

Not Willing, Not Able: Causes of Measurement Error in Business Surveys

Deirdre Giesen¹, Tony Hak²

¹Statistics Netherlands, P.O. Box 4481, 6401 CZ Heerlen, the Netherlands

²Rotterdam School of Management, Erasmus University, P.O. Box 1738
3000 DR Rotterdam, the Netherlands

Abstract

National statistical institutes must collect accurate data from businesses in a timely and cost-effective way and without causing too much response burden. An adequate design of the information request is critical in achieving this goal. This paper describes the lessons we have learned about the design of business survey questionnaires from a thorough evaluation of the questionnaires of a typical business survey for official statistics, the Structural Business Survey. The paper presents a framework for understanding factors that contribute to missing and inaccurate data and draws a number of conclusions regarding how the design of business surveys can be improved to take these factors into account.

Key Words: Data collection, establishment surveys, response process

1. Introduction

The design of a business survey poses challenges additional to the ones that occur in the design of a household survey (e.g. Tomakovic-Devey, Leiter and Thompson, 1994; Cox and Chinnappa, 1995). In both types of survey, a person performs a (partly cognitive and partly administrative) task and is more or less able and willing to do so. However, in business surveys the performance of this task occurs within an organizational setting and may be affected by characteristics of that setting. First, responding to a questionnaire competes with resources spent on the main goals of the organization and usually is an additional task that interrupts the business respondent's normal working day (Phipps, Butani and Chun, 1995). Second, it can be difficult to get access to the right respondent (e.g. Gower, 1994), i.e., the person that knows where the relevant information can be found and has access to it. Third, the requested information is usually not located in the head of a respondent (memory), but may need to be retrieved from several persons and records (e.g. Anderson, Morrison and Brady, 2005). Fourth, decisions about survey participation and about how much maximum effort should go into it may be made at multiple levels in the organization (i.e., it may depend on priorities of both the management and the persons actually involved in completing the questionnaire).

Business surveys for official statistics have additional characteristics that should be taken into account. Usually satisfactory levels of unit-response can be accomplished because most official business surveys are mandatory. However, statistical agencies must spend many resources in getting a timely response (e.g. sending out reminders) and adequate data quality (e.g. data editing). Another important characteristic of official business surveys is that they continue for many years. This implies that some businesses might be sampled more than once and others (because of their impact on population totals) often or every year or month. Whereas households may never or hardly ever be confronted with an official statistics survey request, for certain

businesses this is a regular event. As a result survey organizations collecting data for official statistics often have long term relationships with the larger businesses.

The Multi-dimensional Integral Business Survey Response Model (Bavdaž, 2010a, see Figure 1), which builds on other models of the response process in business surveys (Tourangeau, Rips and Rasinski, 2000; Willimack and Nichols, 2010) takes both the organizational and the temporal setting of business surveys in official statistics into account. The model distinguishes between the cognitive response process on the individual level (comprehension, retrieval, judgment and response) and the processes on the business level that provide the context for the response process (organization, business information systems and authorization). Moreover, the model represents the fact that different actors can be involved in responding: gate-keepers, authorities, coordinators, data providers and respondents. The arrows indicate that responding to the survey may be recurring and that experiences from past surveys can affect the response to future ones.

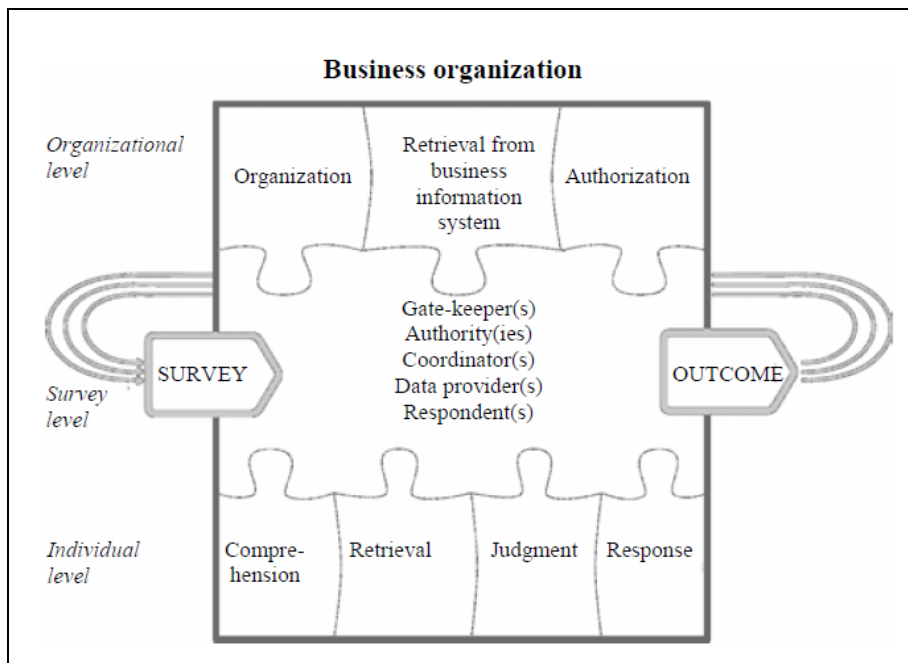


Figure 1: Multi-dimensional Integral Business Survey Response Model
Source: Bavdaž (2010a)

The vast literature on questionnaire design mainly reports findings regarding household surveys in which respondents, in their free time, report about self. The aim of this paper is to contribute to the much smaller literature that specifically deals with business surveys. We describe the lessons we have learned about the development and improvement of business surveys from an extensive evaluation of a typical official business survey questionnaire, the Dutch Structural Business survey (hereafter: SBS). The main goal of this evaluation was to discover and understand sources of measurement error and response burden to guide the redesign of the questionnaires. In section 2 of this paper we briefly describe the SBS survey, the evaluation program, and the main evaluation findings. Building on these findings, we present in section 3 a model of the main factors that affect the response outcome of a business survey. We conclude in section 4 with a discussion of how specific survey design characteristics can be improved.

2. The evaluation of the SBS questionnaires

The Structural Business Survey (SBS) measures annually a large number of indicators of the activity and performance of Dutch businesses. The data are used for the European Structural Business Statistics and for the Dutch system of National Accounts. Almost all industries and all size classes are covered by this survey. Variables collected include detailed information on sales and other revenue, expenses, inventories, purchases and employees. Different sample and follow-up strategies are used for businesses according to their size and relative weight in the published statistics. The larger firms (with 50 or more employees) receive an SBS questionnaire every year. Follow-up strategies for non-respondents are more intensive for statistically “crucial” firms than for other firms. Response is mandatory by law for all firms. More than 84,000 questionnaires were sent out in the year in which we started our evaluation, and the response rate was about 70%.

The questionnaires used were paper forms, mainly sent out for self completion through the mail. Field officers collect the data on site for complicated and very important businesses. The questionnaires are long; more than 15 pages are typical. Questions are printed on the right hand page, and instructions on the opposite left hand page. The number of items ranges between 38 for small health care organizations and 504 for large manufacturers, with an average of 128 items. An integrated set of questionnaires is used. Each questionnaire consists of a uniform core part that is the same for all businesses. This core part specifies all revenues and costs which then must be added up into a summary statement that shows (among other things) the business’ return. Depending on size and branch, the second part of the questionnaire asks for a further specification of certain items.

We summarize here the aims, methods and findings of a thorough evaluation of this survey (see Giesen and Hak, 2005, for a more detailed description). The aim of this evaluation was to find sources of response burden and data error in order to improve the questionnaires. In line with the European Statistics Code of Practice (Eurostat, 2011), reducing response burden is a goal in itself for Statistics Netherlands. Reducing burden and improving data quality were also seen as ways to reduce the costs of data collection and data editing.

For the evaluation we took a two step approach. First, we analyzed information that was already available in the agency to specify and localize problems. We then used these insights to collect information about the response process with respondents in the field, in order to understand causes of reporting errors and response burden. The *agency resources* used for this evaluation were:

- Detailed inspection of completed forms (n=66, including both early and late respondents and responses from businesses of various size classes). We identified problems by looking at features such as crossed out words and numbers, write-in comments from respondents, and calculation errors.
- Coding of respondents’ questions and comments. All questions and comments regarding the SBS questionnaires from respondents by mail, phone, or on the comments sheet that accompanies the questionnaires, are routinely stored in an information system. A field officer coded each comment received in one year (2223 comments from 2149 different respondents).
- Data quality indicators. We explored the use of three indicators to detect problematic questions and groups of respondents: unit non-response, item non-response and plausibility (a plausibility index is routinely calculated for each completed questionnaire).
- Focus group interviews with staff. Eight focus groups were arranged with field officers, helpdesk staff, data editors and data analysts.

In the second step of our evaluation we contacted *respondents* in order to validate the findings from the first step and to further explore causes of data error and response burden. We assumed that this could be done best by studying the activities required from respondents in the circumstances in which they occur (see Giesen, 2007, for a more detailed description of how establishment questionnaires are tested in the field by Statistics Netherlands). We conducted real-time on-site observations of how respondents dealt with the SBS questionnaire (eleven respondents). We also conducted on-site retrospective focused interviews on how respondents had completed their forms (twelve respondents). We also conducted fourteen telephone interviews with non-respondents.

The following complaint by a respondent of the SBS survey illustrates *how* respondents experience the data requests from Statistics Netherlands in general and the request to respond to the SBS in particular: *“Frequently we are faced with piles of Statistic Netherlands' surveys that seem to have no use at all. The all-time low must be the SBS questionnaire with idiotic questions about the quantities of steel and copper wire that are consumed. We really don't keep a record of this. Gathering this data takes days.”* It is this kind of sentiment (an effort that takes days for no use at all) that goes a long way in explaining non-response and data error.

A main finding of our research was that response burden is perceived as too high by many respondents. 30% of the coded respondent remarks were explicit complaints about the response burden. The second largest group of comments (27%) related to the mismatch between the questionnaire items and the respondent's records, which can also be seen as an indicator of response burden. The respondents in our field study took between 45 minutes and 2 and half days to complete the questionnaire (a bit more than 4 hours on average). How difficult and time consuming the response task was varied strongly. We found two main reasons for this variation. First, the actual effort that is required for correct completion varies. This can be caused by differences in the questionnaire design, as smaller firms in the SBS get less detailed questionnaires than larger firms. But it is important to note that also identical questionnaires can require more or less effort depending on characteristics of the respondent and the business. Responding was less burdensome for competent respondents who knew key figures of their organization by heart, kept well organized records, knew how to access them and who worked very precisely while completing the questionnaire. However, for respondents with little knowledge about accounting practices, their own figures and completing these kinds of financial overviews the response task could be extremely difficult. Second, response burden also varied because of differences in how serious respondents took their task, or in other words: how much effort they were willing to spend. Some respondents made careful automated links between their own data systems and the questionnaire items; others took a more time-efficient approach.

In our evaluation study we identified the following main causes of actual response burden (i.e. the time spent on collecting data and completing the questionnaire):

1. Large number of questionnaire items
 2. High level of detail asked
 3. Confusing structure and lay-out of the questionnaire
 4. Lack of match between business records and questionnaire items
 5. Lack of knowledge and skills of respondent
 6. High motivation to complete questionnaire accurately
- Highly motivated respondents might also experience a high response burden because they tend to actually spend more effort and time.

Each of these burdensome aspects might also contribute to an increase of perceived burden, which also might be caused by a perceived lack of usefulness of the survey. Such perceived lack of usefulness makes the response task irritating and burdensome, even if little time is spent on this task.

Regarding *data quality* our analyses of *item non-response* in raw and edited data showed high percentages of empty fields. Businesses without employees showed the highest level of empty fields (about 60%) and businesses with more than 200 employees the lowest levels (about 20%). Only about 10% of the empty fields could be replaced by numbers or amounts in the process of data editing. Although an empty field cannot be straightforwardly interpreted as either an instance of item non-response, 'not applicable' or (correctly rounded off to) zero, we developed the hypothesis that an empty field most often indicates that the item is not applicable, implying that many respondents are faced with a large number of questions that are not relevant for their businesses. This hypothesis was only partly confirmed in our fieldwork. Indeed, for the small companies we visited, many items on the questionnaire were not applicable or could be rounded off to zero. However, for the medium and larger firms many specifications were left blank because the respondents judged that it was not worth their effort to retrieve the relevant number or amount. A typical example was the specification of heating costs by electricity and gas. This information is available to most respondents if they would take the effort to retrieve the invoice of the power company, but many do not. We were able to assess the quality of the information for each item on the questionnaire for ten questionnaires by discussing the completed forms with respondents in a retrospective interview. Information for 23% of the items of these ten questionnaires appeared to be incorrect, i.e., either wrongly left empty or filled out with a wrong number or amount. It must be noted here that one interpretation error can result in an incorrect answer for multiple items.

Lack of motivation and/or time had a negative impact on data quality at almost every step of the response process: the selection of a not very well informed respondent, careless reading resulting in interpretation errors, lack of effort to retrieve the data in the requested format, lack of effort put in consolidating the necessary information, and lack of effort in checking for and correcting errors. Data error most frequently occurred when there was a mismatch between the information that was requested and the data that were readily available for the respondent. Several respondents told us that they made very rough estimates, or even consciously made reporting errors, because calculating the correct number was perceived by them as requiring too much work. Often respondents legitimized their response behavior with the assumption that a correct figure for a specific item is of no importance for the final statistical product (similar to finding as reported by Bavdaž, 2010b). On the other hand, many respondents we talked to saw it as their professional duty to provide high quality data ("If I do this, I do it correctly"). Interestingly, some respondents indicated that they attached more value to providing accurate and timely information than their managers.

3. The crucial role of perceived costs and benefits of accurate reporting

Our findings support the following model of the causes of non-response and data error, consisting of three parts:

- (a) Survey properties as well as respondent characteristics affect the perceived costs and benefits of the response task.
- (b) These perceived costs and benefits affect the response behavior.
- (c) Through the response behavior perceived costs and benefits affect response quality.

This model is the core of various frameworks explaining quality in business statistics (Willimack, Nichols and Sudman, 2002; Haraldsen, 2004; Jones, Rushbrooke, Haraldsen, Dale and Hedlin, 2005; Snijkers, 2007). Our findings also support Bavdaž' (2010a) model on the relationship between answer availability and likely quality of the response outcome, with for example accessible data most likely to result in exact answers, whereas inconceivable or non-existent data most likely results in blunders or item nonresponse.

Our findings also add to the existing models and insights, by showing how in a business surveys:

- A mismatch between required effort and provided effort can result in item-nonresponse and data error.
- This mismatch can be caused by decisions at multiple levels (both on the organizational and personal level) and at multiple points in time.
- These decisions are at least partially dependent on characteristics of the information request.

Our findings suggest that the primary focus in the design of data collection should thus be to make the costs of accurate reporting as low as possible, while at the same time maximizing its perceived benefits. Put differently, less is more: if less effort is required, the likelihood that adequate effort is invested will increase.

Four “moments” of the response process model can be identified as crucial for this task of reducing the perceived costs of responding to the survey:

1. *Assessment of priorities.* Both managers and respondents should be convinced of the usefulness of the survey and of the necessity of information of their company for a correct estimate by the statistical institute. Perceived and actual burden should also be reduced by a reduction of the length of the questionnaire.
2. *Comprehension of the data request.* Questionnaires must be designed such that the structure of the questionnaire is not confusing; that only easily accessible information is asked; and that clearly is indicated what level of precision of the answer is required. Respondents should be assisted in identifying a possible lack of knowledge or skills, so support can be sought.
3. *Retrieval and Judgment.* It should be avoided that a (too) high level of detail is asked. Lack of match between business records and questionnaire items should be avoided as well. Respondents should be informed how important it is that data is reported accurately.
4. *Response.* Respondents should be supported in correctly calculating a total from different items. Questionnaires should be designed such that no additional effort is required for identifying the box or place for an answer.

4. How survey design can affect right match of required and given effort

Concrete recommendations for the SBS questionnaire were the following.

Assessment of priorities: Communication about the survey should explain the usefulness of the requested effort. The communication about the survey should be correct and follow professional standards, for example all letters and requests by respondents must be answered and processed correctly; respondents should also receive a confirmation of receipt of the returned questionnaires. This kind of communication should prevent unnecessary irritation about the agency and signal to the respondents that their contribution is important. Businesses, especially those that are in the sample for several years, should receive feedback on the quality of the data they provide in order to educate them and to motivate them for accurate reporting. Ideally, both the management and the respondent should be informed about this. Enforcement policies should be expanded to the quality of the data provided and respondents must be informed about this.

Comprehension of the data request: Exclude small businesses from the survey or develop tailored questionnaires for this group in very simple language. Include in each survey request a clear description of the targeted business units and time period. A more logical ordering of the questions so that each type of costs or income is only specified once in the questionnaire. Formulate instructions that are considered of crucial importance as questions. The instruction text should be reduced to the absolute minimum and placed directly next to the items concerned. The questionnaire should provide clear guidance on what level of data quality is required. A top down approach instead of the bottom up approach should be used, in which the questionnaire starts with a summary financial statement that gives the respondent overview of the response task and only asks specifications that are needed.

Retrieval and Judgment: Reduce the level of detail asked. Develop an electronically searchable list of financial items that makes it easier for respondents to find where specific items in their administration should be reported in the questionnaire.

Response: Development of an electronic questionnaire with automated additions to prevent calculation errors and to make it easier to correct errors (thus reducing the burden of correcting errors). In the electronic questionnaire “soft” warnings should be given for implausible values to prevent mistakes (but still allowing the continuation of the completion if the respondent accepts these values after the warning). For the paper questionnaires the layout must be improved to make it easier to report numbers in the correct line and make the correct additions of items. Names of contact persons should be asked in the questionnaire and respondents should be asked to sign for accurate reporting in order to increase the psychological barrier for sloppy reporting. Enforcement policies should be expanded to the quality of the data provided.

Many of the recommendations above were implemented in the redesign of the SBS survey. The main changes to the design included the use of tax data to exclude small businesses from the data collection; a reduction of the number of items on the questionnaires; the development of an electronic version of the questionnaire (see Snijkers, Onat and Vis-Visschers, 2007 for more details); the structure of the questionnaire was reorganized so that each topic was only treated once; the instruction text was reduced and placed close to the items and the usability of the paper questionnaires was improved. Redesigning the questionnaires was an iterative process, including pre-tests with respondents in the field. A first evaluation of this new design shows that the response rates have stayed the same but responses come in quicker, also the pick-up rate of the electronic questionnaire is fairly high (about 80% of the respondents use the electronic questionnaire) and the actual and perceived response burden has been reduced (Giesen, Morren and Snijkers, 2009). Further research should show if and how the new design has affected the costs of the data collection and the quality of the collected data.

References

- Anderson, A., Morrison, R. and Brady, C. (2005). Applying knowledge of Business Survey Response Processes to the Design Collection Software at the U.S. Census Bureau. Paper presented at the Federal Committee on Statistical Methodology Research Conferences, Arlington, Virginia, November 14-16 2005.
- Bavdaž, M. (2010a). The multidimensional integral business survey response model. *Survey Methodology*, 36, (1), 81-93.
- Bavdaž, M. (2010b). Sources of Measurement Errors in Business Surveys, *Journal of Official Statistics*, 26 (1), 24-42.
- Cox, B., and Chinnappa. B. (1995). Unique features of business surveys, pp. 1-17 in B. Cox, D. Binder, B. Chinnappa, A. Christianson, M. Colledge, and P. Kott (eds.) *Business Survey Methods*. New York: Wiley.
- Eurostat (2011). *European Statistics Code of Practice - update version 2011*. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-32-11-955/EN/KS-32-11-955-EN.PDF.
- Giesen, D. (2007). *The Response Process Model as a Tool for Evaluating Business Surveys*. Proceedings of the Third International Conference on Establishment Surveys (ICES-3), 18-21 June, Montreal, Canada, pp 871- 880. Alexandria, VA: American Statistical Association.
- Giesen, D & Hak, T. (2005). *Revising the Structural Business Survey: From a Multi-Method Evaluation to Design*. Proceedings of the Federal Committee on Statistical Methodology Research Conference, November 14-16 2005, Arlington, Virginia.
- Giesen, D., Morren, M. & Snijkers, G. (2009). *The Effect of Survey Redesign on Response Burden*. Paper presented at the European Survey Research Association Conference 2009, Warsaw, June 29-July 3 2009.
- Gower, A.R. (1994). Questionnaire Design for Business Surveys, *Survey Methodology*, vol. 20, (2), pp. 125-136.
- Haraldsen, G. (2004). Identifying and reducing response burdens in internet business surveys. *Journal of Official Statistics*, 20(2), 393-410.
- Jones, J., Rushbrooke, J., Haraldsen, G., Dale T., Hedlin, D. (2005). Conceptualising Total Business Survey Burden. *The Survey Methodology Bulletin*. 55. 1-10.
- Phipps, P., Butani, S. and Chun, Y., (1995). Research on establishment-survey questionnaire design, *Journal of Business and Economic Statistics* 13: 337-346.
- Snijkers, G. (2007). *Between chaos and creation*. Inaugural lecture Utrecht University, Statistics Netherlands, Heerlen.
- Snijkers, G., Onat, E. and Vis-Visschers, R. (2007). The Annual Structural Business Survey: Developing and Testing an Electronic Form. *Proceedings of the Third International Conference on Establishment Surveys (ICES-3)*, 18-21 June, Montreal, Canada, pp 456-463. Alexandria, VA: American Statistical Association.
- Tomaskovic-Devey, D., Leiter, J. & Thompson, S. (1994). Organizational Survey Nonresponse *Administrative Science Quarterly*, 39, pp. 439-457.
- Tourangeau, R., L.J. Rips & K. Rasinski (2000). *The Psychology of Survey Response*. New York: Cambridge University Press.
- Willimack, D. and Nichols, E. (2010). Hybrid Response Process Model for Business Surveys. *Journal of Official Statistics*, Vol. 26, No. 1, 2010, pp. 3-24.
- Willimack, D., Nichols, E., & Sudman, S., (2002). Understanding unit and item nonresponse in business surveys. In R.Groves, D. Dillman, J. Eltinge, and R. Little (Eds), *Survey Nonresponse*. New York: Wiley.

ERIM Report Series <i>Research in Management</i>	
ERIM Report Series reference number	ERS-2013-004-ORG
Date of publication	2013-04-16
Version	16-04-2013
Number of pages	9
Persistent URL for paper	http://hdl.handle.net/1765/39658
Email address corresponding author	thak@rsm.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus University Rotterdam PO Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31104081182 Fax: +31104089640 Email: info@erim.eur.nl Internet: http://www.erim.eur.nl
Availability	The ERIM Report Series is distributed through the following platforms: RePub, the EUR institutional repository Social Science Research Network (SSRN) Research Papers in Economics (RePEc)
Classifications	The electronic versions of the papers in the ERIM Report Series contain bibliographic metadata from the following classification systems: Library of Congress Classification (LCC) Journal of Economic Literature (JEL) ACM Computing Classification System Inspec Classification Scheme (ICS)