

**Ann Carton - Geert Loosveldt, Department of Sociology, K.U.Leuven
Ann Carton, Van Evenstraat 2B - 3000 Leuven - Belgium**

Key Words: Interviewer Selection, Data Quality, Total Quality Management

1. Introduction

The concept of "Total Quality Management" has existed in the area of survey research since the beginning of the nineteen-nineties (Colledge & March, 1993, 1997; Biemer & Caspar, 1994; Diplo, 1997; Morganstein & Marker, 1997). Survey research can indeed be viewed as a production process, each stage of which has to meet quality standards. The ultimate objective of a survey is to formulate an answer to a research question. The attainability of this end goal depends on the successful completion of quite a few preliminary stages. The main processes are: conceptualisation and operationalisation of the questions to be researched, the drafting of a questionnaire, construction of the sample group, data collection, coding the information, preparing the final database, and finally analysing and reporting the findings. If one particular stage does not meet the pre-set requirements, corrective processes need to be introduced.

2. Research question and data

This paper deals with evaluating the quality of some aspects of a critical stage of the survey process, specifically with regard to data collection using face-to-face interviewers. In particular, we will evaluate the quality of the procedure for selecting interviewers and the relationship between the socio-demographic background characteristics of the selected interviewers and the quality of their performance. Interviewer performance will be judged on two different levels. The organisational level or the loyalty of the interviewer to the organisation on the one hand, and the data quality on the other hand. When the selection criteria cannot be met, the question arises whether it will result in a group of interviewers who perform less well in the two distinct areas of interviewer performance.

In order to answer the research question we will make use of two different kinds of data sources. The first source is a file which contains data about socio-demographic characteristics of the selected inter-

viewers, and about the organisational relationship of the interviewers with the network of interviewers. This file is kept and updated after each fieldwork by the ISPO-network of interviewers¹. Secondly, to evaluate the quality of data obtained by the interviewers, data from the second wave of the Belgian post-election survey of 1995 is available. Only the data from the Flemish region of Belgium will be used. From the beginning of September 1995 until March 1996, 156 trained interviewers conducted 2,099 face-to-face interviews of which 1,762 were panel respondents - interviewed for the first time after the elections of November 1991 - and 337 were new respondents. Of the 156 interviewers, 54 also participated in the first wave of the election study. All interviewers received a three-hour briefing dealing with the content of the questionnaire and practical issues. On average, the face-to-face interview took 1 hour and 20 minutes (Beerten et al., 1997:12-27).

3. Selection procedure for interviewers and practical application

For Fowler and Mangione (1990:96) an interviewer should possess at least some basic skills. The most obvious are good reading and writing skills. Morton-Williams (1993:9) specifically describes the attributes required for the job, judged from the vital role played by interviewers in the survey process. According to her, interviewers need to have a disciplined mind - in order to follow and respect rules - combined with a sense of initiative during the contacts with respondents. In that sense, Morton-Williams (1993:137) considers interviewing a multi-skilled activity. A skilled interviewer is consequently able quickly to strike up a good relationship with unknown respondents, to obtain their co-operation, and to demonstrate adequate interview behaviour by reacting appropriately to both adequate as inadequate behaviour displayed by respondents. The perfect interviewer would have a certain level of knowledge, the sense of duty required, the necessary social and communicative skills, and dispose of

¹ The Belgian Election Study is supported financially by the Federal Services for Technical, Cultural and Scientific Affairs, under the National Ministry of Science. The Flemish part of the survey was conducted by the Interuniversity Centre for Political Opinion Research (ISPO, Department of Sociology, K.U.Leuven). For the data collection, the ISPO could rely on its own network of face-to-face interviewers.

sufficient time to carry out the interviews within the pre-set period of time.

Apart from socio-demographic and psychological characteristics, more pragmatic considerations are important when selecting interviewers. The effective application of the above mentioned selection criteria also requires a rather specific selection procedure, viz. potential interviewers are invited for an interview. However, given the organisational set-up of the ISPO-network of interviewers, it was impossible to have personal interviews with all the applicants for the job of interviewer. Therefore, the recruitment and selection procedure for interviewers was mainly based upon the written curriculum vitae of the candidates in response to advertisements in newspapers. At first, the criterion level of education was used. A diploma of minimum higher general secondary education was assumed to meet the specified job qualifications. However, once confronted with the high turnover of interviewers, additional selection criteria had to be used. Those were mainly based upon pragmatic considerations and dealt with employment status and assumptions about availability for interviewing work. Interviewers should be willing to work freelance and willing to accept non-regular interviewer assignments. Earlier practical experience in working with face-to-face interviewers also showed that preferably no unemployed people or students should be selected. For those two groups interviewer work is only a temporary solution, an intermediate stage towards a more permanent job. Finally, because previous research has shown that experienced interviewers are not necessarily the best interviewers (Fowler and Mangione, 1990:134), former experience was not retained as a selection criterion.

In sum, selection criteria for the ISPO-network of interviewers referred to level of education and work situation. So much for the general guidelines. Now, let us return to the practical application of the criteria. However, given the working conditions of the job of interviewer (freelance, irregular assignments, no hourly rate), we hypothesise that it will not always be possible to fulfil the specified selection criteria for interviewers.

To test this hypothesis, we will make use of known socio-demographic details of the interviewers deployed during the post-election survey of 1995. When preparing the fieldwork, 172 interviewers accepted an assignment. The education criterion was fairly well followed. Only 9 per cent ($n=15$) did not have a diploma of minimally higher general secondary education. This was not the case for the work criterion. Nearly 19 per cent ($n=32$) of the interview-

ers were still students or were unemployed at that time. Only 13 per cent ($n=23$) worked as self-employed interviewers. More than 11 per cent ($n=20$) were housekeepers and 5 per cent ($n=8$) were retired. So, most of the interviewers had a full time job (35.5 percent or $n=61$) or were working part time (16.3 percent or $n=28$).

In line with expectations, the preconceived standard could not be consistently applied with regard to even those socio-demographic characteristics, due to a shortage of interviewers. The question we now need to address is whether the screening method used - restricted to the "feasible" selection criteria - resulted in a group of interviewers capable of producing good quality interview material, and whether the group of interviewers who failed to meet the criteria would produce less satisfactory data.

4. Evaluation of the interviewer performance

It is important that two separate levels are taken into account when assessing the procedure used for establishing the quality of interviewers' performance during all fieldwork: (1) a level related to the contents of the task and (2) the level of the way in which the interviewer relates with the network. Below, standards of quality are established for these two aspects of the interviewer's task and effectively applied to data gathered during post-electoral research in 1995.

A. The relationship with the ISPO-interviewer network

1. Operationalisation of the concept

The relationship of the interviewers with the network can be made operational by using details about various expectations held by the manager of the network regarding interviewers throughout various stages of the data gathering process. We shall briefly run through these stages, indicating whether the expectations are actually met each time and whether the network can rely on the dedication of its interviewers.

In preparation of the ISPO-network's involvement in a research project, the network manager firstly distributes the interview clusters among the relevant interviewers, who are selected by location. Interviewers are generally informed in writing of a new project. It is up to the interviewers to respond to the offer. The first aspect of the relationship of the interviewer to the network is consequently based on tracking the reaction to the offer, to know which interviewers are definitely not going to co-operate. When an interviewer reacts immediately in a positive way to an offer, it indicates that the network can count on the interviewer. The opposite is true for when the network needs to find out by itself whether an interviewer on its books wishes to co-operate. Interviewers evidently may not be able or willing to accept

an offer because of the particular status and conditions under which they work. In the following, we have only worked with the group of interviewers who initially reacted positively.

Since every research project is unique, a specific training session takes place before the actual fieldwork is started. Training is mandatory before interviews can be conducted. During the training session, lasting approximately three hours, the interviewers are given further background information on the aims of the research, and they usually complete a questionnaire themselves to become familiar with the questions. There is an opportunity to ask questions, and lastly clear, practical arrangements are made. For the ISPO election study, this particular preparatory session was held in various places throughout Flanders on different dates, in order to give every interviewer a chance to attend. An attendance record of interviewers at this training session constitutes the second noteworthy point. The network considers interviewers who stay until the end of the allocated training event to be reliable. The commitment of interviewers who stay away without sending their apologies is consequently considered to be less positive.

The ISPO-network called on supervisors to monitor in detail how interviewers actually performed during the fieldwork for the election study. Each supervisor follows a group of twenty to thirty interviewers from nearby, assessing the commitment of each interviewer by their response to a number of standard questions. Aspects giving positive ratings are: completing the interviews within the time scale, working through the requested number of questionnaires, correct use of the address list, thorough completion of additional forms, e.g. contact sheet (number of attempts made to make contact, reason for nonresponse, etc), completing the entire questionnaire, keeping the supervisor up-to-date with progress in the fieldwork, observing agreements made in the course of fieldwork, and carrying out additional interviews.

Based on this documented information, it is possible to work out a score for each interviewer reflecting how his or her relationship with the network can be typified. As a basis for the calculation of this score, points are allocated according to the reliability of the interviewer's commitment. Points are deducted for a lack of commitment, and the points remain unchanged when the commitment is considered to be neither positive nor negative. The interviewers are assessed on ten different aspects, and the

maximum score possible is seventeen². Interviewers are divided into three groups, depending on their score. A score of 0.85 or more (calculated in proportion to the maximum score) indicates that those interviewers are considered to be very efficient. The network can rely on their near unconditional commitment. Interviewers with a score between 0.60 and 0.85 do not always fulfil the network's expectations, but are still considered to be satisfactory. However, interviewers who achieve a proportion of 0.60 or less, do not relate correctly towards the network.

2. The relation with the network in practice

With regards to the election survey, the network could only rely on the unconditional commitment of a limited group of interviewers. Nearly 40% of the interviewers used for the survey behaved very correctly towards the network. Approximately 40% behaved satisfactorily. Nevertheless, nearly 18% of the interviewers did not comply with a number of formal, evident, and, in some senses, minimal arrangements. The data also show that when an interviewer accepts a job offer, it does not offer any guarantee that they will actually complete the project. After all, more than one in ten interviewers involved in this survey dropped out in the course of the fieldwork.

3. Interviewers' socio-demographic characteristics and their relationship with the network

With regards to the relationship between the socio-demographic background of the interviewers involved in the project and the organisational quality of their work, we expect to find that when the pre-determined selection criteria cannot be met, the interviewers will relate less correctly to the network.

However, none of the bivariate associations between socio-demographic characteristics and the way interviewers relate to the network appears to be statistically significant. Consequently, no prediction can be made on the basis of their background characteristics as to how interviewers will relate to the network. The association between previous interview experience before integration into the network and the relationship with the network may not be significant, but the percentages seem to conform to expectations. Experienced interviewers are not better achievers than inexperienced interviewers, in fact, the reverse appears to be the case. The socio-demographic characteristics relating to level of education and employment status are irrelevant selection criteria with regards to the interviewer's relationship with the network. An important observation, however, is that newly recruited interviewers seem to relate more correctly to the network. Only 11% of this group can be found among the

² The precise method used to calculate the score is available from the authors.

interviewers behaving incorrectly. Furthermore, interviewers who participated in surveys other than the election survey seemed to find it harder to keep to formal arrangements. Of the latter group, 27% did not stick to agreed arrangements.

B. Interviewers and data quality

1. Indicators for data quality

The quality of the data obtained by interviewers can be assessed on different levels. Here we will limit the assessment to the level of the registered answers (Billiet and Loosveldt, 1988). In the preparatory training session, the interviewers learn how they are expected to carry out their job, in order to obtain valid and reliable replies. Adequate interviewer behaviour is, after all, essential if complete and unbiased information is to be obtained. An important aspect in the assessment of interviewers is the amount of information they register for each respondent (Carton, 1995:147-149). The quantity is assessed with the help of three indicators relating to different components of the questionnaire used for the 1995 ISPO-election study.

In general, a user of survey data will delete those records which have a missing value (don't know, no answer). So, a first indicator can be formed by using the substantial information obtained from closed questions. On average, the respondents had to answer 303 closed questions. About 45% of the respondents answered 98% or more of the closed-ended questions in a substantial way (binary dependent variable: 0=less than 98%, 1=98% or more).

A second indicator for data quality is based upon the number of "don't know answers" in a specific part of the questionnaire. In particular, it concerned the respondent's perception of the position of six different political parties on eight different 11-point scales (non Catholic / Catholic; free enterprise / government regulation; immigrants same rights / fewer rights; conservation of environment / jobs; public order / freedom of speech; right to security / right to privacy; quality of life / economic growth; Flanders / Belgium must decide). As this part was considered to be a rather difficult part of the questionnaire, the interviewers were instructed to mention the possibility of a "don't know answer" when presenting the first scale. On average, the respondents answered "don't know" ten times. However, about 42% of the respondents never used the "don't know" answer possibility (continuous dependent variable: minimum 0 and maximum 48 don't know responses).

Finally, the third data quality indicator is the amount of useful information obtained from an open-

ended question. Open-ended questions also belong to the category of "difficult questions". This is not only true from the viewpoint of the respondent (Sudman et al., 1996; Foddy, 1993) but also from the viewpoint of the interviewer. In fact, the job of the interviewer is to ask for clarification, to probe, and to record the whole answer using the words of the respondent. We will make use of one open-ended question. Apart from the respondents who did not know for which party they voted for the Chamber of Representatives or who did not go to vote (voting is compulsory in Belgium), respondents had to answer the following question: "What is the most important reason why you voted for that party / why did you cast a blank or invalid vote". About 71% of the respondents gave one answer, the other 29% gave two or more answers (binary dependent variable: 0=one answer, 1=two or more answers).

2. Method to analyse the impact of the interviewer on data quality

To detect and analyse the effect of the interviewer on the three constructed indicators for data quality, we will follow a similar procedure. The first step checks if there is a significant interviewer effect after taking into account relevant respondent characteristics. In fact, as is shown by other research (e.g. Groves, 1989:441-445; Alwin and Krosnick, 1991:154-155) the obtained data quality can depend upon the respondent's age, level of education, and interest in the topic. Here, we will control for five different respondent characteristics: age, level of education, gender, political interest and whether the respondent was a panel or a new respondent. Furthermore, we have to take the survey design into account. As a matter of fact, the respondents were not randomly assigned to the interviewers. A method that deals very well with such a kind of design is multilevel analysis (Hox, 1994:300-304). The Mln software was used to fit a regression analysis with two binary dependent variables (substantial information from closed questions and the open-ended question reason for voting) and one analysis with the number of don't know answers which was considered to be a continuous variable (Rasbash and Woodhouse, 1996; Yang et al., 1996)³.

As a second step, when there is a significant interviewer effect, interviewer characteristics will be introduced into the analysis (stepwise method) in order to explain the effect of the interviewer on the data quality. First of all, we will make use of the socio-demographic selection criteria. However, given the results from other research (e.g. Fowler and Mangione, 1990:104-105;

³ We are grateful to Jan Pickery (Dept. of Sociology, K.U.Leuven-Belgium) for his useful help when fitting the different models.

Groves, 1989:405; de Leeuw and Collins, 1997:213; Hox et al., 1991:440-442; Pickery and Loosveldt, 1998), we hypothesise that we will not find a significant effect of the socio-demographic selection criteria on the quality of the interviewer performance. In other words, the fact of not observing the selection criteria will not have a detrimental impact on the three aforementioned indicators for data quality. Apart from these socio-demographic interviewer characteristics, we will make use of four other interviewer characteristics: former experience before the interviewers were accepted into the network, the number of interviews realised by the interviewers during the post-election survey, whether the interviewer participated in the first post-election survey of 1991 or whether it was a newly recruited interviewer, and the relation of the interviewers to the network.

3. The effect of socio-demographic and "other" interviewer characteristics on data quality

In general, the effect of the respondent characteristics are consistent with the results of other studies. The higher the level of education, the more the respondents are interested in the topic, the more substantial information the interviewers can get. However, for none of the three constructed indicators for data quality did we find a significant effect of the respondent's age. Furthermore, female respondents give less substantial information.

The results indicate that after controlling for different respondent characteristics, there still remains a significant interviewer effect. Some interviewers obtain more/less substantial information from closed questions, more/less from an open-ended question and more/less don't know answers for a particular part of the questionnaire than other interviewers. As hypothesised, the socio-demographic interviewer characteristics cannot explain the interviewer effect. However, there is one exception to this conclusion. The self-employed interviewers obtain less don't know answers. Therefore, it should be useful to investigate the way the interviewers handled this particular part of the questionnaire. Unfortunately, this kind of analysis is only possible when there are tape recordings of the interviews available. Another more important result is the effect of interviewers who also participated in the first wave of the post-election survey. Those interviewers obtain less information from closed questions and get more don't know answers. So, the quality of the data collected by those interviewers is not as good as that obtained by the other interviewers. This result indicates that it seems better not to deploy the same

interviewers in panel surveys - a result that is consistent with the analysis of the effect of interviewer and respondent characteristics on refusals in the same post-election survey. Loosveldt et al. (1997) show that the interviewers of 1991 obtain more refusals than other interviewers. This is another indication that the same interviewers should not be used in panel surveys. For the data quality indicator information from one open-ended question, only the number of interviews done by the interviewer has a rather small but still significant effect. The more interviews the interviewer realised, the more reasons he/she wrote down. Finally, also important to note is that there is no significant effect for the relationship of the interviewers with the network on any of the three indicators for data quality. Interviewers who behave correctly or incorrectly do not show differences in data quality.

5. Conclusions and recommendations for further research

The first and foremost conclusion derived from the TQM-approach, is that the ISPO-network should screen its interviewers more thoroughly. After all, a not insignificant number of interviewers (18%) fail to conform to some basic, straightforward formal arrangements. During general and specific training sessions on interviewing, it seems advisable to lay more emphasis on the importance of the interviewer's correct relationship with the network.

The criteria used so far cannot improve the selection process. The socio-demographic characteristics of the interviewers involved did not seem to relate significantly to the organisational aspect of the way the interviewers carried out their duties. Despite the fact that the pre-set selection criteria could not always be adhered to (because the demand for interviewers exceeded the supply of interviewers on the job market), it did not seem to have decisive repercussions for the way formal agreements were adhered to.

This conclusion also applies for the association between socio-demographic selection criteria and the quality of the gathered data. Only the work situation, after all, showed a modest, but significant association with the achieved quality of the data. A more important observation is that the interviewers who had already co-operated in the post-electoral research of 1991 seemed to perform worse than new interviewers. This finding seems to indicate that new interviewers are preferable for panel surveys. It also shows, together with the fact that new interviewers seem to adhere better to formal arrangements, that interviewers who have been involved with projects for a longer period should be repeatedly reminded of how interviewers are expected to do their job. If they are given

feedback regarding the way they carried out a project, combined with additional training sessions to refresh the basic principles underlying a standard interview, the relationship with the network can be highlighted, and the quality standards can again be clarified. This conclusion is compatible with the TQM concept. When certain processes no longer satisfy quality standards, improvements need to be made.

Lastly, the findings do not seem very useful for future interviewer selection. The association between the performance and known socio-demographic characteristics of the interviewers did not produce many significant results. The correlation between relationship to the network and performance did not prove to be significant either. For the TQM-approach, however, it means that the selection procedure and the criteria used so far need some fine tuning. As far as the criteria are concerned, personality traits and social skills need to be considered. It is clear from the interviewer's job description that psychological factors and social skills may play an important role. It immediately implies that the selection and recruitment procedures may require some additional attention. One possibility is to invite all the interviewers for an assessment. This could include questions regarding motivation, perseverance and discipline, but it could also contain role playing to assess how the potential interviewer would behave in real interview conditions. The question remains whether enough interviewers are available, given the state of the labour market and the interviewer's status.

Therefore, total quality management is a useful approach within survey research, particularly for the specific application of an important component of the survey, namely the gathering of data through personal interviews. TQM clearly offers possibilities for drawing up an assessment procedure for interviewers employed by a permanent interviewer network.

References

- ALWIN D.F. & KROSNICK J.A. (1991), "The reliability of survey attitude measurement. The influence of question and respondent attributes". In: *Sociological Methods & Research*, 20(1): 139-181.
- BEERTEN R., BILLIET J., CARTON A. & SWYNGEDOUW M. (1997), *1995 General Election Study, Flanders - Belgium: codebook and questionnaire*. Leuven, ISPO - K.U.Leuven.
- BIEMER P.P. & CASPAR R. (1994), "Continuous quality improvement for survey operations: Some general principles and applications". In: *Journal of Official Statistics*, 10(3): 307-326.
- BILLIET J. & LOOSVELDT G. (1988), "Improvement of the quality of responses to factual survey questions by interviewer training". In: *Public Opinion Quarterly*, 52(2):190-211.
- CARTON A. (1995), "Towards an effective evaluation of interviewers working within the context of a permanent network of interviewers". In: *Proceedings of the International Conference on Survey Measurement and Process Quality*, Alexandria: American Statistical Association, p. 145-150.
- COLLEDGE M. & MARCH M. (1993), "Quality management: Development of a framework for a statistical agency". In: *Journal of Business & Economic Statistics*, 11(2):157-165.
- COLLEDGE M. & MARCH M. (1997), "Quality policies, standards, guidelines, and recommended practices at national statistical agencies". In: LYBERG L.E. et al. (eds.), *Survey measurement and process quality*. New York: John Wiley, p. 501-522.
- DE LEEUW E.D. & COLLINS M. (1997), "Data collection methods and survey quality". In: LYBERG L.E. et al. (eds.), *Survey measurement and process quality*. New York: John Wiley, p. 199-220.
- DIPPO C.S. (1997), "Survey measurement and process improvement: Concepts and integration". In: LYBERG L.E. et al. (eds.), *Survey measurement and process quality*. New York: John Wiley, p. 457-474.
- FODDY W. (1993), *Constructing questions for interviews and questionnaires. Theory and practice in social research*. Cambridge: University Press.
- FOWLER F.J. Jr. & MANGIONE T.W. (1990), *Standardized survey interviewing: Minimizing interviewer-related error*. Newbury Park, Sage Publications.
- GROVES R.M. (1989), *Survey errors and survey costs*. New York: John Wiley.
- HOX J.J., DE LEEUW E.D. & KREFT I.G. (1991), "The effect of interviewer and respondent characteristics on the quality of survey data". In: BIEMER P.P. et al. (eds.), *Measurement errors in surveys*. New York: John Wiley, p. 439-461.
- HOX J.J. (1994), "Hierarchical regression models for interviewer and respondent effects". In: *Sociological Methods & Research*, 22(3): 300-318.
- LOOSVELDT G., CARTON A. & PICKERY J. (1998), "The effect of interviewer and respondent characteristics on refusals in a panel survey". In: *ZUMA Nachrichten Special*. Mannheim (forthcoming).
- MORGANSTEIN D. & MARKER D.A. (1997), "Continuous quality improvement in statistical agencies". LYBERG Lars E. et al. (eds.), *Survey measurement and process quality*. New York: John Wiley, p.475-500.
- MORTON-WILLIAMS J. (1993), *Interviewer approaches*. Aldershot, Dartmouth Publishing.
- PICKERY J. & LOOSVELDT G. (1998), "The impact of respondent and interviewer characteristics on the number of 'no opinion' answers". In: *Quality and Quantity*, 32: 31-45.
- RASBASH J. & WOODHOUSE G. (1996) *Mln command reference. Version 1.0a*. London: Multilevel Models Project. Institute of Education, University of London.
- SUDMAN S., BRADBURN N.M. & SCHWARZ N. (1996), *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass Publishers.
- YANG M., GOLDSTEIN H. & RASBASH J. (1996) *Mln macros for advanced multilevel modelling. Version 1.0*. London: Multilevel Models Project. Institute of Education, University of London.