepal has several distinct ethnic groups with their own languages and social customs. While members of these groups do interact, close interactions such as intermarriage are quite rare. Thus ethnic differences between respondent and interviewer may also influence the behavior of each, affecting data quality.

This paper explores the hypothesis that interviewers will be more careful during irregular social interactions, committing fewer errors. Greater care, in this instance, means slower paced interviewing, more exact reading of the questionnaire, greater attention to the respondent's answer and fewer errors in following instructions (both those in the questionnaire and those given by supervisors). The paper examines both cross-gender and cross-ethnicity interviews as examples of irregular social interactions. Ethnicity, in this paper, is defined as the respondent's ethnic self-identification for respondents and the interviewer's ethnic self-identification for interviewers. While cross-gender and cross-ethnicity interactions may be regular in some situations, the structured survey interview, which calls for the interviewer to play a dominating role, it is an abnormal situation. In this abnormal situation cross-gender and cross-ethnicity interactions may be irregular. In general, the more unusual the social interaction, the more likely the interviewer is to be careful, and the more careful the interviewer is, the higher the quality of the data. However, findings from studies in other settings suggest that the respondent may provide more complete responses and accurate information when being interviewed by someone with similar characteristics (Bradburn (1983, pp. 312–315)). Thus, in some survey work considerable effort is taken to match the characteristics of interviewers to respondents in order to generate high quality data. If this matched-characteristics hypothesis is correct, we would expect to find that our cross-gender and cross-ethnicity interviews produce low quality data.

In addition to examining the effects of

interviewer/interviewee characteristics on data quality, this paper also looks at the sex of the interviewer per se. Females, clearly the subordinates in Nepalese society, are less threatening, dominating, and influential in social situations than males. Respondents may be more inclined to provide complete responses and accurate information to less forceful interviewers, so female interviewers may be more likely to obtain higher quality information than their male counterparts.

In order to assess the effect of interviewer characteristics, we shall explore three dimensions of the technical quality of data: first, we explore the number of technical errors interviewers committed, such as skipping questions, checkpoints, etc.; then variations in the number of "don't know" responses the respondent gave during the interview; finally, whether or not interviewers gathered information found to be false by an anthropologist's cross check.

Variations in technical errors are an outcome relating to the interviewer, not the respondent. The amount of care the interviewer takes should be related to the number of technical errors. Therefore, we expect to find fewer technical errors generated by cross-gender and cross-ethnicity interviews, but no relation between the interviewer's gender and the number of technical errors.

Variations in the number of "don't knows" and the amount of misinformation gathered are a function of both the respondent's willingness to provide and the interviewer's ability to elicit complete and accurate information. To the extent that the matched-characteristics hypothesis is correct, we expect to find more "don't knows" and misinformation among the cross-gender and cross-ethnicity interviews. Furthermore, to the extent that the hypothesis that females are less dominating is correct, we expect to find fewer "don't knows" and less misinformation among interviews conducted by females.

3. The Setting

The Tamang Family Research Project is a comparative study of two clusters of villages in Nepal. The data for this analysis are from a cluster of villages in the Central-Hill region. It was chosen to represent a rural Tamang community with some exposure to primary education and the wage labor economy. A variety of information related to social change and family processes was collected during a three-month period of residence in this cluster of villages.

Male and female interviewers worked in both settings. Both settings were rugged, with an array of physical dangers ranging from monsoon mud and leaches to snowy, treacherous mountain trails. Neither site had public sanitation, indoor plumbing, or transportation facilities better than a rocky mountain trail. While interviewers of both genders suffered some minor injuries and illnesses, their gender never imposed any problems on the progress of project efforts. Co-residence of male and female interviewers also failed to result in any difficulties for project staff, as sibling-like relationships developed among the interviewers.

Steps were taken by the study staff to establish a congenial relationship between themselves and respondents in each setting, including participation in local events, sharing of tea, picture-taking and some assistance with medical care. These efforts were aimed at keeping participation in the study and information quality high, and interviewers felt they were successful.

4. The Interviewers

Interviewers used in this study were recruited by spreading an open invitation to an

organizational meeting. This invitation was spread among local agencies which conduct surveys on a regular basis, college campuses, and branches of the government doing work in rural areas near our study sites. Those who attended this meeting were asked to write responses to some written questions and read some questions aloud. On the basis of their reading, writing, and speaking abilities, thirty individuals from this group were invited to participate in a one-week training course. This course involved five days of instruction in the survey process and in interviewing technique (based on the Survey Research Center's *Interviewer's Manual* (1976)) for eight hours per day and two days of supervised field interviewing in a rural area. Various tests and evaluations made throughout the week culminated in invitations to sixteen trainees to join the inter-

Table 1. Interviewers' characteristics and mean scores on data quality measures

Interviewer	Sex	Ethnicity	Education (years)	Previous experience (months)	Average technical errors	Average "don't knows"	Average misinformation
1	F	Tamang	10	0	0.45	0.95	0.33
2	M	Tamang	10	1	0.36	1.71	0.71
3	M	Non-Tamang	13	30	0.79	0.30	0.70
4	M	Tamang	14	0	0.58	1.20	1.00
5	M	Tamang	10	0	0.57	2.00	0.50
6	F	Non-Tamang	14	0	0.18	1.57	0.29
7	F	Non-Tamang	17	10	0.25	0.93	0.50
8	M	Tamang	10	0	0.30	0.05	0.30
9	M	Non-Tamang	12	11	0.20	1.05	0.45
10	F	Non-Tamang	14	0	0.24	0.66	0.00
11	F	Non-Tamang	17	3	0.18	0.65	0.30
12	M	Non-Tamang	15	24	0.38	0.28	0.70
13	M	Non-Tamang	15	10	0.12	0.87	0.33
14	F	Tamang	10	1	0.21	0.17	0.10
15	M	Tamang	10	0	0.63	1.35	0.33
16	M	Tamang	10	0	0.17	1.55	0.18

Key to variable lables

AV TEC ER = Average number of technical errors
AV DK = Average number of don't know responses
MISINF = Misinformation collected during interview (1 = YES, 0 = NO)
INT NUM = Interview number
I EXPER = Interviewer's previous interviewing experience (months)
I AGE = Interviewer's age in years
I EDUC = Interviewer's eduction in years
I GEND = Interviewer's gender (0 = MALE, 1 = FEMALE)
I ETHNIC = Interviewer Tamang or not (0 = YES, 1 = NO)
R GEND = Respondent's gender (0 = MALE, 1 = FEMALE)
MULT R = Multiple respondents gave answers (0 = NO, 1 = YES)
SAME SEX = Interviewer and respondent had the same gender (0 = NO, 1 = YES)

viewing staff. These 16 interviewers then resided at the study site for another four days of intensive training. Finally, even as interviewers began to work in the field, daily field supervision and one-hundred percent interview checking were used to give interviewers constant feedback on their work.

The interviewers themselves represented varying characteristics. Table 1 shows the composition of the interviewers by sex, ethnicity, educational attainment, and previous experience. Ten of the interviewers were male and six were female. Eight interviewers were Tamang and eight were non-Tamang. The mean educational attainment was slightly higher among the females, 13.6 years, than among the males, 11.9 years. The mean education among the Tamang, 10.5 years, was a good deal lower than the mean education among non-Tamang, 14.6 years, with all but one of the Tamang interviewers having only ten years of formal education. One result of this is a correlation of 0.82 between ethnicity and education (see Table 2). The implications of this high correlation will be discussed with the results. On the average, female interviewers had less previous experience than male interviewers, a mean of 2.3 months for females compared to a mean of 7.6 months for males. Tamang were also much less experienced than non-Tamang, with 0.25 months and 11 months average previous experience respectively. The interviewers' ages also ranged from 19 years to 39 years (not shown in tables).

Finally, an examination of mean values on the dependent variables arranged by interviewer is warranted. While some patterns do appear, note that no one interviewer is consistently either highest or lowest on any of the dependent variable measures. In fact no one interviewer is consistently among the top or bottom two. Also note that the highest and lowest values on each dependent variable never represent a large change from the next closest value. Thus, even though only 16 interviewers generated the set of interviews analyzed here, no one extremely bad, or extremely good, interviewer can be responsible for the findings reported below.

5. The Data

This village study interviewed all the residents over age twelve, not just a sample of residents. While the study interviewed every household in the village, for the purposes of this investigation the 145 cases analyzed are considered a simple random sample of the population of possible interviews with Tamang respondents living under the conditions this village was chosen to represent. The sample is considered random because respondents and interviewers were paired via a drawing of random numbers. Because of the extensive resources required to collect the data utilized for this analysis, it is restricted to a subset of the full study. The data used here come from the collection of census and genealogical information in one community of this cluster of villages. One hundred percent of that community's eligible residents participated in this data collection. There were no cases of missing data among the materials considered here.

The community these data come from is composed of 145 separate households. A household was defined as all those individuals who ate and slept regularly around a single hearth. Interviewers were assigned to households by drawing random numbers. At each household the interviewer was asked to complete a census and a series of genealogy sheets for every member of the household who either was married or had parents who did not reside in the same household. Each genealogy is four pages

long. The information asked on the genealogies include the names, birthdays, clans, current residence, place of birth and brief marital and fertility histories of the household member involved, his/her parents and his/her siblings. One member of each household was asked to provide this information as well as possible for every member of the household. Clearly, the larger the family, and the more complicated the family structure, the more difficult the task.

Once interviewers completed the genealogies, they were given to a supervisor. The supervisor checked the genealogies, tabulating on a separate form the number of technical errors, the number of "don't know" responses and the number of genealogy sheets. Interviewers were sent back to the households as necessary, but only data from their first visit to each household are analyzed here. Some time after the completion of the genealogies, an anthropologist who resided in the community part-time cross-checked the information. Any discrepancies discovered in the cross-check were also tabulated. Information about the interviewers was gathered informally over two months of co-residence with the investigator and with a short formal interview.

Each of the 145 households produced one interview. Characteristics of the interviewer are treated as characteristics of that interview, so for each of the three dependent variables, 145 separate cases are examined. While the household is the unit of analysis used here, there are actually only 16 interviewers. The interviews might also be thought of as clustered by interviewer. Significance tests discussed in the text are based on the assumption of a simple random sample. Estimates of the sampling error resulting from the clustered design were also calculated. Those calculations demonstrate that the design effects for the variables included in the ordinary least squares (OLS) models were in the ranges 0.81 to 1.22, with most falling below the value 1.00. In no case did this slight design effect change the *t*-value used in significance tests enough to make a relationship that had been significant under the assumption of simple random insignificant under the assumption of a sample clustered by interviewer.

The respondents were all of the Tamang ethnic group and either male or female, depending solely on who was home at the time of the interview and considered knowledgeable among members of the household on these matters. While interviewers were randomly assigned to the respondent's household, they might be expected to select respondents of the same gender. The data indicate this was not the case, with a correlation of only -0.02 between interviewer gender and respondent gender (see Table 2). Each of the 16 interviewers completed between seven and eleven interviews.

6. Methods

From each interview three measures of the quality of information were created. The first is the average number of technical errors an interviewer produces per interview sheet at a household (a continuous variable). The second is the average number of "don't know" responses per interview sheet (a continuous variable). The third is whether or not the interviewer collected any misinformation (a dichotomous variable).

It is clear that the more questions are asked, the more the interviewer is exposed to the risk of collecting poor quality data. To adjust for this problem, counts of technical errors and "don't know" responses are divided by the total number of questionnaire forms used at that household. The result is an average number of technical

Table 2. Correlation matrix of variables used to examine the influence of interviewer characteristics on the technical quality of survey data

	AV TEC ER	AV DK	MISINF	INT NUM	I EXPER	I AGE
AV TEC ER	1.00					
AV DK	0.18	1.00				
MISINF	0.18	0.05	1.00			
INT NUM	−0.38**	−0.21*	−0.17	1.00		
I EXPER	0.14	−0.23*	0.23	0.06	1.00	
I AGE	−0.03	0.08	0.10	0.03	0.19	1.00
I EDUC	−0.10	−0.13	0.10	−0.06	0.41**	0.13
I GEND	−0.15	−0.10	−0.24*	−0.07	−0.30**	−0.21*
I ETHNIC	−0.10	−0.16	0.03	−0.01	0.64**	0.41**
R GEND	0.06	0.10	−0.05	−0.03	−0.04	0.06
MULT R	−0.14	0.16	−0.02	0.13	−0.03	0.19
SAME SEX	0.14	−0.01	−0.01	−0.02	0.00	−0.03

	I EDUC	I GEND	I ETHNIC	R GEND	MULT R	GEND INT
I EDUC	1.00					
I GEND	0.32**	1.00				
I ETHNIC	0.82**	0.19	1.00			
R GEND	0.06	−0.02	0.02	1.00		
MULT R	0.04	−0.02	0.15	0.13	1.00	
SAME SEX	0.03	0.08	−0.01	−0.29**	−0.12	1.00

N of cases: 145 *$P < 0.01$ Two-tail **$P < 0.001$ Two-tail

errors and "don't know" responses per form. However, the collection of any false information during an interview is conceptualized here as being equally likely in all interactions. Rather than counts of incidents of collecting misinformation, this variable is operationalized as dichotomous (equal to 1 if any misinformation was gathered), so no such transformation is necessary.

To ensure valid comparisons, several control variables are used in models of the effect of interviewer gender and ethnicity on the quality of data. First, it is possible that the more interviews interviewers have completed the more likely they are to perform the task successfully. So, the number of previous interviews completed by interviewers while working on this study is used as a control variable. Likewise, interviewers' previous interviewing experience, education, and age are all entered as control variables. More experience with this study, more overall experience, and more education are expected to be associated with fewer technical errors, "don't know" responses, and less misinformation. However, no direction is assumed for the effect of age.

Multivariate statistical methods are used to estimate models of the three data quality criteria examined here. The first two criteria, the average number of technical errors and the average number of "don't know" responses, are both measured by interval level variables. These two variables are analyzed using ordinary least squares regression procedures. The third criterion of data quality, whether or not the interviewer gathered any misinformation during the interview, is measured by a dichotomous variable. This variable is also analyzed using ordinary least squares regression (a linear probability

Table 3. Means, standard deviations and ranges for variables used to construct models of interviewer effects

Variable	Mean	Standard deviation	Minimum	Maximum
Sex (Female = 1)	0.35	0.48	0	1
Ethnicity (Tamang = 0)	0.50	0.50	0	1
Interviewer and respondent had same sex (Yes = 1)	0.48	0.50	0	1
Controls				
Previous experience (months)	5.94	9.42	0	30
Interview number	5.14	2.78	1	11
Education (years)	12.46	2.57	10	17
Age (years)	24.86	4.91	19	39
Dependent variables				
Average technical errors	0.35	0.47	0	2.80
Average "don't knows"	0.93	1.10	0	6.50
False information was collected (Yes = 1)	0.41	0.49	0	1

model). However, because regression is known to be an imperfect estimation technique for linear probability models (Kmenta (1986)), both contingency tables and logistic regression estimation procedures are also used.

7. Results

Estimations of bivariate and multivariate models of the average number of technical errors, the average number of "don't know" responses, and the collection of false information are used below to evaluate the hypotheses outlined above. Table 3 displays the means, standard deviations, minimum and maximum values for each of the variables used in these estimations.

7.1. Technical errors

The average number of technical errors committed by interviewers varied by both interviewer gender and interviewer ethnicity. In a bivariate examination, male interviewers committed a mean of 0.40 such errors per questionnaire while females committed only 0.25 (not shown in tables). This

result may reflect the effects of cross-gender and same-gender interactions since interviewers who were the same sex as the respondent committed 0.42 errors and those who were the opposite sex committed only 0.29 errors (not shown in tables). Tamang interviewers committed more errors per questionnaire, 0.40 compared to only 0.30 among non-Tamang.

Next we test the hypotheses outlined above by simultaneously including these effects and series of control variables in a multivariate model. Table 4 displays the results of this estimation of the influence of interviewers' characteristics on the average number of technical errors. When interviewing respondents of the same gender, interviewers committed 0.12 more technical errors per genealogy sheet than when interviewing respondents of the opposite gender. This result is consistent with the hypothesis that when in unusual social situations interviewers do their job more carefully. Interviewers of a different ethnicity than that of the respondent (that is, non-Tamang since all respondents were Tamang) committed 0.40 fewer technical errors per sheet than interviewers of the same ethnicity as the respondent. This result is also consistent with the greater care in irregular situations hypothesis. Notice that the interviewer's gender had no significant influence on the number of technical errors he or she committed.

Table 4. Multivariate ordinary least squares estimation of the influence of interviewer characteristics on the average technical errors per form

	b	*t*
Interviewer sex (Female = 1)	0.02	0.162
Ethnicity (Tamang = 0)	−0.40	−2.099**
Interviewer and respondent had same sex (Yes = 1)	0.12	1.673*
Controls		
Previous experience (months)	0.02	3.257**
Interview number	−0.07	−5.501**
Education (years)	0.01	0.243
Age (years)	0.01	0.775
Constant	0.46	1.063
Adjusted R-square	0.23	
F	6.97***	
N	145	

*$P < 0.1$ (Two-tail) **$P < 0.05$ (Two-tail)
***$P < 0.005$ (One-tail)

As discussed earlier, interviewers' ethnicities and educational attainments are closely associated. When ethnicity is removed from this multivariate model, the effect of education on technical errors becomes negative and statistically significant, implying that more educated interviewers commit fewer errors. Thus, an alternative explanation for the ethnicity effect is that the more educated, non-Tamang interviewers simply made fewer mistakes. Unfortunately, highly educated Tamang were unavailable for this experiment, making it impossible to choose between these alternative explanations. However, removing the ethnicity measure has no effect on any of the other parameters of the model.

The greater the interviewer's previous experience, the more likely she or he is to make technical errors. This outcome is explained by the fact that the highly struc-

tured genealogy forms involved in this data collection are quite different than most survey questionnaires used in Nepal. Since interviewers with previous experience in Nepal are likely to have used very different questionnaires, their errors with this particular type of genealogy form are understandable. As expected, the more practice the interviewers had with the genealogies, the fewer errors they made. Finally, the characteristics of the interviewers account for 23% of the variance in the number of technical errors they make.

7.2. *Don't know responses*

The average number of "don't know" responses the interviewers obtained also varied by their gender and ethnicity. From a bivariate examination it is clear that male interviewers collected substantially more "don't know" responses, a mean of 1.02 per questionnaire compared to 0.78 among females (not shown in tables). It does not seem likely that differences between cross-gender and same-gender interactions explains this relationship since interviews where respondent and interviewer were of the same sex generated a mean of 0.93 "don't knows" while those in which they were of the opposite sex generated 0.94 (not shown in tables). A bivariate analysis also reveals a relationship between ethnicity and "don't knows", with Tamang collecting a mean of 1.10 per questionnaire compared to 0.76 among non-Tamang. Next, a multivariate model is used to test the effects of these factors simultaneously and control for other important factors.

Parameter estimates for the multivariate model of variations in "don't know" responses are displayed in Table 5. No significant difference in the number of "don't know" responses exists between interviews in which the respondent and interviewer were of the same gender and those in which they were of opposite genders. Likewise, no significant difference in the number of "don't know" responses exists between interviews in which the respondent and the interviewer were of the same ethnicity and those in which they were of different ethnicities. These two findings are inconsistent with the interpretation that carefulness in irregular situations is responsible for variations in the number of "don't know" responses col-

Table 5. Multivariate ordinary least squares estimation of the influence of interviewer characteristics on the average don't knows per form

	b	*t*
Interviewer sex (Female = 1)	−0.50	−2.002**
Ethnicity (Tamang = 0)	0.17	0.354
Interviewer and respondent had same sex (Yes = 1)	0.02	0.131
Controls		
Previous experience (months)	−0.04	−2.533**
Interview number	−0.08	−2.577**
Education (years)	−0.002	−0.037
Age (years)	0.02	0.743
Constant	1.26	1.172
Adjusted R-square	0.10	
F	3.21***	
N	145	

*$P < 0.1$ (Two-tail) **$P < 0.05$ (Two-tail)
***$P < 0.005$ (One-tail)

lected. The same findings are also inconsistent with the interpretation that increased willingness to cooperate among the respondents because they shared characteristics with the interviewer is responsible for variations in the number of "don't know" responses collected.

Female interviewers collected 0.5 fewer "don't know" responses per genealogy sheet than male interviewers. This finding is consistent with the hypothesis that respondents interviewed by females will be more inclined to provide complete responses than respondents interviewed by males. The reason for this may be that in male dominanted Nepal female interviewers are perceived as less forceful, dominating, and influential than male interviewers by both male and female respondents. Perceiving less threat, respondents may be more willing to furnish complete answers to the best of their ability.

As expected, interviewers with more previous interviewing experience with other studies and interviewers with more interviews completed working for this study gathered significantly fewer "don't know" responses. The age and education of the interviewers had no significant influence on the number of "don't know" responses obtained. This model explains 10% of the variance in the number of "don't know" responses collected.

It is important to note that dropping the ethnicity measure from estimations of the model described here had no effect on the other parameter estimates. Even though educational attainment and ethnicity are closely associated, the effect of education failed to be statistically significant when ethnicity was removed from the model. This finding suggests that, once other important effects are taken into account, neither education nor ethnicity has any significant effect on the number of "don't knows" collected.

One might wonder if particular interviewers received "don't know" responses at the expense of gathering more misinformation. This hypothesis was tested by regressing a dependent variable adjusted for the amount of misinformation, using the amount of misinformation as a regressor (both continuously and as a dummy), and through various crosstabular analyses. All demonstrated that the number of "don't knows" collected is not related to the amount of misinformation gathered and the amount of misinformation gathered does not mediate the impact of any of the other independent variables.

7.3. *Incorrect information*

Misinformation was treated as a dichotomous variable coded "1" if any misinformation had been gathered in an interview. Crosstabular analyses indicate only one characteristic of the interviewer is related to the collection of misinformation, the interviewer's gender. Female interviewers were less likely to collect misinformation than male interviewers. Male interviewers gathered misinformation in exactly 50% of their interviews while female interviewers gathered misinformation in only 25.5% of their interviews (Chi-square = 7.2, $p < 0.01$) (not shown in tables). This relationship is consistent with the hypothesis that female interviewers are less threatening to respondents and therefore respondents provide them with more accurate information.

Other interviewer characteristics, such as ethnicity, did not affect the gathering of misinformation. Tamang interviewers gathered misinformation in 39.7% of their interviews while non-Tamang gathered misinformation 43.1% of their interviews, an insignificant difference. Likewise, interviewers who were the same gender as their respondent gathered misinformation in

Table 6. Multivariate models of the probability an interviewer collects misinformation

	(1 = False information was collected)					
	Ordinary least squares (linear probability model)			Logistic regression		
	b	*t*			*b*	*t*
Interviewer sex (Female = 1)	−0.28	−3.266**			−1.27	−3.12**
Ethnicity (Tamang = 0)	0.08	1.031			0.39	1.08
Respondent and interviewer had same sex (Yes = 1)	0.001	0.019			0.01	0.04
Control						
Interview number	−0.03	−2.411*			−0.16	−2.37*
Constant	0.65	6.263**			0.65	1.42
Adjusted R-square	0.07		Degrees of freedom		4	
F	3.91***		Chi-square		15.40***	
N	145		*N*		145	

*$P < 0.05$ (Two-tail) **$P < 0.01$ (Two-tail) ***$P < 0.005$ (One-tail)

46.7% of their interviews while interviewers who were of a gender opposite that of the respondent gathered misinformation in 53.3% of their interviews, also an insignificant difference (not shown in tables).

A multivariate examination of the probability of collecting misinformation demonstrates results similar to those for the model of "don't know" responses discussed above (see Table 6). Both ordinary least squares and logistic regression estimations of the misinformation model are displayed in Table 6. Because the results of the model tested using a logit specification of the dependent variable were identical to the ordinary least squares results, the more interpretable ordinary least squares results (from the linear probability model) are discussed in the text.

Interviews in which the respondent and the interviewer shared the same gender were no more likely to generate false information than those in which the two had opposite genders. Tamang interviewers appear just as likely to gather misinformation as non-Tamang interviewers in the multivariate model. These findings fail to support the idea that the matching of respondent and interviewer characteristics influences the amount of false information collected during an interview.

The interviewer's gender, on the other hand, did have a significant influence on the likelihood of gathering false information.

Female interviewers were significantly less likely to gather misinformation than male interviewers. This finding is consistent with the hypothesis stated above that respondents perceive female interviewers as less threatening than males and are therefore more willing to provide accurate information when being interviewed by a female.

As expected, more practiced interviewers (those who had completed more interviews during this study) gathered significantly less misinformation. Finally the interviewer characteristics explain 8% of the variance in the probability of obtaining misinformation.

8. Conclusion

The findings described above are consistent with two conclusions. First, the evidence provides some support for the idea that interviewers are less likely to commit technical errors while interviewing in irregular social situations. Second, when conducting interviews, females, the subordinate gender in Nepalese society, are less likely to gather "don't know" responses and false information than males.

The analyses discussed here support the hypothesis that in abnormal social interactions interviewers will perform their task with more care than in normal interactions. In both cross-gender and cross-ethnicity interactions interviewers made fewer technical errors such as skipping questions, miss-marking checkpoints or asking extra questions. Interviewers may have been more successful during these irregular interactions because they conducted the interview more slowly, gave more attention to reading the questions exactly, and followed instructions properly. However, this conclusion must be tempered, at least regarding cross-ethnicity interviews, since education and ethnicity are closely associated. When ethnicity is removed from a model of technical errors education has a negative effect on errors indicating more educated interviewers commit fewer errors.

The results described above do not support the hypothesis that respondents will provide more complete and accurate information when interviewed by individuals with similar characteristics. Matching of either gender or ethnicity appeared to have no influence on variations in the number of "don't knows" or the amount of misinformation interviewers collected. The efforts study staff took to ensure a congenial relationship between themselves and the respondent population may be one reason for this outcome.

If that is the reason matching of interviewer and respondent characteristics had no influence on "don't knows" and misinformation, the finding that female interviewers gathered significantly less of both is that much more striking. Female interviewers may have been perceived as less threatening, dominating, and forceful than male interviewers by respondents. Because females were perceived as less threatening, respondents may have been more inclined to provide complete and accurate responses during the interview.

One strength of this study was the random assignment of interviewers to respondents, but one limitation was the relatively small number of interviewers. Further research, making use of larger numbers of interviewers and interviews, will certainly help to clarify the issues presented here. Also, the subject matter of this study, family relations, marital histories, and fertility histories may have favored female interviewers. Nevertheless, these topics are a common focus of surveys in less developed countries, and the findings reported here may be useful for improving data collection in such countries. Certainly further research, using a

broader range of substantive topics, will demonstrate the extent to which these findings can be generalized across survey topics.

This study shows that characteristics of the interviewer do have a significant effect on the technical quality of data gathered via survey interviews in Nepal, suggesting that the effect of interviewer characteristics on the outcome of interviews deserves attention in the application of survey methods to less-developed settings. Manipulation of the characteristics represented in a pool of interviewers by the investigator may provide the means for reducing non-sampling error in applications of survey research in such settings. Such a reduction in non-sampling error may in turn increase the validity of findings based on survey research carried out in less developed settings.

Furthermore, the findings reported here suggest that female interviewers should not be excluded from data collection efforts. In some situations the gender of the interviewer may have no effect on the quality of data. However, in situations where the gender of the interviewer does have an effect on the quality of the data, female interviewers provided higher quality data in every case examined here. Perhaps even under conditions of "arduous field work" the employment of female interviewers in the application of survey research should be considered. Certainly in Nepal there is evidence that female interviewers are better able than male interviewers to obtain high quality information from their respondents.

9. References

Acharya, M. and Bennett, L. (1981): The Rural Women of Nepal: An Aggregate Analysis and Summary of 8 Village Studies. Vol. II, Field Studies in The Status of Women in Nepal, Centre for Economic Development and Administration, Tribhuvan University, Kathmandu, Nepal.

Anderson, B.A., Silver, B.D., and Abramson, P.R. (1987): The Effects of the Race of the Interviewer on Race-Related Attitudes of Black Respondents in SRC National Election Studies. Paper presented at the Annual Meeting of the American Political Science Association, Chicago, Illinois, The Palmer House Hotel, September 2–6.

Anderson, B.A., Silver, B.D., and Ambramson, P.R. (1988): The Effects of Race of the Interviewer on Measures of Electoral Participation by Blacks in SRC National Election Studies. Public Opinion Quarterly, 52, pp. 53–83.

Back, K.W. and Stycos, J.M. (1959): The Survey Under Unusual Conditions, Methodological Facets of the Jamaica Human Fertility Investigation. Journal of Applied Anthropology, pp. 1–50.

Bennett, L. (1983): Dangerous Wives and Sacred Sisters. Columbia University Press, New York.

Bradburn, N.M. (1983): Response Effects. In Handbook of Survey Research, edited by Peter Rossi, James Wright, and Andy Anderson, Academic Press, New York.

Caldwell, J.C. (1985): Strengths and Limitations of the Survey Approach for Measuring and Understanding Fertility Change: Alternative Possibilities. In Reproductive Change in Developing Countries, Insights from the World Fertility Survey, edited by John Cleland and John Hobcraft, Oxford University Press, pp. 45–63.

Casterline, J.B. and Chidambaram, V.C. (1984): The Presence of Others During the Interview and the Reporting of Contraceptive Knowledge and Use. In Survey Analysis for the Guidance of Family Planning Programs, edited by John

Ross and Regina McNamara, ORDINA Editions, Liege, Belgium.

Fürer-Haimendorf, C.V. (1956): Ethnographic Notes on the Tamang of Nepal. The Eastern Anthropologist, 9(3–4), pp. 166–177.

Groves, R.M. and Fultz, N.H. (1985): Gender Effects Among Telephone Interviewers in a Survey of Economic Attitudes. Sociological Methods and Research, 14, pp. 31–52.

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

Kmenta, J. (1986): Elements of Econometrics. Second Edition, Macmillan Publishing Company, pp. 547–560.

Ministry of Health, His Majesty's Government (1976): Nepal Fertility Survey. First Report, World Fertility Survey Nepal Project, FP/MCH Project, Katmandu.

Stone, L. and Campbell, J.G. (1984): The Use and Misuse of Surveys in International Development: An Experiment from Nepal. Human Organization 43, pp. 27–37.

Sudman, S. and Bradburn, N.M. (1974): Response Effects in Surveys. Aldine, Chicago.

Survey Research Center, Institute for Social Research (1976): Interviewer's Manual. The University of Michigan, Ann Arbor.

Received September 1988
Revised August 1989