Contributions to Methodology and Statistics
A. Ferligoj and A. Kramberger (Editors)
Metodološki zvezki, 10, Ljubljana: FDV 1995

Interviewers' Effects in Telephone Surveys

The case of International Victim Survey

Eva Belak[1]and Vasja Vehovar[2]

Abstract

The interviewers' effects is expressed in terms of the interviewers' variance in the case of the International Victim survey. Different methods for variance calculation (Jackknife, Taylor series) were employed to estimate the interviewers' variance and the interviewers' intraclass correlation. It was found that these estimates were very unstable. This was mainly due to the relatively large and variable interviewer's workloads. However, despite this instability the interviewers' effects substantially distorted the precision of the estimates. For factual variables (victimization variables), the variance of the estimates was about two times larger than the variance where the interviewers' effects were ignored. For attitudinal variables (opinions about the survey itself) the increase was even higher.

1 Introduction

The interviewers are the agents who implement the entire survey design. They can also contribute to many errors in the survey. It is hard, however, to separate their effects from other measurement errors, coverage errors and nonresponse errors. Interviewers' effects on measurement error in surveys occur in four ways (Groves, 1989: 395):

1. Social psychologists sometimes view the interview as a structured social interaction, subject to many social influences. Thus, the demographic and socioeconomic characteristics of interviewers can affect the behavior of the respondents.

2. Interviewers can administer the questionnaire in different ways. Despite the instructions, they can reword questions, fail to ask some questions, or record the wrong answer. Attention to this problem is integral to a psycholinguistic view or a cognitive science view of the survey interview.

[1]Statistical Office of the Republic of Slovenia, Vožarski pot 12, 61 000 Ljubljana, Slovenia
[2]Faculty of Social Sciences, University of Ljubljana, P.O. Box 47, 61 109 Ljubljana, Slovenia

3. Even if the questions are read exactly as written, interviewers can emphasize different words or use different intonation in delivering the words.

4. When reacting to respondent's difficulties (e.g., failure to understand a question), interviewers assist the respondent in different ways, use different probing techniques, and thus yield variation in responses.

All four factors are reasons for errors associated with interviewers.

Several studies not only showed the existence of interviewers' effects, but also indicated that the magnitude of these effects differs from item to item. As Groves and Magilavy (1986) pointed out, this depends on whether the question was factual or nonfactual. They suggested, however, that the effects are actually a function of the amount of possible interviewer interference. This interference can take a variety of forms and can exist in both factual and nonfactual items. Questions which seems to be most susceptible to interviewer interference are those which concern sensitive topics leading to resistance in asking and responding, open-ended questions (especially those involving probes), and questions requiring a rating or subjective assessment from the interviewer. A number of studies have also found that interviewers' effects are related to characteristics of the interviewer and the respondent. Older respondents are most open to interviewers' effects (Groves and Magilavy, 1986). Interviewers' effects seem to be related to interviewer competence (determined in a variety of ways) and to the interviewer's prior expectations of the survey results. There is also evidence to suggest that younger interviewers are less susceptible to interviewer effects. Finally, several studies have shown that interviewers' effects can be the product of an interaction between interviewer, respondent and item characteristics (Tucker, 1983).

The interviewers' effects can be expressed in terms of the interviewers' variance which reflects the tendency for answers being correlated with interviewers. When such effects are present, the total variance of the estimates can be substantially larger than the sampling variance alone.

We will evaluate interviewers' variance in the case of the Victim Survey (1992) and compare the average magnitude of this effect with the studies of Groves and Magilavy (1986), Tucker (1983), Kish (1962).

2 Description of the study

The Victim Survey is part of an international project and was conducted by the Institute of Criminology at the School of Law in Ljubljana and by the Center of Methodology and Informatics at Faculty of Social Sciences in Ljubljana. The data were collected during three weeks in September 1992.

A mixed mode survey and a dual frame was used: out of 1000 households in the sample, 700 interviews were conducted by telephone (CATI), and 300 by face-to-face (CAPI). Thus, computer interviewing was used in both modes. The telephone numbers were randomly selected (simple random sample) from the directory of Ljubljana.

We analyzed the interviewers' effect only in the telephone part of the survey, where interviewers' assignments were randomized.

The interviewing team consisted of nine (experienced) interviewers, who received a two-day instruction course. They were familiarized with interview techniques, computers and software, and received detailed instructions of the procedures. The particular sensitivity of the topics was emphasized.

To obtain 700 telephone interviews, 834 residential numbers were contacted. There were 782 eligible numbers, with a nonresponse rate of 7 % and a refusal rate of 3 %.

Interviewers' variance was estimated for 28 target[1] variables: 18 factual and 10 attitudinal variables. Factual variables are those for which there is a knowable, verifiable answer (e.g., THEFT FROM CARS, etc.), unlike attitudinal variables which cannot be verified by external, objective data (e.g., "How often do you think you can trust the results of surveys", etc.).

3 Estimation of interviewers' effects

The goals of the study are:

- to estimate the interviewers' effects measured with the interviewers' variance and intraclass correlation ρ_{int};
- to estimate the actual increase in variance of the estimates due to the interviewers' effects;
- to estimate the precision of estimate $\widehat{\rho}_{int}$ for intraclass correlation ρ_{int}.

3.1 Model for Simple Random Sampling (SRS)

The interviewers are randomly assigned to the respondents. Each response from the sample can be written as (Kish, 1962: 108):

$$y_{ij} = y'_{ij} + A_i \, , \tag{1}$$

with the following notation:

- y_{ij} - observed response on a survey item for ith respondent belonging to jth interviewer.
- y'_{ij} - a response that would have been obtained if there were no interviewers' effects, $y'_{ij} = Y_{ij} + e_{ij}$, where Y_{ij} labels the true value and e_{ij} stands as the error term in the simple response variance model,
- A_i - average "effect" of the ith interviewer.

[1] All variables were dichotomized. Detailed description of the variables is presented in the paper by T. Kolenc "Explaining victims in Slovenia – Vulnerability and attractiveness of crime victims" and in the paper by M. Gnidovec and S. Kropivnik "Latent classification of survey respondents based on respondents' and interviewers' evaluations" in this volume.

We use a interviewers as a simple random sample from A interviewers. We assume that the sum of the interviewers' effects is zero for the population of A interviewers. Thus, the expected bias of the interviewers' effect is also zero.

The element variance, expressed in standard notation (Kish, 1965),

$$Var(y) \doteq S_y{}^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_{ij} - \overline{Y})^2$$

can be decomposed into two components:

$$S_y{}^2 = S_a{}^2 + S_b{}^2 , \tag{2}$$

where $S_a{}^2$ is the between-interviewers' variance component, and $S_b{}^2$ the within-interviewers' variance component in a standard analysis of variance.

The sampling variance, ignoring interviewers' effects and the finite correction factor, equals the standard SRS variance, where n is a sample size:

$$Var_{SRS}(\overline{y}) \doteq \frac{1}{n}S_y{}^2 = \frac{1}{n}(S_a{}^2 + S_b{}^2) . \tag{3}$$

With the introduction of proportions we have the following well known simple estimate of variance:

$$var_{SRS}(p) = \widehat{Var}_{SRS}(p) \doteq \frac{p(1-p)}{n-1} . \tag{4}$$

3.2 Correlated response variance

If we incorporate interviewers' effects, the variance of the estimate equals SRS variance multiplied by the design effect attributable to interviewers (Kish, 1965: 161):

$$Var_{int}(\overline{y}) = Var_{SRS}(\overline{y}) * Deff_{int}(\overline{y}) . \tag{5}$$

In a simple random sample design with equal workloads per interviewer the design effect can be expressed as a function of ρ_{int} and m:

$$Deff_{int}(\overline{y}) = 1 + \rho_{int}(\overline{y})(m-1) , \tag{6}$$

where:

m - number of interviews per interviewer (workload),

$\rho_{int}(\overline{y})$ - interviewer intraclass correlation; in the proceeding discussion we will use ρ_{int} instead of $\rho_{int}(\overline{y})$.

Intraclass correlation ρ_{int} can be approximated with:

$$\rho_{int} \approx \frac{S_a{}^2}{S_a{}^2 + S_b{}^2} , \tag{7}$$

where S_a^2 is the "between interviewer" and S_b^2 the "within interviewer" variance component. Thus, ρ_{int} is a proportion of element variance induced by interviewers. If we have a ρ_{int} of only $\rho_{int} = 0.01$ with the average workload of $\overline{m} = 78$ interviews we have $deff_{int}(\overline{y}) = 1.77$, a 77 percent increase of variance of the estimate. Obviously, even apparently small intraclass correlations can produce important losses in precision when combined with a large workload.

3.3 Weighting

In our case, however, we have different workloads per interviewer:

interviewer	1	2	3	4	5	6	7	8	9
workload	31	81	22	55	44	120	74	152	121

Because of unequal workloads we had to weight the data. This was done in the following way:

- first, we calculated the average workload - number of interviews per interviewer; in our case $\overline{m} = 78$;
- then we divided the average workload by the actual number of interviews completed by each interviewer. If interviewer i $(i = 1, ..., a)$ had completed m_i interviews, and the average workload was $\overline{m}$, then the weight for each interview equals $w_i^0 = m/m_i$;
- finally, we standardized the weights: $w_i = w_i^0 \,/\, (\sum_i^a w_i^0/n)$. In our case $n = 700$;
- the estimator is also changed in the case of weighting: instead of $\bar{y}$ for the estimation of the population average $\bar{Y}$ we have $\bar{y}_{wgh} = \sum_i \sum_j w_i y_{ij} / \sum_i \sum_j w_{ij}$.

Since we are dealing exclusively with proportions the corresponding notation will be p and p_{wgh} for the estimates of the population proportion P.

Through weighting we eliminate the impact of unequal workloads. As a consequence, the estimate of the target variable may change. We will express this change as the relative bias. We calculate the estimate in the following form:

$$bias(p) = \frac{|p_{wgh} - p|}{p_{wgh}} \times 100 \,. \qquad (8)$$

3.4 Estimation

To estimate the interviewers' variance, first we calculated the actual variance - including the interviewers' effects - of the estimate. Then, the design effect (deff) is calculated as ratio of the actual variance and the simple random sampling variance:

$$Deff_{int}(p) = \frac{Var(p)}{Var_{SRS}(p)} \,. \qquad (9)$$

From (6) and from the estimated design effect $deff(p)$ we calculate the intraclass correlation:

$$\widehat{\rho}_{int} = \frac{deff(p) - 1}{\bar{m} - 1} , \tag{10}$$

where $\bar{m}$ is the average workload per interviewer.

Of course, weights have to be used to compensate the unequal workloads, and a ratio estimate is needed for this purpose. We used procedures based on the **Taylor series** (Kish, 1965, chp. 6) for estimating the variance of ratio (proportion) p.

However, to estimate the sampling variance of $\widehat{\rho}_{int}$ we had to use the **Jackknife estimator.** For this purpose we divide the sample into k subsamples associated with each interviewer. We have $k = 9$ jackknife samples obtained by the omission of the corresponding interviewer. This enables the estimation of the variance of p (Wolter, 1985):

$$var(p) = \frac{k-1}{k} \sum_{i=1}^{k} (p_i - p)^2 , \tag{11}$$

where p_i is the estimate in the jackknife subsample where the interviewer i was omitted. After this, using (9) and (10) we can calculate $\widehat{\rho}_{int}$.

However, to calculate the variance of $\widehat{\rho}_{int}$ we first need to calculate $\widehat{\rho}_{int,i}$, for all $i = 1..9$ jackknife subsamples. For this purpose we create 8 additional jackknife replicates within each of the 9 initial jackknife replicates. Using the corresponding forms of (9), (10) and (11) we can estimate $Var_i(p)$, $Deff_i(p)$ and $\rho_{int,i}$. Finally, the variance of $\widehat{\rho}_{int}$ is estimated as:

$$var(\widehat{\rho}_{int}) = \frac{k-1}{k} \sum_{i=1}^{k} (\widehat{\rho}_{int_i} - \widehat{\rho}_{int})^2 . \tag{12}$$

Altogether we needed 72 jackknife subsamples to calculate the variance of $\widehat{\rho}_{int}$.

4 Results

4.1 Estimation of proportions p

As an example, let us look at the typical target variable THEFT FROM CARS (Table 1). If we assume simple random sampling (SRS), we see that 24.4 % of respondents were victimized. The standard error is $se(p) = 1.6$ %, so the confidence interval is [24.4 ± 3.2]% with ($\alpha = .05$). The design effect is $Deff(p) = 1$ by definition.

It is convenient to measure the relative precision of the estimate with the coefficient of variation CV(p). This will be estimated as:

$$cv(p) = \frac{se(p)}{p} . \tag{13}$$

We use the following rough approximations for critical levels of precision:

$cv(p) \leq 5$ % - preferable level,

$cv(p) \leq 10$ % - acceptable,

Table 1: Jackknife estimates for key victimization variables

variable	p	$se(p)$	deff(p)	bias(p) %	cv(p) %
THEFT OF CARS					
SRS	1.4	0.004	1.00		32
unweighted data	1.4	0.005	1.40	21.35	38
weighted data	1.8	0.010	4.18		58
THEFT FROM CARS					
SRS	24.4	0.016	1.00		9
unweighted data	24.4	0.028	3.00	7.37	15
weighted data	26.3	0.029	3.12		15
CAR VANDALISM					
SRS	32.9	0.018	1.00		8
unweighted data	32.9	0.035	3.89	1.29	16
weighted data	33.3	0.028	2.50		13
THEFT OF MOPED					
SRS	2.4	0.006	1.00		25
unweighted data	2.4	0.008	1.73	28.34	32
weighted data	1.9	0.006	1.49		34
THEFT OF BICYCLES					
SRS	13.9	0.013	1.00		11
unweighted data	13.9	0.011	0.74	5.38	9
weighted data	14.7	0.015	1.24		12
BURGLARY					
SRS	7.7	0.010	1.00		14
unweighted data	7.7	0.018	3.36	1.58	26
weighted data	7.6	0.016	2.43		22
ATTEMPTED BURGLARY					
SRS	7.4	0.010	1.00		14
unweighted data	7.4	0.014	2.09	6.32	21
weighted data	7.0	0.016	2.75		25
THEFT FROM GARAGES					
SRS	9.0	0.011	1.00		13
unweighted data	9.0	0.014	1.60	9.64	17
weighted data	10.0	0.015	1.82		17

Table 2: Design effect comparison: Ratio and Jackknife estimation for key victimization variables

variable	p	design effect ratio	design effect jack.
THEFT OF CARS			
unweighted data	1.43	1.26	1.40
weighted data	1.79	4.25	4.18
THEFT FROM CARS			
unweighted data	24.4	2.12	3.00
weighted data	26.3	2.48	3.12
CAR VANDALISM			
unweighted data	32.8	2.68	3.89
weighted data	33.3	1.97	2.50
THEFT OF MOPED			
unweighted data	2.43	1.42	1.73
weighted data	1.88	1.45	1.49
THEFT OF BICYCLES			
unweighted data	13.86	0.62	0.74
weighted data	14.71	1.23	1.24
BURGLARY			
unweighted data	7.71	2.97	3.36
weighted data	7.59	2.62	2.43
ATTEMPTED BURGLARY			
unweighted data	7.43	2.01	2.09
weighted data	6.94	2.92	2.75
THEFT FROM GARAGES			
unweighted data	9.00	1.49	1.60
weighted data	9.96	1.82	1.82

$cv(p) > 33$ % - not acceptable.

In the case of the above variable we have - assuming SRS, the coefficient of variation $cv(p) = 9$ % - an acceptable level of precision.

If we incorporate the interviewers' effects, proportion p remains the same, but the variance is much larger. The design effect is $deff(p) = 3$, so the actual variance is 3 times larger. The confidence interval is: $[24.4 \pm 5.6\ \%]$ at $(\alpha = .05)$ and the coefficient of variation is $cv(p) = 15$ %. Clearly, the actual estimates of our statistics are considerably less precise compared to those obtained by use of SRS formulas ignoring the interviewers' effects. It is also proper, of course, to eliminate - with weighting - the impact of unequal workloads. The weighted estimate can be interpreted as the estimate that would have been obtained if all interviewers had the same workloads. The relative bias in the case of THEFT FROM CARS was $rbias(p) = 7.37$ %. We can also observe that weighting additionally increases the variance.

Table 3: Estimate of intraclass correlation ($\widehat{\rho}_{int}$) and Jackknife estimate of standard error of $\widehat{\rho}_{int}$ - for key victimization variables

variable	$\widehat{\rho}_{int}$	se($\widehat{\rho}_{int}$)	cv %
THEFT OF CARS			
unweighted data	0.0032	0.0077	236
weighted data	0.0414	0.0462	111
THEFT FROM CARS			
unweighted data	0.0260	0.0206	79
weighted data	0.0277	0.0187	68
CAR VANDALISM			
unweighted data	0.0375	0.0213	57
weighted data	0.0194	0.0101	52
THEFT OF MOPED			
unweighted data	0.0085	0.0104	122
weighted data	0.0063	0.0101	159
THEFT OF BICYCLES			
unweighted data	-0.0043	0.0059	136
weighted data	0.0026	0.0129	501
BURGLARY			
unweighted data	0.0307	0.0255	83
weighted data	0.0216	0.0134	62
ATTEMPTED BURGLARY			
unweighted data	0.0146	0.0103	71
weighted data	0.0252	0.0129	51
THEFT FROM GARAGES			
unweighted data	0.0090	0.0124	138
weighted data	0.0105	0.0053	51

To establish the stability of our estimates, we calculated the variance and design effect with two different estimators. Results show (Table 2) that estimates obtained with the jackknife estimator are the same as or only slightly smaller than the estimates obtained by the Taylor series estimator.

4.2 Estimates of intraclass correlation $\widehat{\rho}_{int}$

We can observe from Table 3 that the coefficients of variation of $\widehat{\rho}_{int}$ are extremely high. This means that we measured the interviewers' effects very imprecisely.

4.3 Attitudinal variables

We can observe from Table 4 that attitudinal variables are even more susceptible to the interviewers' effects.

Table 4: Estimate of intraclass correlation ($\widehat{\rho}_{int}$) and Jackknife estimate of standard error of $\widehat{\rho}_{int}$ - for attitudinal variables

variable	$\widehat{\rho}_{\text{int}}$	se($\widehat{\rho}_{int}$)	cv %
DIFFICULTY			
unweighted data	0.0043	0.0094	219
weighted data	0.0132	0.0129	98
USFULNESS			
unweighted data	0.0177	0.0169	95
weighted data	0.0572	0.0223	39
APPLICABILTTY			
unweighted data	0.0567	0.0419	74
weighted data	0.0297	0.0178	60
ACCURACY			
unweighted data	0.0330	0.0294	89
weighted data	0.0751	0.0572	76
COOPERATION			
unweighted data	0.0122	0.0111	91
weighted data	0.1027	0.0975	95
UNDERSTANDING			
unweighted data	0.0276	0.0365	132
weighted data	0.1902	0.1881	99

5 Discussion

1. On average the actual variance of the estimator is more than 2 times larger than the SRS sampling variance ignoring interviewers' effects. Similarly, standard errors, coefficients of variations and confidence intervals are also increased. For a typical factual variable THEFT FROM CARS the confidence interval in the case of simple random sample is $[24.4 \pm 3.2]\%$, but if we incorporate the interviewers' effects the interval is much larger $[24.4 \pm 5.8]\%$. This, no doubt, reveals a considerable loss of precision.

2. Analysis of interviewers' variance enables us to consider the economic aspects of survey designs. The optimal design is obtained according to the following formula (Kish, 1962):

$$m_{opt} = \sqrt{\frac{C_a}{c} \frac{(1 - \rho_{int})}{\rho_{int}}}\,, \tag{14}$$

   where we used the following estimates:

   Ca - costs of hiring and training an interviewer ($\widehat{C}_a = 50DEM$),

   c - costs of completing one interview ($\widehat{c} = 3DEM$),

   $\rho_{\mathbf{int}}$ - intraclass correlation (median: $\widehat{\overline{\rho}}_{int} = 0.014$).

   The result implies an optimal workload of $m_{opt} = 34$ interviews. Thus, it would be optimal to have 20 interviewers. By comparing this with the actual workload of $\overline{m} = 78$ we can conclude that the existing workload was not optimal, but considerably too large.

3. Because of the relatively small number of interviewers and the variable workloads, the estimates of interviewers' variance are found to be very unstable. Despite this instability we can conclude that interviewers' variance substantially increased the sampling variance. To obtain more stable estimates of the interviewers' effects we have two alternatives:

   a) Altering the survey design to improve precision - chiefly by increasing the number of interviewers. Increasing the number of interviewers is not recommended in this case, because the hiring and training of larger numbers of interviewers may not be feasible. Also higher costs may occur.

   b) increasing precision through the replication of the survey. Another way to obtain useful knowledge about the average levels of interviewer variability is to repeat measurement of interviewers' effects over surveys. By cumulating results over surveys we hope to take advantage of the fact that averages of the same survey statistics have greater stability than does a single measure. Ideally, averages would be constructed from identical measures of the same populations, using the same interviewer selection and staffing rules, and the same statistics.

Table 5: Summary of overall findings

study	Range of $\hat{\rho}_{int}$	$\bar{\hat{\rho}}_{int}$
Kish	(-.03,.09)	.020
O'Muircheartaigh	(.00,.19)	.070
Feather	(-.01,.03)	.006
Freeman and Butler		.040
Tucker	(-.003,.008)	.004
Groves	(-.03,.09)	.020
Victim Survey	(-.005,.19)	.030

4. Let us compare the results with some other research. Groves and Magialavy presented some results (Groves and Magialavy, 1986: 253) from the past interviewers' variance research (see Table 5).

   It is not easy, of course, to compare different studies as there exist very specific circumstances. The first four cases refer to face-to-face interviews:

   - Kish's study was dealing with the job attitudes of factory workers;
   - O'Muircheartaigh and Marckwardt presented fertility statistics most susceptible to interviewers;
   - Feather was analyzing a health survey;
   - Freeman in Butler were dealing with the attitudes toward mental retardation (they found unusually high values of $\hat{\rho}_{int}$).

   The overall average for all these statistics for personal interview survey is $\hat{\rho}_{int} = 0.018$. The other two studies deal with telephone interviewing:

   - Tucker presented mean $\hat{\rho}_{int}$'s over 11 surveys dealing with political and social attitudes.
   - Groves presented mean $\hat{\rho}_{int}$'s from 10 telephone surveys (general opinion surveys).

   For the Victim Survey estimates of ρ_{int} for factual items the range is $[-0.0053, 0.0432]$ with mean $\bar{\rho}_{int} = 0.021$ and median $\rho_{int,Me} = 0.023$. For attitudinal items we obtained a range $[-0.0024, 0.1902]$ with mean: $\bar{\rho}_{int} = 0.055$ and median $\rho_{int,Me} = 0.037$. The overall mean was 0.03. We can conclude that our results are in accord with the results found in the above studies.

5. Perhaps the most important weakness of the past research is the relatively slight attention given to the stability of estimates $\hat{\rho}_{int}$. Only Groves and Magialavy (1986) treated this problem and they also found similiarly high $cv(\hat{\rho}_{int})$. They suggest that for more stable estimates, measurement should be repeated over several surveys.

6. The explanations for the high interviewers' effects might lie within the nature of interviewer training guidelines - wording, voice intonation and lack of attention given to sensitive topics. These can be improved by the following measures:

   - Interviewer **training** and **supervision**. Relatively untrained interviewers, who have been trained for less than a day, generally will not have mastered the basic techniques for standardized interviewing. So more through training is preferable. Also, the monitoring of the telephone interviews and tape-recording of the face to face interviews clearly help, although this might be costly. Of course, not all questions in surveys are significantly affected by interviewers; however, for questions in a survey that are likely to be significantly affected by interviewers, training and monitoring are very efficient measures.
   - **Selection of the interviewers** - for most surveys, getting the best trained and supervised interviewers possible, and giving them good instruments, is probably the best way of reducing interviewer-related error. Choosing interviewers with particular characteristics probably may be useful, although it is not the key instrument for controlling interviewer - related error.
   - One of the most important ways of minimizing interviewer-related error is to give interviewers **questions that can be administered in standardized way**. Although better question evaluation requires some time, as well as expense, better evaluation of questions in the laboratory and during the pretesting is probably the most cost-effective way of reducing interviewer-related error. At the same time, it is very likely to reduce response error for other reasons as well.
   - Finally, researchers should seriously consider **limiting the number of interviews per interviewer** in any survey. For those items that are most susceptible to interviewers' effects, the resulting reduction in total error can be significant. In addition, limiting the assignment sizes can reduce the effects of interviewer fatigue.

6 Conclusions

Let us summarize the key findings:

- there exists a considerable interviewers' effect in the Victim Survey;
- the actual precision of the estimates is significantly lower than the apparent precision ignoring interviewers' effects based on SRS formulas - these formulas overestimate the width of the confidence intervals by about 50%;
- the estimates of the interviewers' variance themselves are also subject to extremely large imprecision;

- the factual variables are less susceptible to interviewers' effects than are the attitudinal variables;
- the results relating to the interviewers' variance from the Victim Survey basically confirm the findings found in other studies.

If the Victim survey is to be repeated we would recommend the following:

- interviewers' effects should not be ignored;
- smaller workloads should be employed;
- stronger emphasis should be placed on interviewer's training and supervision.

References

[1] Belak E. (1994): *Anketa in anketar.* Diplomska naloga. Ljubljana: Fakulteta za družbene vede.

[2] Fowler F. J. Jr. (1991): Reducing interviewer-related error through interviewer training, supervision, and other means. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman (Eds.): *Measurement errors in surveys.* New York: Wiley, 259-278.

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

[4] Groves R. M. and Magilavy L. J. (1986): Measuring and explaining interviewers' effects in centralized telephone surveys. *Public Opinion Quarterly*, **50**, 251-266.

[5] Kish L. (1962): Studies of interviewers' variance for attitudinal variables. *Journal of the American Statistical Association*, **57**, 91-115.

[6] Kish L. (1965): *Survey Sampling.* New York: Wiley.

[7] Tucker C. (1983): Interviewers' effects in telephone surveys. *Public Opinion Quarterly*, **47**, 84-95.

[8] Wolter K.M. (1985): *Introduction to Variance Estimation.* New York: Springer-Verlag.