

Meyermann, Alexia: The use of Behavior Coding to Analyze Data Quality in the SOEP Establishment Survey 2012/2013

draft version
August, 16th 2013

1. INTRODUCTION.....	1
2. BEHAVIOR CODING AND ITS APPLICATION ON ESTABLISHMENT SURVEYS	1
2.1. BEHAVIOR CODING	1
2.2. THEORETICAL CONSIDERATIONS	2
2.3. THE CODING SCHEME.....	3
3. ANALYSES AND OUTLOOK OF THIS RESEARCH	5
4. LITERATURE	6

1. INTRODUCTION

The understanding of the survey response process is crucial to achieving high data quality. This holds with regard to either the prevention of errors based on the implementation of special design features or the reduction of errors based on the application of special statistical methods. When it comes to surveys of organizations too little is known about the actual response process and its mechanisms yet, that could be used to ensure data quality. This research addresses the topic of response quality in establishment surveys by analyzing 31 audio-taped f2f-survey interviews of the SOEP-LEE establishment survey 2012/2013. The audio-recordings are analyzed by applying the method of Behavior Coding (Ongena/Dijkstra 2006). According to this every observable respondent behavior shown during an interview and with regard to a single question-answer-sequence will be coded based on a pre-defined coding scheme. This paper presents the design of the Behavior Coding application and possible ways of analysis. This first version of the paper will be supplemented by empirical results later on.

2. BEHAVIOR CODING AND ITS APPLICATION ON ESTABLISHMENT SURVEYS

2.1. Behavior Coding

Behavior Coding (BC) is a method of observing the question-answer process (Ongena/Dijkstra 2006). By Behavior Coding interviewer and respondent behaviors shown during an interview will be coded. Thus, BC is used to standardize non-standardized, qualitative information by coding (Schnell 2012, 150). All kinds of behaviors can be coded such as wording or paralinguistic utterances, pauses, even body language, and facial expressions. The resulting information (i.e. the coded behaviors) can then be analyzed statistically and by so called “diagnostic” methods, using the full richness of the qualitative material (van der Zouwen/Smit 2004).

BC is used since the end of the 1960s (Cannell et al. 1968), and was originally applied to observe interviewer behavior as a form of interviewer monitoring. Until today, it has often been used during questionnaire pretesting and evaluation (Ongena 2005, 62). Applications of BC include researches on interviewer behaviors, respondent behaviors, the interaction between interviewer and respondent,

the influence of certain types of questions on interviewer or respondent behavior, and consequently on determinants and consequences of the response process overall (Ongena/Dijkstra 2006, 434, Ongena 2005, 59pp). It seems that an application of the behavior coding method on establishment surveys and establishments' response behaviors has not been done so far.

The use of BC to infer response quality is based on the assumptions of the standardized interviewing paradigm (e.g. Beatty 1995). According to this every departure from the standardized interviewing rules – such as not reading the question exactly as scripted – can be considered as threats to data quality. Behavior coding will then be used to identify these departures from the standardized or “paradigmatic” interview. But, the interpretation of departures as being an indication of low response quality is limited (e.g. Dykema et al. 1997, 288, Ongena 2005, 89). *Firstly*, deviating behaviors such as respondent's requests for clarification can be resolved by interviewers explaining unclear terminology correctly. Although, if the interviewer does not probe neutrally but suggestively instead, responses are influenced in a certain direction limiting response quality. *Secondly*, deviances can have several causes, which cannot be distinguished by the method itself. For example respondents' requests for clarification could be either caused by comprehension problems or by high respondents' motivation or diligence. *Thirdly*, with BC only obvious (visible or audible) behaviors are observed. Thoughts, other kinds of problematic behaviors, which are not expressed, or misinterpretations of which the respondent is not aware of will not be coded anyway. But, obvious behaviors still signify response quality following Dykema and her colleagues: They refer to the phenomenon of the “classroom question” which shows that “... questions by only one or two students indicate that the class at large or in this case respondents in general, do not understand.”(Dykema et al. 1997, 305)¹

Despite of these critics BC might be a valuable tool for a thorough analysis of the question-answer-processes. The basis of BC is the observation of the complete and real-time response process, it is not as retrospective as respondent debriefings and it is not as intrusive as think-aloud methods for example. Through the standardization of qualitative material BC offers a quantitative look, the application of statistical analysis methods, and herewith the opportunity of making inferences about the distribution of phenomena. Quantitative findings can then be supplemented by a closer look into the qualitative material. The richness of the qualitative material can be used in a more focused way and with a direct link to the quantitative findings (van der Zouwen/Smit 2004).

Behavior Coding is based on a coding scheme, which is a list of codes used to observe and count all types of behaviors a researcher is interested in. Coding schemes used in the past vary much depending on the level of detailedness, the research purpose and theoretical considerations. The coding scheme applied here is based on certain theoretical considerations of the establishment survey response process.

2.2. Theoretical considerations

The establishment or organizational survey response process is different to the one in individual surveys in (at least) two central aspects. Firstly, establishment surveys rely on the use of proxy

¹ There are studies which analyze the qualification of deviating behaviors as indicators of low response quality (Belli/Lepkowski/Kabeto 2001, Dijkstra/Ongena 2006; Dykema et al. 1997, Smit/Dijkstra/Van der Zouwen 1997, Van der Zouwen/Smit 2004). These studies show for example, that suggestive probes reduce response quality (Smit et al. 1997, Ongena 2005), or that certain types of problematic deviances correlate negatively with response quality (Dijkstra/Ongena 2006). In the contrary other authors find mixed results (Dykema et al. 1997, Mathiowetz 1999).

informants who respond acting as an organizations' representative and secondly, they rely on the so called organizational information system (OIS) or record-keeping system.² It follows that factors on the individual as well as on the organizational level are relevant for response processes and their outcomes.

Looking at individual level factors survey response theory names the two aspects *motivation* and *ability* as determining responses (e.g. Krosnick 1991, Groves et al. 1992, 486, Beatty/Herrmann 2002, 72). An individual respondent needs to be willing and able to respond accurate. In organization surveys the work situation should be considered in addition: Extending Tomaskovic-Deveys and his colleagues' (1995) work, the aspect *ability* can be differentiated in the individual's capacity and authority to respond. Capacity comprises basically the informant's knowledge of the organization itself and of the organizational information system (Bavdaz 2010). Authority can further be distinguished into the individual's authority to access and to publish information.

Organizational level factors are relevant for response quality as the OIS determines what kind of information is gathered (*existence*), how the information is recorded (*accuracy*) and where as well as the terms of accessing this information by its employees (distribution and *accessibility*). Information that is requested by a survey institute might not be collected for the time period in question or for the organizational unit under review. It might as well be that there are lacks of accuracy in the documentation itself (Lorenc 2006). The distribution of knowledge throughout organizations linked to the division of labor leads to the fact that not every employee or organizational unit is equally knowledgable. Sometimes information is gathered even outside the organization itself for example due to the outsourcing of personnel or tax reporting tasks.

Summing up, it is expected that the OIS (information's *existence*, *accuracy*, and *accessibility*) as well as informant's *motivation*, *capacity*, and *authority* determine response quality in organizational surveys.³ It follows that as far as organizations differ in their information systems and as far as individuals selected to being proxy informant differ there is a risk of response bias. This research analyzes to what extent quality problems arise due to differences between establishments and between informants. Having said this, a methodological design is needed that allows for the observation of these aspects and processes. How far the behavior coding method – as a tool of observing the response process – allows observing these aspects is an open research question yet.

2.3. The coding scheme

The coding scheme that was developed within this research consists of 35 different codes and it includes codes regarding interviewer as well as respondent behavior. It allows coding on the level of each question-answer-sequence. The codes generally allow to distinguishing between actors, occurrences of behaviors, types of deviating behaviors, and the intensity of certain behaviors (minor

² Organizational response process models and theories have been developed so far and amongst others by Willimack/Nichols (2001, 2010), Bavdaz (2010), Lorenc (2006), and Tomaskovic-Devey et al. (1995). Willimack/Nichols extended the cognitive response process model from Tourangeau (1984) with steps on the organizational level. These steps refer to the relevance of the information system ("retrieval from record", "record formation") and the informant ("selection and identification of respondent/s"). Bavdaz (2010) proposes a broader perspective on the response process, distinguishing between organizational and individual level factors and specifically introducing the aspect of recurring surveys. Lorenc (2006) focuses on the paths of information processing and looking at the organization as a single cognitive unit (based on Socially Distributed Cognition Theory). Tomaskovic-Devey et al. (1995) introduce a theoretical perspective based on organization studies focusing on the organization as an actor.

³ This paper focuses solely on the respondent's (organization's and informant's) influence on response quality. Of course, response quality also depends on the interview situation, the survey question and the interviewer.

or major departures). Table 1 lists a selection of the codes for respondent behaviors and gives additional information on relations with underlying theoretical constructs.

Table 1: coding of respondent behavior (selection of the full coding scheme)

	Indicating ... on the individual level	Indicating ... of the OIS	Effect on response quality of a single response	Indicator of response quality on the aggregate level
Question delivery and answering stage				
Interruption of question reading	-	-	negative	low response quality
Comprehension issue	capacity -	-	negative	low
Sensitivity issue	motivation -	-	negative	low
Authority issue	authority -	-	negative	Low
Knowledge issue	capacity -	-	negative	Low
Retrieval from records	capacity - motivation +	existence of information + accessibility of information +	positive	low (burden +)
Consultation of colleagues	capacity - motivation +	existence of information + accessibility of information +	positive	low (burden +)
Complaints	motivation -	-	negative	low
Response and related behaviors:				
Response:	-	-		
1. Direct and adequate response (=paradigmatic sequence)			positive	high
2. Qualified response (e.g. "approx.", "roughly")			negative	low
3. other behavior (=problematic sequence)			negative	low
Inadequate response, relevant but not within response options	-	-	negative	low
Invalid response, showing misinterpretation	capacity -	-	negative	low
Additional comments and reports	capacity + motivation +	-	positive	low
Irrelevant comments	motivation -	-	negative	low
Extraordinary effort	motivation +	-	positive	-
Final codes				
Problem could not be solved (=Inadequate sequence)	-	-	negative	low
Certainty of response evaluated by coders	capacity +	-	-	-

These codes, i.e. the observed behaviors, can be interpreted as indicating the individual level factors motivation, capacity and authority which were outlined above as determining response quality on the individual level. The use of the OIS, in particular the existence and accessibility of information can be observed with these codes as well. Certain behaviors can be interpreted as having a negative or positive effect on response quality of that specific response at which they were observed. In addition, a high prevalence rate of certain behaviors regarding one survey question can be interpreted as indicating a higher chance of low response quality over all responses to that question. With these codes paradigmatic sequences can be distinguished from problematic and inadequate sequences, where a problem that arose (such as a request for clarification) could not be repaired by the interviewer. Some behaviors (such as interruptions) indicate response problems rather due to poor question formulation instead of respondent factors, which needs to be considered in the interpretation.

3. ANALYSES AND OUTLOOK OF THIS RESEARCH

The data for the analyses stem from a representative establishment survey of German employers (N=1708) that was conducted in 2012 and 2013 using f2f- and paper-and-pencil-interviews.⁴ Establishments were sampled based on address information given by employed participants in the Socio-Economic Panel Study (SOEP). Information from both surveys can be linked in order to create a linked employer-employee data set on organizational strategies and labor market outcomes. Within the SOEP-LEE establishment survey 31 interviews were audiotaped and form the sample where behavior coding is applied.⁵ There is a total of 145 coding instances or behavior observations per interview.

Codes will be used firstly to analyze prevalence rates of behaviors, and secondly to build a typology of respondents, if the data allows to. The first question to be answered will be how often do behaviors that are expected in establishment surveys arise overall and per question. The literature gives information on criteria for problem questions (van der Zouwen/Smit 2004). Next, it will be analyzed, if types of establishments and types of informants differ in their response behaviors as expected. Unfortunately, as the number of observations (31 interviews) is rather low the possibilities of these statistical analyses are limited. Last but not least, certain behaviors and interviews will be analyzed further by looking at the qualitative material more deeply (van der Zouwen/Smit 2004).

Future research could realize additional value of BC if the findings of these analyses were combined with other types of data as Ongena (2005, 52) suggests. This can be done in two directions. First, hypotheses can be built upon the findings of BC which can then be analyzed using other response quality indicators such as Item Nonresponse or quality indicators which are based on editing information or interviewer debriefings. The question behind is, whether additional analyses show similar results. The SOEP-LEE data includes several types of (para-)data that can be used in that sense: interviewer survey data, interviewer debriefings regarding every interview situation⁶, data regarding the editing process⁷, and data based on the handwritten responses and comments in the original paper questionnaires⁸. Second, findings based on quantitative analyses of the whole establishment sample using the different sources of data and paradata can be supplemented by the BC method. In this case, BC and especially the qualitative material can be helpful in an explanatory sense as it allows a closer look into the actual response process.

⁴ For more information, please see http://www.diw.de/sixcms/detail.php?id=diw_01.c.394066.en. The technical report and the dataset will be available in 2014.

⁵ Originally, 100 audio-recordings were planned and 109 of 502 field interviewers had been ordered to record interviews. Coding was primarily based on written transcripts, as each interview had been fully transcribed. For the application of certain codes the audio-recordings were consulted in addition. In order to ensure the reliability of the coding, random parts of multiple interviews will be double-coded.

⁶ These evaluations include questions regarding the respondent (difficulties of identifying, his/her accuracy and knowledge), regarding the interview situation (interruptions), and regarding the response process (information retrieval, use of records, response burden).

⁷ Within the editing process, several procedures of data cleaning were done such as checks of plausibility, consistency, sums and filters. While the checks themselves were done automatically, breaches were settled based on personal inspection done by survey institute staff. Settlement was done by looking at the original questionnaires and by re-contacting the establishments. This kind of paradata includes quantitative information on the number of checks done, the number of items checked, and the number of breaches found, the number of re-contacts, whether a clarification was based on a re-contact with a person other than the original respondent, or whether an attempt to re-contact was unsuccessful. It also includes qualitative information stemming from the editing group's comments regarding the telephone contacts.

⁸ During the process of digitalizing paper questionnaires all of the responses that were not encodable within the coding scheme for the corresponding question were documented. Examples are handwritten comments next to a survey question or inadequate responses, such as the wrong use of scales.

4. LITERATURE

- Bavdaz, Mojca. 2010. Sources of Measurement Errors in Business Surveys. *Journal of Official Statistics* 1: 25–44.
- Beatty, Paul. 1995. Understanding the Standardized/Non-Standardized Interviewing controversy. *Journal of Official Statistics* 11 (2): 147–160.
- Belli, Robert F.; Lepkowski, James M.; Kabeto, Mohammed U. 2001. The respective roles of cognitive processing difficulty and conversational rapport on the accuracy of retrospective reports of doctor's office visits. In *Seventh Conference on Health Survey Research Methods*, eds. Richard A. Kulka and Marcie L. Cynamon. Hyattsville, MD: U.S. Government Printing Office (DHHS Publication No. (PHS) 01-1013): 197–203.
- Cannell, C.F., F.J. Fowler, and K.H. Marquis. 1968. The influence of interviewer and respondent psychological and behavioral variables on the reporting of household interviews. *Vital and Health Statistics*, 2 (26).
- Dijkstra, Wil; Ongena, Yfke. 2006. Question-Answer Sequences in Survey-Interviews. *Qual Quant* 40 (6): 983–1011. DOI: 10.1007/s11135-005-5076-4.
- Dykema, Jennifer; Lepkowski, James M.; Blixt, Steven. 1997. The Effect of Interviewer and Respondent Behavior on Data Quality: Analysis of Interaction Coding in a Validation Study. In *Survey measurement and process quality*, ed. Lars Lyberg. New York [etc.]. Wiley: 287–310.
- Groves, Robert M., Robert B. Cialdini and Mick P. Couper. 1992. Understanding the decision to participate in a survey. *Public Opinion Quarterly* 56: 475–495.
- Loosveldt, Geert. 1997. Interaction Characteristics of the Difficult-to-Interview Respondent. *International Journal of Public Opinion Research* 9 (4): 386–394. DOI: 10.1093/ijpor/9.4.386.
- Lorenc, Boris. 2006. Propagation of Representational States in a Survey of Schools. In *Proceedings of Q2006*, eds. European Conference on Quality in Survey Statistics.
- Mathiowetz, Nancy A. 1999. Research note. Respondent uncertainty as indicator of response quality. *International Journal of Public Opinion Research* 11 (3): 289–296. DOI: 10.1093/ijpor/11.3.289.
- Ongena, Yfke Pieternel. 2005. Interviewer and respondent interaction in survey interviews. [S.l.: s.n.].
- Ongena, Yfke Pieternel; Dijkstra, Wil. 2006. Methods of Behavior Coding of Survey Interviews. *Journal of Official Statistics* 22 (3): 419–451.
- Schnell, Rainer. 2012. Survey-Interviews. Methoden standardisierter Befragungen. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Smit, Johannes H.; Dijkstra, Wil; van der Zouwen, Johannes. 1997. Suggestive Interviewer Behaviour in Surveys: An Experimental Study. *Journal of Official Statistics* 13 (1): 19–28.
- Tomaskovic-Devey, Donald, Jeffrey Leiter, and Shealy Thompson. 1995. Item Nonresponse in Organizational Surveys. *Sociological Methodology* 25: 77-110.
- Tourangeau, Roger. 1984. Cognitive Sciences and Survey Methods. In *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*, eds. T.B. Jabine, M.L. Straf, J.M. Tanur, R. Tourangeau, 73-100. Washington, DC: National Academy Press.
- van der Zouwen, Johannes; Smit, Johannes H. 2004. Evaluating Survey Questions by Analyzing Patterns of Behavior Codes and Question-Answer Sequences: A Diagnostic Approach. In *Methods for testing and evaluating survey questionnaires*, eds. Stanley Presser, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin und Eleanor Singer. Hoboken, NJ: Wiley-Interscience (Wiley series in survey methodology).
- Willimack, Diane K., Elizabeth Nichols. 2001. Building an Alternative Response Process Model for Business Surveys. In *Proceedings of the Annual Meeting of the American Statistical Association*, eds. American Statistical Association.
- Willimack, Diane K., Elizabeth Nichols. 2010. A Hybrid Response Process Model for Business Surveys 2010, *Journal of Official Statistics* 26 (1): 3–24.