

The impact of interviewer fieldwork performance on the response process and data quality

Geert Loosveldt & Koen Beullens, Centre for Sociological Research, KU Leuven

Introduction

The two main categories of interviewer tasks are: first, the fieldwork activities concerning contacting respondents and securing their participation, and second, the interviewer behaviour during the interaction with respondents in order to obtain answers to survey questions.

The way interviewers perform the fieldwork activities of contacting respondents and obtaining their participation can have an impact on the quality of the realized sample. The interviewer's workload, the time needed to finish the assignment, and the response and contact rates can be considered as important interviewer fieldwork performance indicators. Based on these performance indicators, the description of an ideal interviewer is one who is willing to process a sufficient amount of interviews (workload), realizes a high response rate and contact rate, and completes the assignment within the agreed time. During the fieldwork period of a survey, the fieldwork management sometimes uses this 'ideal' interviewer for refusal conversion activities or to carry out additional clusters of interviews. The implicit assumption is that these interviewers can make a positive contribution to the quality of the realized sample. Previous research clarifies that there are differences between interviewers in terms of fieldwork performance. This means that interviewer effects can influence the fieldwork performance indicators.

Interviewer behaviour during the interaction with a respondent can also influence the quality of the responses. To reduce interviewer-related effects during the interaction, it is important that all interviewers perform their task in a standardized way (the key principle of standardized interviewing). The ideal interviewer systematically applies the key principle of standardized interviewing in an appropriate way. Research about interviewer effects on data quality indicators (e.g. item nonresponse, the tendency to select the same response category, etc.) shows significant effects. This means that there are significant differences between interviewers and that some interviewers produce better data quality than others.

Research question and hypothesis

The main research question for this paper concerns the relevance of some interviewer characteristics based on fieldwork performance in explaining interviewer effects on the data quality indicators related to the response process.

To explain interviewer effects on the data quality indicators, only a limited set of interviewer characteristics are used: experience, relevant interviewer background characteristics (gender, race, etc.), interviewers' opinions, and their attitudes to the tasks. These are all interviewer role-dependent characteristics. It is assumed that this category of characteristics is more relevant in explaining the differences between interviewers. Interviewer fieldwork performance characteristics are also interviewer role-dependent

characteristics and it can further be assumed that these characteristics are also relevant in explaining interviewer effects on data quality indicators.

Different hypotheses can be formulated about the relationship between fieldwork performance characteristics and data quality indicators. It should be noted that we assume there are significant differences between interviewers in the way they perform their fieldwork activities, and that consequently there are significant interviewer effects on the data quality. In fact, this first assumption is in contrast with what would be expected if all interviewers perform fieldwork activities in an optimal way and ask questions according to the key principles of standardized interviewing. In this ideal situation, significant differences between interviewers would not be expected and as a consequence there would be no correlation between fieldwork performance and the data quality.

- Interviewer satisficing hypothesis

The concept of 'interviewer satisficing' can be defined by analogy with 'respondent satisficing'. Interviewer satisficing means that interviewers do not make the necessary effort to carry out their tasks as prescribed. Interviewer satisficing can generate nonresponse error and measurement errors (Japac, 2008). During fieldwork activities, these interviewers choose the easy cases and minimize efforts to contact and persuade respondents. This probably results in lowered response rates and selective responses (low or moderate response rates, high non-contact rates and high refusal rates). Interviewer satisficing during the interaction with respondents results in shorter interviews. It can further be assumed that interviewer satisficing during interviews is negatively correlated with the quality of the resulting data, because respondents are relatively poorly stimulated and supported in terms of carrying out the cognitive answering process in an optimal way. This means that interviewer satisficing creates an atmosphere in which respondent satisficing behaviour is easily accepted. The interviewer satisficing hypothesis implies that the interviewer satisficing has a negative effect on interviewer fieldwork performance and on the data quality indicators that are related to respondent satisficing. This implies a positive relationship between the interviewer performance indicators and the data quality indicators.

— Cost-benefit hypothesis

The cost-benefit hypothesis is more specific than the satisficing hypothesis and focuses on interviewers' financial remuneration for work. Most interviewers are paid per realized interview. Under this system they can try to optimize their cost-benefit balance through compensating for the effort needed to achieve a high response rate (an increase in benefits) by minimizing their investment during interviews (e.g. shorter interviews, acceptance of respondent satisficing behaviour, etc.). In this context, a high response rate is more important than high-quality data. Based on this hypothesis, a negative relationship between the interviewer performance indicators and the data quality indicators can be expected.

Data

Datasets from the fifth round of the European Social Survey (ESS), organized in autumn 2010 and spring 2011 are used in the analysis. For all the participating countries, a representative sample of the population aged 15 and above was surveyed by means of

face-to-face interviews. The ESS questionnaire contains several core modules, which remain relatively constant from round to round (e.g. socio-demographic, national, ethnic and religious identity, socio-political values, etc.) and two rotating modules. The first rotating module in the fifth round concerns trust in the police and courts and the second module deals with work, family and wellbeing. In addition, a contact form is used to register in detail all the paradata from the fieldwork activities (e.g. dates and results of contact attempts, respondent's reason for refusal, type of dwelling, neighbourhood characteristics, etc.).

Interviewer fieldwork performance indicators and data-quality indicators

To test the link between interviewer performance during the fieldwork and the quality of the data, we first define the fieldwork performance and data-quality indicators to be used in our analysis.

- Data-quality indicators

Three data-quality indicators are used to evaluate the effect of the interviewers' fieldwork performance indicators on the data quality. There are two established indicators: (1) the percentage of don't know (DK) answers; and (2) the percentage of middle-category answers to questions with an 11-point scale. The third indicator is the interview length, which can be considered as a more global or indirect indicator of the quality of the interaction between interviewer and respondent. This is an indicator that is eminently suitable for testing the satisficing and the cost-benefit hypotheses. Interview length is measured in minutes, and values in the interval from 30 to 180 are used in the analysis.

With regard to the interviewer fieldwork indicators, we can make a distinction between output and process indicators.

- Interviewer fieldwork performance: output indicators

The output indicators are related to the final result of the fieldwork activities for a particular sample unit (e.g. response, refusal, non-contact, etc.). In the analysis, the co-operation rate and the contact rate are used. The co-operation rate is the proportion of respondents out of the total of the respondents and refusals. The contact rate is the proportion of contacted sample units that was contacted. It should be noted that sometimes more than one interviewer is involved in processing a sample unit (e.g. in a refusal-conversion process). Two alternatives can be used to deal with this situation in the analysis. The first option is to use only the interviewer with the final code. This means that the activities of other interviewers are not taken into account and consequently there is a loss of information. The alternative option is to use the best result for each interviewer for a particular sample unit. This means that all interviewer activities related to a sample unit are used, and that there are more records than sample units in the analysis. This results in lower effective response rates, but in an interviewer-oriented approach, this option is more appropriate than the former.

To measure the interviewer fieldwork performance related to the final results of the fieldwork activities, we specified a random intercept, logistic multi-level model. In this model, the respondent's binary response outcome (contact or not, co-operation or not) are

the measurements at the first (respondent) level and the interviewers are the second level. The independent variables used in the analysis are the respondent's type of dwelling and the interviewer's workload. This random intercept model results in unique parts of the intercepts for each interviewer. The interviewer fieldwork performance indicator concerning the dependent variable is the decile of the distributions of the random parts of the intercepts. A score of 0 means that the interviewer (compared to other interviewers) belongs in the group with the worst performance; a score of 9 means that the interviewer belongs in the group with the best performance.

For all the countries used in the analysis there are significant differences between interviewers in the realized co-operation rate after controlling for the type of dwelling and the interviewer's workload. Based on the deciles of the distribution of the unique parts of the intercepts, the interviewers obtain a score for the fieldwork indicator: *intcoop*

After controlling for the type of dwelling and the interviewer's workload, there are also significant differences between interviewers in terms of contact rates. This means that for all the countries, some interviewers realized significantly higher contact rates than did others. Based on the deciles of the distribution of the unique parts of the intercepts, the interviewers obtain a score for the fieldwork indicator: *intcontact*

- *Interviewer fieldwork performance: process indicators*

In contrast with the output indicators, the process indicators are based on information that is generated and becomes available during the fieldwork activities. Four process indicators are used in the analysis.

(1) Order. The most obvious and simple process indicator is the order of the interviews carried out by an interviewer. The order can be considered as an indicator of the interviewer's experience during the survey project and can have an impact on the interviewer's perception of his or her workload, cost and benefits. A higher order can stimulate an interviewer's satisficing behaviour, resulting in shorter interviews with consequently lower data quality.

(2) Initial response rate. The next process indicator is the interviewer's response rate after the first contact with all assigned sample units. Based on this initial response rate (*initialRR*), an interviewer can make an estimate of the responsiveness of the assigned sample units and make an initial cost-benefit analysis. It can be assumed that an initial low response rate will not stimulate interviewers to increase their investment in the fieldwork and interview activities.

(3) Line of least resistance. For each contact attempt with a sample unit, the response propensity can be estimated, based on information about previous contact attempts (number of contacts, language, etc.). The correlation between this response propensity and the decision of the interviewer to revisit the sample unit can be used as an indicator of the interviewer's tendency to follow the line of least resistance (*lineleast*). High values (correlations) indicate that interviewers have a tendency to revisit sample units with a high response propensity. This means that they follow the line of least resistance.

(4) Endure/perseverance. This represents the correlation between the number of realized interviews at a particular moment during the fieldwork period and the effort taken to contact the remaining sample units (endure).

Evaluation of the impact of the interviewer fieldwork performance on the data quality

It should be noted that the interviewer fieldwork performance indicators are those interviewer characteristics we use to explain the interviewer effects on the data quality indicators. Therefore, we specify a random intercept, multi-level model with the data-quality indicators as dependent variables and the interviewer fieldwork performance indicators as independent variables in the fixed part of the model. To evaluate the impact of the interviewers, the intra-interviewer correlation is calculated (IIC). The analysis is carried out separately for each country.

First we evaluate the impact of the two output indicators (intcoop and intcontact), then the process indicators are addressed (order, lineleast, initialRR and endure).

- *Effects of interviewer fieldwork performance: output indicators*

For all countries there are significant ($p < 0.05$) differences in interview length between interviewers, and in most countries the intra-interviewer correlation (IIC) is rather high (Table 1). This means that the interviewers explain a considerable part of the variability in interview length. However, the effect of the two output interviewer performance indicators is (with one exception) not significant. It is not even possible to detect a pattern in the direction of the effects. For some countries the effect is positive, for others it is negative. This means that the interviewer's scores for the co-operation rate and contact rate have no effect on the length of an interview and cannot be used to explain the differences between interviewers in terms of interview length. This means that there is no empirical support for either the satisficing or the cost-benefit hypothesis.

Table 1: Effect of the interviewer output fieldwork performance indicators on interview length

Obs	CTRY	IIC	interc	intcoop	intcontact
1	BE	0.070	58.02	0.043	0.408*
2	BG	0.213	61.52	-0.265	0.131
3	CY	0.220	68.23	0.750	0.530
4	DK	0.110	63.69	0.018	0.152
5	EE	0.000	82.05	-2.018	0.000
6	ES	0.011	69.56	-0.034	0.088
7	FR	0.210	67.31	0.200	-0.224
8	GR	0.320	72.39	-0.398	-0.466
9	HU	0.206	68.38	-0.101	0.146
10	IL	0.394	54.24	-0.609	-0.207
11	NL	0.126	61.97	0.272	-0.182
12	PL	0.020	75.85	0.270	-0.034
13	PT	0.081	56.50	-0.137	-0.026
14	RU	0.301	64.62	0.141	0.053
15	SI	0.107	47.95	0.244	0.303
16	SK	0.193	62.14	0.359	0.745
17	UA	0.535	69.83	-0.466	-0.409

* $P < 0.05$

We obtain similar results when the other two data-quality indicators are used (Tables 2 and 3). In Table 2, the effects of the interviewer output fieldwork performance indicators on the percentage of DK answers are presented. There are no significant differences between interviewers only for the Netherlands and Poland. For Belgium, France and Spain there are problems in estimating the standard error of the variance component. For the other countries, the differences between interviewers are significant but the value of the IIC varies between countries. The effects of the fieldwork performance indicators are not significant (with two exceptions: Estonia and Spain) and there is also no general trend in these effects.

Table 2: Effect of the interviewer output fieldwork performance indicators on the percentage of DK answers

Obs	CTRY	IIC	interc	intcoop	intcontact
1	BE	0.041	0.0041	0.0000053	-0.0000290
2	BG	0.261	0.0199	-0.0003951	0.0003193
3	CY	0.130	0.0122	0.0002698	0.0001335
4	DK	0.039	0.0043	0.0000760	0.0000408
5	EE	0.129	0.0084	-0.0003124	0.0008156 *
6	ES	0.000	0.0063	-0.0001171	0.0001684 *
7	FR	0.095	0.0035	0.0000770	-0.0000071
8	GR	0.159	0.0082	-0.0001021	0.0002128
9	HU	0.111	0.0125	-0.0001665	0.0001024
10	IL	0.136	0.0158	0.0001620	-0.0005258
11	NL	0.017	0.0062	-0.0001148	0.0001382
12	PL	0.005	0.0143	-0.0000190	0.0000158
13	PT	0.124	0.0172	-0.0002036	0.0001097
14	RU	0.259	0.0227	0.0002203	-0.0001713
15	SI	0.075	0.0131	-0.0001101	0.0001343
16	SK	0.283	0.0130	-0.0005871	0.0006043
17	UA	0.416	0.0364	-0.0006752	-0.0007203

* P < 0.05

In Table 3, the results for the third data-quality indicator are presented. Once again, we observe significant interviewer effects for most countries, though not for Belgium, Spain and the Netherlands. There are problems in estimating the standard error of the variance component for Denmark and Poland. There are also no significant effects from the interviewer's scores based on the co-operation rate and the contact rate.

Table 3: Effect of the interviewer output fieldwork performance indicators on the percentage of middle-category answers on questions with an 11-point response scale

Obs	CTRY	IIC	interc	intcoop	intcontact
1	BE	0.00763	0.20652	-0.000954248	-0.001310313
2	BG	0.12428	0.17414	0.001193694	-0.000926601
3	CY	0.07015	0.19002	0.003609705	-0.003109422
4	DK	0.00000	0.16777	0.000616761	-0.001373793
5	EE	0.02974	0.21865	0.002964629	-0.003885142 *
6	ES	0.00559	0.20756	-0.000612091	0.000120956
7	FR	0.09006	0.22017	-0.001819304	0.002917519
8	GR	0.06373	0.16934	-0.002812040	0.000123254 *

9	HU	0.06759	0.21229	-0.001116539	0.002465244
10	IL	0.03156	0.19057	-0.000699615	-0.001742221
11	NL	0.00342	0.16876	-0.000327611	0.000435264
12	PL	0.00000	0.20356	0.000098625	-0.000401901
13	PT	0.10070	0.21782	-0.002675641	0.000074234
14	RU	0.07092	0.19733	0.000341135	-0.000383248
15	SI	0.05477	0.16840	0.002301556	0.002958198
16	SK	0.13779	0.24063	-0.002359805	-0.000559461
17	UA	0.10240	0.18810	-0.000640313	-0.001451536

- *Effects of interviewer fieldwork performance: process indicators*

In Table 4, the results of the multi-level analysis (random intercept) are presented, with the four interviewer process fieldwork indicators as independent variables and interview length as the dependent variable. The results show that the effect of order is negative and significant for nearly all countries. This means that the interview length decreases when interviewers have carried out more interviews. This result is consistent with previous research by Olson and Peytchev (2007) and indicates that interviewer satisficing increases when interviewers carry out more interviews. Further, the effect of the initial response rate is negative and significant for half of the countries. This means that the interview length decreases when the initial response rate increases, which seems to support the cost-benefit hypothesis. Interviewers who realize a high initial response rate reduce the interview length. In this way, they optimize their cost-benefit balance. There are no consistent effects on interview length from either the tendency to follow the line of least resistance or from perseverance.

Table 4: Effect of the interviewer process fieldwork performance indicators on interview length

Obs	CTRY	interc	order	lineleast	endure	initialRR
1	BE	64.5890	-0.12830 *	-0.48952	-0.07722	12.0815
2	BG	68.1548	0.03922	-0.26249	-0.09487	-15.0591 *
3	DK	69.0317	-0.30242 *	-0.14846	0.08173	0.8647
4	ES	75.9179	-0.21707 *	0.23546	0.73693	-30.7083 *
5	FR	69.3003	-0.27179 *	0.88316 *	0.43814	-35.8451 *
6	GB	69.4871	-0.37959 *	0.01860	-0.37242	-37.2991 *
7	HU	72.7262	-0.17182 *	0.01487	-0.43188	-4.9761
8	NL	65.5961	-0.18300 *	-0.26442	0.19062	-35.0797 *
9	PT	56.2178	0.02383	-0.05064	-0.15596	-1.3644
10	RU	73.0482	-0.29804 *	0.15413	-0.55242	-8.3265

*p < 0.05

The results in Table 5 show less clear patterns that are less clear. There are a few significant effects. However, these effects are for different countries and related to different independent variables. In fact, there seems to be no empirical support for the effects of the process fieldwork performance indicators on the percentage of DK answers. Further, these indicators are not relevant in explaining the differences between interviewers in terms of the percentage of DK answers.

Table 5: Effect of the interviewer process fieldwork performance indicators on the percentage of DK answers

Obs	CTRY	interc	order	lineleast	endure	initialRR
1	BE	0.004628	0.0000217	-0.0000982	-0.0000647	0.0008
2	BG	0.019032	-0.0002553 *	-0.0004594	0.0001348	0.0135
3	DK	0.005812	-0.0000429	-0.0003659 *	0.0002746	0.0017
4	ES	0.004863	-0.0000110	0.0001237	0.0004518 *	-0.0031
5	FR	0.004928	-0.0000203	-0.0001470	-0.0000040	0.0001
6	GB	0.018055	0.0000184	0.0002830	-0.0006042	0.0013
7	HU	0.013396	0.0000098	-0.0001807	0.0000245	-0.0004
8	NL	0.005590	0.0001043 *	0.0000004	0.0000015	0.0009
9	PT	0.019209	0.0000180	-0.0011713 *	-0.0000148	0.0091
10	RU	0.019952	0.0000205	0.0002550	0.0000922	0.0091

The results in Table 6 show that the effect of order is mostly positive and is significant for some countries. Interviewers obtain more middle-category answers when they carry out more interviews. This result supports the satisficing hypothesis. In addition, the effect of the line of least resistance is in line with this hypothesis. The effect is mainly positive and indicates that more middle-category answers are recorded when the tendency to follow the line of least resistance increases. There is no clear pattern for the effects of endure and initial response rate.

Table 6: Effect of the interviewer process fieldwork performance indicators on the percentage of middle-category answers on questions with an 11-point response scale

Obs	CTRY	interc	order	lineleast10	endure10	initialRR
1	BE	0.19406	0.000834 *	-0.002614 *	0.001475	-0.0177
2	BG	0.17981	0.000217	-0.001243	0.001438	-0.0112
3	DK	0.16085	-0.000450	0.000712	0.000187	0.0488
4	ES	0.20222	0.000209	0.000710	0.001245	-0.0638
5	FR	0.18943	0.001610 *	0.002949	0.001328	0.0031
6	GB	0.18026	0.000535	0.000858	-0.000903	0.0157
7	HU	0.21059	0.000147	0.000468	-0.001924	0.0464
8	NL	0.17427	-0.000849 *	0.001453	-0.000189	-0.0668
9	PT	0.18443	0.000463 *	0.006444 *	-0.003162	-0.0140
10	RU	0.20736	-0.000635	0.002072	-0.003481	0.0216

* P < 0.05

Conclusion and discussion

For almost all countries, there are significant interviewer effects on the data quality indicators used in the analysis. The results clearly illustrate that there are significant differences between interviewers in terms of the realized data quality. In this paper, two hypotheses were formulated concerning the link between interviewers fieldwork performance and the quality of the obtained data. Based on the interviewer satisficing hypothesis, a positive relationship between the interviewer performance indicators and the data quality indicators was expected. The cost-benefit hypothesis predicted a negative relationship. The impact on the data quality indicators of the interviewer output performance indicators was not significant. This means that there is no support for either hypothesis. It appears that interviewer output fieldwork performance is not relevant in explaining the differences between interviewers. It is also possible that the assessment of interviewer fieldwork performance based on the output indicators is too inaccurate and not suitable for explaining interviewer performance during interviews. In the next step of the analysis, interviewer process fieldwork indicators are used. These indicators are more related to some characteristics of interviewers' fieldwork activities. Based on the results concerning the impact of these process indicators, there seems to be some support for the interviewer satisficing hypothesis. Although the use of process indicators seems to be more promising, the results are still not truly convincing. In fact, most empirical evidence suggests that the way interviewers organize their fieldwork activities and how successful they are during these fieldwork activities are not related to the quality of the answers they obtain during interviews.