Revising the Structural Business Survey: From A Multi-Method Evaluation to Design[1]

Deirdre Giesen
Statistics Netherlands, Department of Methods & Informatics, Kloosterweg 1, 6412 CN Heerlen, the Netherlands
igin@cbs.nl
&
Tony Hak
RSM Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
thak@rsm.nl

Introduction

At Statistics Netherlands (SN) there is an increasing interest in how data collection for establishment surveys can be improved to gain efficiency, reduce response burden and increase data quality. In 2003, a Data Collection Expertise Program was established to professionalize and coordinate data collection procedures for establishments (Snijkers, Göttgens & Luppes, 2003). As part of this program a drastic revision of the questionnaires of the Structural Business Surveys (SBS) was planned. Three main projects can be distinguished in this revision process: 1) a project aimed at reducing the output variables by critically examining the legal and statistical necessity of each variable; 2) a project aimed at reducing the data collection by using administrative data; and 3) a project aimed at improving the remaining primary data collection. This third project includes both an evaluation and redesign of the current, paper, questionnaire as well as the development of an electronic questionnaire.

This paper describes the evaluation and redesign of the paper questionnaire. We describe how and to what extent various (qualitative and quantitative) evaluation methods contributed to our understanding of the main problems of the SBS questionnaires and how some of the evaluation results could be translated straightforwardly into solutions. We discuss in some detail why our evaluation findings were not decisive on the issue of the desirable approach to overall structure of the questionnaires and how we tried to collect this evidence in a pretest. We conclude the paper with an overview of the main lessons we learned with respect to the methods used in this project.

An overview of the revision process

The SBS program

The SBS questionnaires - also known as the Production Surveys or Annual Business Inquiry - measure a large number of indicators of the activity and performance of Dutch businesses. Variables collected include detailed information on sales and other revenue, expenses, inventories, purchases and employees. 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 program. Data are collected annually with an integrated set of questionnaires. Each questionnaire consists of a uniform core part that is the same for all businesses and a more or less unique part, containing questions specific for type of activity and size class. There are over 180 industry specific versions of the SBS questionnaire. Questionnaires are automatically generated from a question database.

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 and follow-up strategies for non-respondents are more intensive for statistically "crucial" firms. For all firms response is mandatory by law. In 2003 over 84,000 questionnaires were sent out, with a response rate of 70%. Of all SN establishment surveys, the SBS ranks second with respect to response burden, measured as the time needed to fill out the questionnaire.

[1] The research reported in this paper was conducted by a project team at Statistics Netherlands. The authors of this paper gratefully acknowledge the contribution of all members of the project team. The views expressed in this paper are those of the authors and do not necessarily reflect those of the members of the team or the policies of Statistics Netherlands.

The questionnaires
SBS data are (still) collected through the mail by means of paper forms. The questionnaires are long; more than 15 pages are typical. The structure of the questionnaire consists of three main parts. The first part taps sales, other revenue and costs according to the definitions of SN. This part is practically identical for all industries and all size classes. The second part of the questionnaire is a summary of the profit and loss account. This summary starts with the total amounts of the revenues and costs reported in part one. Together with the financial results, the provisions and extraordinary items this sums up to the operating profit before taxes. The third part of the questionnaire contains industry specific specifications of revenue and costs.

Each questionnaire is sent out with an introductory letter and a so called 'comments sheet', which respondents can use for comments about the questionnaires as well as for reporting a change of address or of the name of the firm and for requesting an extension of the submission deadline. In the questionnaire, questions are printed on the right hand page and instructions on the opposite left hand page. See Figure 1 for an example of this design.

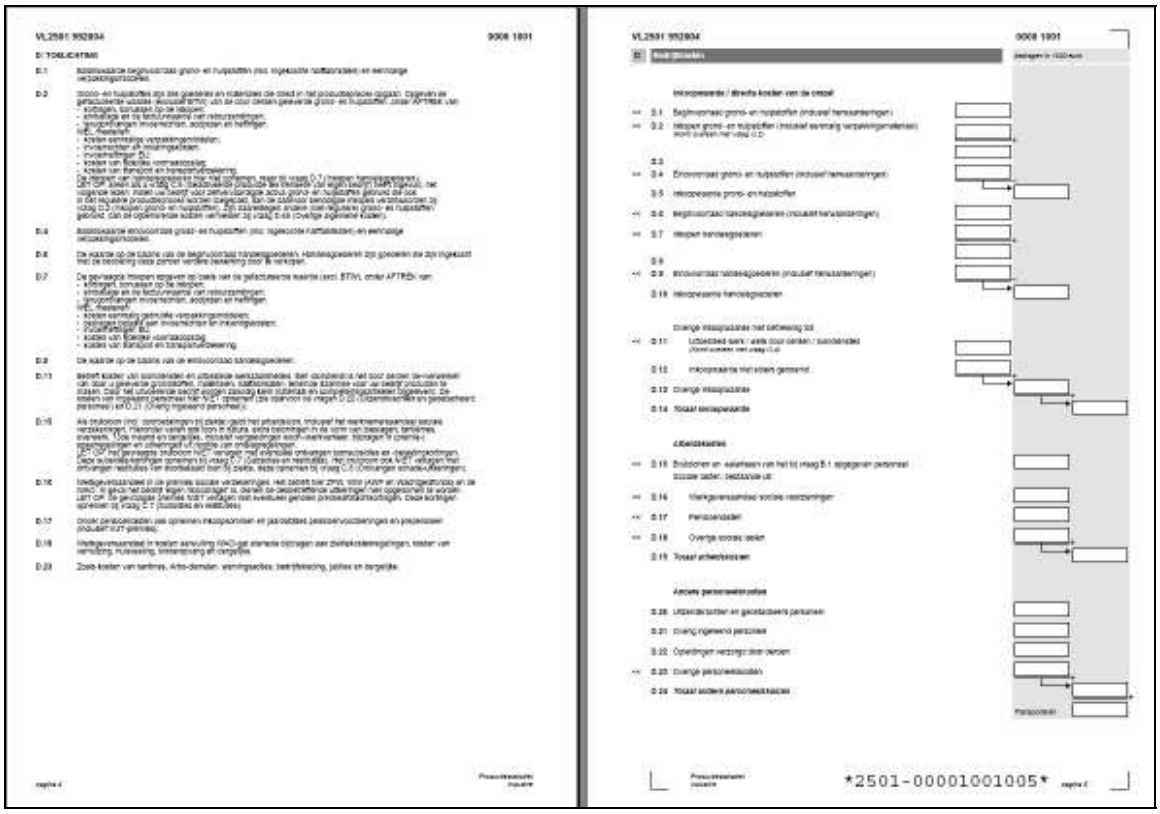


Figure 1: Design of the SBS Questionnaires

Outline of the revision process
The SBS revision process encompasses five main phases. First, in the **problem finding** phase, the quality of the current questionnaires, in terms of response burden and data quality, was assessed by means of multiple evaluation methods. The results of this phase served as an input for the **diagnostic** phase, in which the results of the problem finding phase were validated with qualitative field research and likely causes of response burden and data error determined. These two steps resulted in recommendations for the revision of the questionnaires that were implemented in the **design phase**. Next, in the **test phase**, two versions of a new questionnaire were designed and tested. The original planning was that in the **decision phase** (Fall 2005), decisions are taken on the new questionnaires which would be fielded in spring 2006. However, as will be explained later in this paper a new design and test cycle was added and the introduction of the new questionnaire is now planned for 2007. This paper reports on the first four phases of the SBS revision process and will conclude with our recommendations regarding content, design and structure of the SBS2005 questionnaire. It will, therefore, not include a description of the decision phase or of methods that will be used in a final test of the questionnaires.

The Problem finding phase

Aims and procedures

As the current questionnaires have already been in the field for a number of years, there was quite some process and survey data available that could be used for problem finding. Therefore, we first made a thorough analysis of available information. The aim of this phase was to identify groups of respondents, questionnaires and questions that seemed to be problematic with respect to data quality and/or response burden. In the next phase, the diagnostic phase, we would then explore the relevance of the initially identified problems and also diagnose them.

Previous reports. We identified seven previous reports about problems with the SBS questionnaires that were based on either information from respondents or from field staff. Apart from many interesting details about problems with question wording and overall design, we took two main insights from these reports:

- A recurrent theme in these reports was that comprehension of question wording is industry specific. We concluded that, ideally, the new questionnaires should be tested in all industries.
- It appeared that, in the past, respondents had complained about not having been informed timely about changes in the questionnaire. Many respondents receive the questionnaire every year and have made preparations in their records to be able to easily answer the questions. We concluded that we should inform our respondents timely of the changes that we were going to make.

Completed forms. We took a convenience sample of 66 questionnaires that had been returned by respondents in the previous years. The sample consisted of questionnaires from all main industries, it included both early and late respondents, and also contained both so-called crucial (mostly: large) firms and non-crucial firms. We tried to identify problems by looking at features such as crossed out words and numbers, write-in comments from respondents, and calculation errors. A side-effect of this approach was that we ourselves got a feeling for the structure and content of the questionnaires.

Figure 2: Example Of Completed Questionnaire With Calculation Errors

We learned two main things from looking at these completed forms:

- We observed many wrongly calculated (sub)totals. By inspecting the questionnaire itself (instead of looking at a dataset) we got an idea of how the *design* of the questionnaire caused these calculating errors. In Figure 2, for instance, it is not clear to the respondent that a total must be calculated of all numbers in the right hand column of this page and that this total should be written in the box at the bottom of the column. We concluded that we should look into possibilities to create a clearer questionnaire (page) *design.*
- We also learned that the *response process* is important. We concluded this from, for instance, seeing two kinds of handwriting on a form (raising questions about division of labor in the response process) or seeing that corrections had been made.

Respondents' questions, complaints and comments. Comments and questions 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. In 2003, 2232 comments (from 2149 different respondents) were received. Each comment was coded by a field officer. Table 1 gives an overview of the main codes used. For each main code additional sub codes were used, for example to indicate which specific question was addressed by a remark.

Table 1: Main Topics Of Respondents` Remarks About The Questionnaire

Main topic	Example	#	%
Specific questions	"Our company is run by the owners in partnership firm. Where should I put this on the form?"	291	13
General questionnaire problem	"For which company should we provide the numbers?"	26	1
Match with own records	"A small company like ours does not keep the records that are needed to answer these questions."	610	27
Response burden	"This takes far too much time to complete! Please remove us from your sample."	665	30
Quality of the answers	"These are all estimates; we do not have the final figures yet."	179	8
Other	"Can't you send a form in English? Our books are kept in Austria."	461	21
Total		2232	100

These comments from respondents helped us to identify problems in the questionnaires as well as specific groups of respondents that seemed to encounter more problems than other respondents.

Statistics Netherlands staff. Because we assumed that SN staff that work with either respondents or questionnaires or data had developed opinions about strengths and weaknesses of the SBS questionnaires, we arranged eight focus groups with field officers, helpdesk staff, data editors and data analysts (Heinink, Oostrum & Snijkers, 2004). These focus groups were homogeneous in the sense that in each group only one category of staff was represented. We arranged different focus groups for data analysts and editors in differing industries, e.g., a different group for analysts of the Construction industry and another one for analysts of the Manufacturing industry. These focus groups resulted in a rich overview of the problems originating from specific groups of questions, of particularly problematic categories of respondents, as well as of problematic aspects of the general SN approach regarding the SBS program. The focus group also yielded many concrete ideas for improvement. An additional benefit from the focus groups was that a large group of SN staff with an interest in the SBS questionnaires got involved in the evaluation process.

Statistical process indicators. We explored three statistical process indicators: unit non-response, item non-response and plausibility[2]. However, the interpretation of these indicators appeared to be difficult. We encountered the following problems in this part of the evaluation:

- Data quality is the result of the entire data collection process. If response and plausibility rates are low for specific categories of questionnaires or respondents, it is not clear how causes could be identified.
- We distinguished 42 broad categories of SBS questionnaires, each consisting of a group of largely similar questionnaires for businesses with similar economic activities and size classes (for example: the category of questionnaires for large wholesale companies, with slightly different questionnaires for different types of wholesale). It appeared that average plausibility varies considerably between subcategories. For example, data for wholesale of mineral products had a much higher plausibility than data for wholesale of livestock.
- Item non-response is difficult to interpret. It is often not possible to determine whether an empty field or a zero means 'not applicable', 'applicable but zero', 'refusal to answer' or 'don't know'. Therefore we made two analyses for item non-response. First we described questionnaires and questions by the percentage of empty fields. Secondly, we examined how many of these empty fields had been replaced by a figure in the process of (manual) data editing.

Despite these difficulties we learned several things from analyzing these statistical process indicators. First, based on unit and item non-response rates we could identify questionnaires and questions that seemed to be more problematic than others. Second, we were amazed by finding very high percentages of empty fields overall. For the 42 broad categories of SBS questionnaires, these percentages ranged from 34% to 66%. The smaller the company, the higher is the percentage of empty fields. The highest level of empty fields (more than 60%) is found in questionnaires submitted by businesses without employees. After data editing, only about 10% of these empty fields was replaced by figures. Although an empty field cannot be straightforwardly interpreted as either an instance of item non-response or as meaning 'not applicable', we developed the hypothesis that an empty field most often indicates that the item is not applicable, implying that many respondents might be faced with a large number of questions that are not relevant for their business.

Stakeholder check. After having completed these different forms of problem finding, in which we had tapped different sources of existing information in SN, we compiled a list of main problems concerning the overall communication with respondents (the way respondents are informed about the program and requested to complete the questionnaire), the selection of questions to be included in the questionnaire, the design of the questionnaire, and question wording. We wanted to conclude this problem finding phase by presenting our main results to a number of carefully selected SN colleagues who had experience with the SBS questionnaires and would be core participants (or at least facilitators) of the revision process. The main goals of this stakeholder check were (1) to make sure that we had a complete overview of the main problems with the questionnaire as known within SN, and (2) to create a consensus between stakeholders about the problem definition on which the revision of the SBS questionnaire would be based.

Main result of the Problem finding phase

The overall conclusion of the problem finding phase was that the questionnaires seem to cause a high response burden and that the quality of the data collected may be questionable for certain items and subgroups. We concluded from this that it would be useful to spend resources to further investigate these problems and try to solve them. The problem finding phase resulted in a tentative list of questionnaire problems, varying from more general issues, for example the communication with respondents, to specific questions/variables that were identified as problematic. This list served as a guideline to select respondents and create a topic list for the diagnostic phase.

In our analysis of the data we compared the different sources to interpret and "weigh" our findings. We interpreted overlap in findings as a validation of the problems identified. In table 2 we present the main findings of the problem finding phase as well as the sources in which evidence was found for these problems. With respect to the more general questionnaire issues the sources were fairly consistent. However, the evidence of the process indicators with respect to the issues of response burden and motivation must be characterized as 'circumstantial'. We interpreted low rates of spontaneous unit response and a high level of empty items of the questionnaire as a corroboration of the findings from other sources.

---

[2] For each submitted form a *plausibility index* is routinely calculated to help decide whether a form should be manually edited or not. The plausibility index is mainly an index of deviation from expected values that are computed from tax data, data from previous years and data from comparable firms.

As for the specific problems with questions, respondent groups or design issues, the table shows that all sources were used in the determination of these problems. These sources were often, but not always consistent with respect to specific issues. For example, in the focus groups it was noted that the employee questions were easy to complete. However, we found some severe problems in the layout of these questions in the analysis of the completed forms that were well supported by data inconsistencies found in the item response analyses.

Table 2: Main Findings from Problem Finding Phase, with Sources of Evidence

Findings	Example	Evidence from				
		Reports	Completed Forms	Respondent Comments	Staff focus groups	Process indicators
Response burden is high	Respondents report that the questionnaire is too time consuming.	X		X	X	(X)
Motivation to complete form is low	Respondents do not understand why they should provide the detailed information asked.	X		X	X	(X)
Communication about the data request is not satisfactory	Respondents who receive the questionnaire every year complain about the advance letter that tells them they are in a sample.	X		X	X	
Structure of the questionnaire seems a source of data error and response burden	Questions in general part remain empty, whereas specifications of these questions in second part of questions are answered.	X	X	X	X	X
List of problematic questions that seem to be prone to data error	Questions about quantities in stead of costs are not well answered.	X	X	X	X	X
List of problematic groups of respondents that seem to have more reporting problems	Small firms are problematic respondents			X	X	X
List of possible flaws in the form design	From the layout it is not always clear which items should be summed up	X	X		X	X

The Diagnostic phase

Aims and procedures

The aim of the diagnostic phase of the revision process was to validate the results of the problem finding phase, to explore the likely causes of response burden and data error, and to generally describe the response process of the SBS questionnaires in order to understand how errors happen. We assumed that this kind of information could best be gained by studying how actual respondents in actual practice deal with a questionnaire. We conducted real-time on-site observations in eleven companies of how respondents dealt with the SBS questionnaire. We also conducted twelve on-site retrospective focused interviews on how respondents had actually completed their forms. In our on-site data collection we used topic lists based on the response process model for establishment surveys (Sudman et al, 2000; Willimack&Nichols, 2001). Additional data were collected in four less focused interviews with respondents, fourteen interviews with non-respondents about their reasons for non-response, and one expert review with a document designer.

Expert review. A document designer from an external design firm was asked to review the questionnaires and give us advice about desirable changes in the design. He identified several flaws in the design of the questionnaire that could explain a number of the problems found. Chronologically this expert review was part of the problem finding phase. This meant we could use the expert's comments for our topic lists for the field work in the diagnostic phase. However, as the expert review was not on based on empirical experiences with the SBS questionnaires from the SN organization, we place this step in the diagnostic phase.

Real-time on-site observations. Each visit was conducted by a team of an SN field officer and an interviewer. We visited eleven firms and observed how the respondent completed the forms. Visits typically started in some representative meeting room and ended at the respondent's desk (and PC). During the observation we tried to refrain from interrupting or influencing the response process. Afterwards we conducted a debriefing interview and then the field officer made corrections on the form – if necessary – or gave additional information about the questionnaire. The advantage of observing the actual completion of a questionnaire (on-site and in real-time) is that things can be noticed that would be hard to reconstruct retrospectively.

Retrospective focused interviews. There are also drawbacks to observation. Previous research (Hak & Van Sebille 2002) showed that many respondents do not normally complete the questionnaire in one sitting but space this over multiple sittings on as many days. Requesting them to do everything in one sitting would be a disruption of their work routines and might change their response process considerably. Distortions might also be the result of the fact that respondents who are observed might make a bigger effort, might use the observers as informants when they encounter problems, or might feel that are under pressure to finish quickly. Because of these potential drawbacks of observation we also visited twelve firms who had already completed and returned their SBS questionnaires. Following the method used in the pilot study by Hak and Van Sebille, we interviewed these respondents about how they had completed their form, carefully reconstructing how they had arrived at their answers.

Respondent interviews. We interviewed four other respondents about their experiences with and opinions of the SBS questionnaire. We conducted these interviews with two respondents who were not able to complete their forms because their annual financial report was not yet available. Two other respondents worked at an administrative office that handled the SBS questionnaires for many different firms.

Non-respondent interviews. To explore whether non-respondents had specific problems with the questionnaire that might differ from problems experienced by respondents who had completed the questionnaire, we also conducted telephone interviews with fourteen non-respondents who had not responded to the SBS2003 data request. It appeared that most non-respondents had refused for other reasons than questionnaire characteristics. Many had not even opened the envelope.

Main results of the Diagnostic phase

To fully complete the SBS questionnaire, respondents need to carefully map figures from their own records on items of the questionnaire. The time spent on completing the questionnaire varies widely: In our study the (observed or reported) completion time differs between 45 minutes and two and a half days. Even professional respondents who conscientiously try to complete the questionnaires make reporting errors. A particular source of frustration was the summary of the profit and loss account in the second part of the questionnaire. Here respondents often discovered that the sum of the revenues and costs reported on the first part of the questionnaire did not sum up to the totals in their own financial reports. We observed three strategies to deal with this. Some respondents went back to the first part of the questionnaire and checked all figures for

missing or double counted entries. This took often as much time as they had spent to fill out this part in the first place. Some respondents accepted (even large) discrepancies and did not bother to make the numbers match. Others used an "other costs" category to make the numbers fit. One respondent explained that she would always start with the second part of the questionnaire and than go back to the first part to fill out the details. This is also the strategy of the field officers when they have to fill out a questionnaire.

The most important causes of observed errors were: lack of motivation / time, reporting about another business unit than intended by SN, interpretation errors, calculation errors and writing errors. Overall, the SBS questionnaires seem to cause a high response burden and to be prone to reporting errors.

Most respondents said that they have the impression that their work is not important to SN and that it does not matter whether they respond at all or how seriously they take their task. Most respondents did not see the purpose of being asked to report their data in such detail. The standardized letters and follow-up strategies used by SN seemed to cause additional resistance. Overall, the communication of SN towards respondents about the SBS questionnaires seemed to be unsatisfactory.

The previously mentioned high item non-response in small businesses and the large number of complaints by respondents about the fact that the questionnaire does not apply well to small businesses were corroborated by various observations. For example, when we observed how the owner of a small nursery completed the questionnaire, we found that most of the concepts of the questionnaire were unfamiliar and irrelevant to this respondent. The respondent did not understand the questions *and* did not understand his own financial reports which were produced by his accountant. Even though he made a serious and rather frustrating effort to complete the questionnaire, he made so many errors that the obtained data were virtually useless. Such observations confirmed our idea that the SBS questionnaires are especially ill-suited for small firms and respondents without a background in bookkeeping.

Based on the results of both the problem finding and diagnostic phases, the following recommendations were made to the group responsible for the SBS questionnaires.

Recommendations concerning the general data collection strategy:

- Respondents should be convinced that the data request is legitimate, their participation is important and that they are treated respectfully.
- It should be made clearer about which business unit(s) the respondent is supposed to report.
- The questionnaire for small businesses (with zero or one employees) should either be abolished or completely redesigned as to fit with the knowledge and situation of this group of respondents.
- The necessity of all variables measured in the questionnaires should be critically assessed.
- The fundamental response problem of the SBS questionnaire is the fact that respondents need to reallocate items from one category in their own records to another one as defined by SN. This problem will never be completely solved, as businesses in the Netherlands are allowed to have very heterogeneous bookkeeping practices and SN needs comparable data for all businesses. This is in additional reason to further explore alternative ways of data collection by automated tapping of administrations with XBRL. Also, more expensive data collection methods, as for example visits by field officers might be necessary for the collection of high quality data for certain important businesses.

Recommendations concerning the questionnaire:
Next to several detailed comments about specific questions we formulated the following recommendations concerning the questionnaire:

- The questionnaire should be accompanied with a short general instruction on how to handle the questionnaire.
- The structure of the questionnaire should be changed so that
  - The questions of the first part of the questionnaire and the industry/business specific questions of the third part are integrated and all questions about a topic are asked in the same place.
  - The questionnaire should start with a summary of the profit and loss account; this could have a psychological advance of starting with something familiar and gives a good impression of the total response task that will follow.
- Essential instruction text should be placed near the question.
- References to additional instruction text should be more clearly indicated.
- Design should be improved to make clearer which items should be included in the computation of a total.
- The wording of the questions and instructions should be edited to make the texts more readable and

understandable. This should preferably be done by a team of a text-expert, a content matter specialist and a field officer.
- An electronic search tool should be developed in which respondents can automatically search where specific items in their records should be placed on the questionnaire.

The findings of the evaluation were documented in reports, but also presented to both the SBS group (the group responsible SBS data collection) as well as to the management of the Division of Business Statistiscs. In these presentations we used film fragments from our field work to illustrate our findings. For many in the audience this was the first time that they saw how respondents actually deal with the questionnaires.

The Design phase

Aims and procedures
The redesign of the SBS questionnaire was organized by the SBS group. This group consists of representatives from the departments that prepare and analyze the data and the department that actually makes the questionnaires. For the redesign phase a field officer, a form designer and the project manager responsible for the problem finding and diagnostic phase were added as advisers to this group. The goal of the design phase was to develop a new questionnaire that would address the response burden and data quality issues discovered in the evaluation of the questionnaire but that would also provide the data needed for all SBS data users.

The SBS group made decisions on both the content (variables measured) as well on the design and wording of the questionnaire. As a result of the 'output variables' project, project 1 mentioned in the introduction of this paper, a number of detailed questions could be removed from the questionnaire. The SBS group decided which general guidelines should be used for the redesign. For example, an important decision was that the wording of the instruction text could be edited to make the text more readable, even if this meant that not all legal content could be covered. A small team consisting of the form designer, the field officer and a methodologist from the data collection department than prepared concrete design and text proposals according to this new variable lists en design guidelines. This redesign was than applied to two examples of the SBS questionnaires.

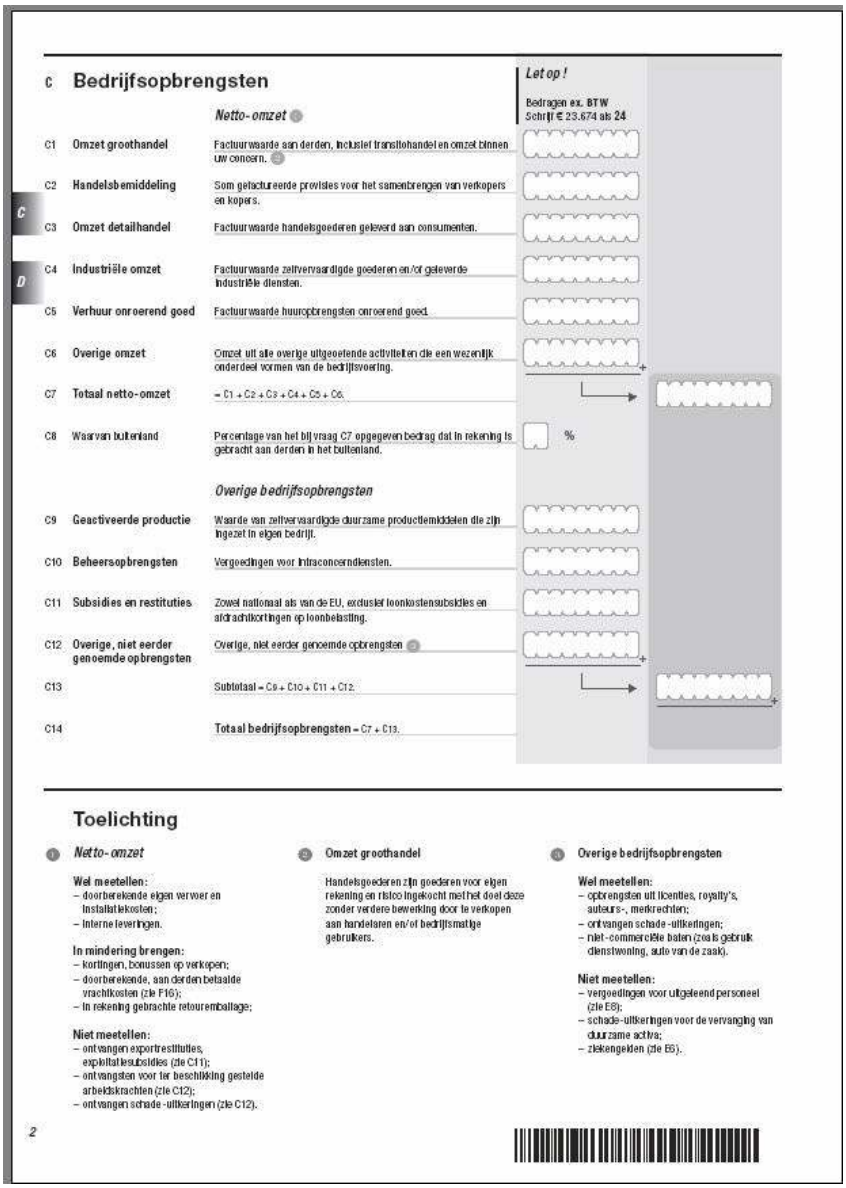
Most, but not all recommendations from the evaluation were accepted by the SBS group. It was decided not to spend scarce resources on developing a special questionnaire for small businesses. The reason for this was that SN can already reduce the number of questionnaires sent out to small businesses by over 18000 questionnaires as a result of the use of administrative data (project 2 mentioned in the introduction). In the future SN hopes to find ways to exclude all small businesses from the extensive annual data collection as done in the SBS questionnaires.

Main results of the Design phase
The questionnaire that resulted from the design phase differs from the old questionnaire with respect to the content (variables measured), the design, the wording of the questions and instructions and the structure of the questionnaire.

Contents. The content of the questionnaire was adjusted according to the results of the output project. This resulted in a small reduction of the items of the questionnaire. The general goal of this was to reduce the response burden by asking less detailed information.

Design. The following format for the presentation of items was developed. Each item is presented by three elements: on the left of the page a short label or keyword (such as "total revenue") indicates the topic of the item; in a space next to the right of the keyword a short explanation or instruction is provided and then on the right hand side of the page there is a box in which an answer can be written. If more text is needed than can be provided in the three lines allowed for the short instruction, additional text is displayed in a footnote. See figure 3 for an example of this design. The goal of this design is to increase the likelihood that respondents read the instruction texts. The respondent will "meet" the explanation text when scanning between the keyword and the answer box. By presenting the text in a short text column we hoped to increase the likelihood that respondents will not only notice the text but will actually read it. Many other adjustments were made in the design, such as the font used, the design of the items that should be counted together, and the addition of a small guideline between the question text and the answer box.

Wording. We tried to increase the readability of all text by using more understandable wording and decreasing the text length. Here the field officer had important input as to what concepts and terms would be understandable for the respondents.



C **Bedrijfsopbrengsten**

Let op!
Bedragen ex. BTW
Schrijf € 23.674 als 24

Netto-omzet

C1	Omzet groothandel	Factuurwaarde aan derden, inclusief transitohandel en omzet binnen uw concern.
C2	Handelsbemiddeling	Som gefactureerde provisies voor het samenbrengen van verkopers en kopers.
C3	Omzet detailhandel	Factuurwaarde handelsgoederen geleverd aan consumenten.
C4	Industriële omzet	Factuurwaarde zelfvervaardigde goederen en/of geleverde industriële diensten.
C5	Verhuur onroerend goed	Factuurwaarde huuropbrengsten onroerend goed.
C6	Overige omzet	Omzet uit alle overige uitgeoefende activiteiten die een wezenlijk onderdeel vormen van de bedrijfsvoering.
C7	Totaal netto-omzet	= C1 + C2 + C3 + C4 + C5 + C6.
C8	Waarvan buitenland	Percentage van het bij vraag C7 opgegeven bedrag dat in rekening is gebracht aan derden in het buitenland.

Overige bedrijfsopbrengsten

C9	Geactiveerde productie	Waarde van zelfvervaardigde duurzame productiemiddelen die zijn ingezet in eigen bedrijf.
C10	Beheersopbrengsten	Vergoedingen voor intraconcerndiensten.
C11	Subsidies en restituties	Zowel nationaal als van de EU, exclusief loonkostensubsidies en afdrachtkortingen op loonbelasting.
C12	Overige, niet eerder genoemde opbrengsten	Overige, niet eerder genoemde opbrengsten
C13		Subtotaal = C9 + C10 + C11 + C12.
C14		Totaal bedrijfsopbrengsten = C7 + C13.

Toelichting

1 *Netto-omzet*

Wel meetellen:
- doorberekende eigen vervoer en installatiekosten;
- interne leveringen.

In mindering brengen:
- kortingen, bonussen op verkopen;
- doorberekende, aan derden betaalde vrachtkosten (zie F16);
- in rekening gebrachte retouremballage;

Niet meetellen:
- ontvangen exportrestituties, exploitatiesubsidies (zie C11);
- ontvangsten voor ter beschikking gestelde arbeidskrachten (zie C12);
- ontvangen schade-uitkeringen (zie C12).

2 **Omzet groothandel**

Handelsgoederen zijn goederen voor eigen rekening en risico ingekocht met het doel deze zonder verdere bewerking door te verkopen aan handelaren en/of bedrijfsmatige gebruikers.

3 **Overige bedrijfsopbrengsten**

Wel meetellen:
- opbrengsten uit licenties, royalty's, auteurs-, merkrechten;
- ontvangen schade-uitkeringen;
- niet-commerciële baten (zoals gebruik dienstwoning, auto van de zaak).

Niet meetellen:
- vergoedingen voor uitgeleend personeel (zie E6);
- schade-uitkeringen voor de vervanging van duurzame activa;
- ziekengelden (zie E6).

2

Figure 3: Example Of The New Questionnaire Design

Structure. The questionnaire structure was changed in two ways. First, the industry specific questions (part 3 of the old questionnaire) were integrated with the general revenue and costs questions (part 1 of the old questionnaire). The result of this is that all revenue and costs groups now only appear once in the questionnaire. For example, the detailed costs of an industry specific tax are now asked directly after the general questions about taxes paid.

Second, we experimented with the placing of the summary of the profit and loss account. Two contradictory approaches were tried out: a "bottom-up" and a "top-down" approach. The structure of the present SBS questionnaire is bottom-up. It begins with asking for details about specified categories of revenue and costs and the aggregate of these items is the basis of the summary of the profit and loss account. The findings from the diagnostic phase suggested that a top down approach might be easier for respondents. In this approach the questionnaire starts with a summary of the core financial data *mostly according to the respondents own definitions* and then asks details according to the SN definitions on the items within these broad categories. In this approach response burden would still be relatively high in the more detailed ("down") part of the questionnaire, which still could result in data error and item non-response. However, we expected that providing the core financial data in the beginning of the questionnaire would reduce the overall response burden in several ways. First of all, it would be a positive experience in the beginning of the questionnaire to start with a rather easy task which results in a general

and recognizable overview of the financial situation of the business. Second, this overview would also give the respondents a general outline of the content of the questionnaire. Third, in this version of the top-down approach it was no longer necessary that the total costs and revenue in the detailed questions would add up to the numbers provided in the overview.

For certain revenue items, for example income from rent, it is only asked of respondents to report how much this income was and in which item of their profit and loss account is included. One of the consequences of this approach is that more data editing will be needed to translate the reported numbers to the desired output variables. Another consequence of this approach is that the respondent is not forced by the questionnaire to check if all numbers add up, and might thus make more reporting errors. Because we did not have sufficient evidence to make a good choice between the "bottom-up" and "top-down" approaches, we decided to develop and pretest both approaches.

Testing phase

Two pairs of questionnaires were designed by applying both the "top-down" and "bottom-up" approach to a large questionnaire (for wholesale food) and a small questionnaire (for IT service companies). The four draft questionnaires were tested in a lab test with eight SN staff. These staff was asked to complete a questionnaire using a set of financial data of a virtual company. These draft questionnaires were also discussed with other SN staff in individual expert interviews, and written evaluations were solicited and received from SN field officers. Major flaws, identified in these expert commentaries, were immediately corrected. Many other comments were documented and kept for being used in the final stage of the SBS revision process. With this preparation we hoped to prevent wasting expensive field-test time on errors that could easily be detected by SN staff.

Field-test

Our field-test consisted of the same two approaches as used in the diagnostic phase, on-site real-time observations (with debriefing) and retrospective focused interviews. In fifteen companies we conducted on-site real-time observations of respondents completing these questionnaires. These respondents were visited by two SN staff, a field officer and an interviewer. They observed the response process and then conducted a debriefing interview. In seven companies we conducted retrospective interviews. We sent these respondents a questionnaire and asked them to complete it before the interview in which we discussed the way they had understood the data request (in its entirety as well as for separate items) and had organized their response. In four companies we could only discuss the questionnaire because it could not be completed by the respondent for various reasons. We had planned to have about as much retrospective interviews as observations, but as some respondents had not completed the questionnaire when we visited some retrospective interviews were changed into observations. For this field-test we selected only non-respondents of the latest SBS data request (the SBS2003, fielded in 2004). The reason for making this selection was that we wanted to request the considerable additional burden of this field-test only from respondents who had a legal obligation to complete the questionnaire and had not yet done so. We made 26 company visits, twelve with food wholesalers and fourteen in IT service companies. Three visits were video taped. We made use of checklists in the observations with debriefing as well as in the retrospective interviews.

Pretest results

Our main overall result is that, from the respondent viewpoint, the new questionnaire has a better design than the existing one, but response burden is still considered too high.

Contents. Although the new questionnaire was shorter than the existing one, a minimal completion still requires about 90 minutes on average. Most respondents do not see any use (neither for themselves nor for others) in reporting the requested information in the required level of detail.

Design. Respondents report that this questionnaire is 'prettier' and is more respondent-friendly. The short clarification or explanation that is given with each item is highly appreciated. We observed how most respondents indeed read the short explanation in the second column. Respondents also easily understand the logic of requesting information by first offering a keyword, followed by an explanation, and then a response box. This was apparent on occasions where this design was not consistently applied. For instance, respondents tended to skip questions where the first key word was missing. With the new design hardly any errors were made by writing the answer in the wrong box (one line up or down). Also, much less errors were made in the calculation of totals, because it is much clearer which boxes should be included in the calculation and which ones not.

Wording. From the fact that hardly any difficulty in understanding the questionnaire text was observed (by us) or mentioned (by respondents), we conclude that wording has improved sufficiently. This does not imply, however, that every respondent is able to fully understand the data request.

Structure. The topic-oriented integration of the questionnaire seemed to work well. That is, we heard no more negative comments about this aspect of the structure of the questionnaire. Our conclusions concerning the top-down or bottom-up approach are less straightforward. Preferences regarding these to approaches differ between respondents, but most like the "top-down" approach more (which was expected). Those who like this approach tend to be very enthusiastic about it. They are particularly positive about the fact that they recognize this general outline from other reports they make, that they need less adjustment of their own logic (i.e., the logic of their annual reports and their bookkeeping system) to the SN logic and that they do not need to go through the frustrating process of checking all the numbers once they have completed the summary of the profits and loss account. This might have a huge positive effect on (real and subjective) response burden. One respondent mentioned that he spends usually about two hours on completing the SBS questionnaire, but now it took only 50 minutes. This fits our overall observation that completion of the "bottom-up" questionnaire takes twice as much time as completion of the "top-down" version. We found however two disadvantages of the "top-down" approach. The way the top-down approach was implemented still asked for some consistency between the numbers provided in the summary of the profit and loss accounts according to the respondents own books and the numbers provided according to the SN specifications. This was confusing for some respondents. Another import draw back of this approach that as less controls are forced by the questionnaire, respondents seem to be more likely to make errors. A closer inspection of the errors made by respondents in our sample showed that two respondents of the top-down approach had accidentally skipped questions, something that did not happen in the bottom-up approach. Otherwise we could not, as we had expected, detect a clear tendency in the data to make more reporting errors.

We concluded from the test that the trade-off of a considerably more respondent-friendly questionnaire, with considerably less response burden, might be that data quality (at least: the completeness) is reduced. However, our qualitative test approach does not allow for any reliable conclusions as to how much response burden and data quality might be influenced by the top-down approach in our population.

Response burden and data quality. The overall conclusion of the pretest is that reduced content, improved design, and improved wording result in a reduction of both response burden and data error. A change in questionnaire structure ("top-down" rather than "bottom-up") seems to be the most effective way of further reducing response burden (though still considered too high by respondents) but it is likely that is also reduces data quality. This does not mean that all sources of data error are targeted by the changes made in the questionnaire. A lot of error sources still exist, such as calculation mistakes, unwillingness to search for a correct figure in the records (resulting in sometimes poor estimates or item non-response), and the fact that some respondents just do not understand the data request. In (very) small IT service companies in particular, we encountered respondents without any knowledge of bookkeeping or financial reporting who did not understand basic concepts as used in the questionnaire. It is unlikely that any revision of questionnaires could address that problem without a substantial reduction of the variables measured.

Recommendations

Our main recommendations are that new questionnaires should have the design as tested in the pretest and that (both real and subjective) response burden should still further be reduced. The later can be achieved by further reducing the number of items in the questionnaire (reduction of "content"). Adoption of a "top-down" structure would result in a further considerable reduction of response burden but, when adopted, this should be done a bit different from the tested approach: the respondent should not be forced at all to make the figures of their own summary of the profit and loss account consistent with the specifications of SN. Where necessary for data editing it can be asked for more items where they are included in the profit and loss summary. Further research might be needed to determine how the completion (and submission) of a minimum level of necessary data can be achieved and to what extent the data quality of the output variables is influenced. Also it was recommended that SN should look into ways to further motivate respondents to provide the requested data. This motivation should work both on positive and negative incentives. Positive incentives could include a better communication with respondents, specific feed-back for respondents with bench-mark information and the development of a general PR strategy to explain the usefulness of the work of SN. The negative incentives should make it clear to respondents that they are not only legally obliged to fill in the forms, but that they are also obliged to provide correct numbers. This awareness might be increased by adding a statement to the questionnaire in which respondents have to declare that the information provided is correct. Also, SN could announce and actually perform random quality controls, for example with field officers on site.

Finally, it is (again) recommended that an entirely different approach is developed for respondents in small businesses who have not a sufficient understanding of bookkeeping to be able to answer all SBS questions about their records.

Based on our findings it was decided that the new design and wording should be implemented for the SBS questionnaire. A further reduction of variables was not deemed feasible. Also it was decided that the new SBS questionnaire will have a top-down structure in principle but that should be investigated how the data editing systems should be changed in order to process the data from this structure. Currently the SBS group is working on the implementation of these decisions. The final questionnaire that will be the result of this process will again be tested. The specific set up for this test depends on to what extent the final questionnaire will differ from the previously tested one. A minimum test should include a test for industry specific understanding of the wording. The new questionnaire will hopefully be fielded in 2007.

Overall conclusions

Next to the conclusions about the questionnaire itself, we learned several methodological lessons from this project. In the problem finding phase the more qualitative methods were very useful as a structured preparation for the field work with respondents in the diagnostic phase. In this study the quantitative analyses (the unit response, item response and plausibility indices) were less useful as a tool to detect questionnaire problems. Given the complex structure of the SBS questionnaires it was often impossible to disentangle effects of questionnaires, respondents and approach strategies. However, the quantitative data were very useful in the interpretation of the findings of the qualitative analyses.

In the diagnostic and testing phase we found that the combination of both 'concurrent' observation as well as retrospective reconstruction was a very useful method to collect data to help understand which problems occur in questionnaires and why they happen. These different ways of data collecting complement each other very well. The observations provide essential details about what actually happens when filling out the questionnaire, whereas the retrospective interview gives insight in the overall response process as it has occurred, without any disturbance by the researchers. We also found that collecting data with a team of an interviewer and a field officer was a good approach. The field officers have the expertise to determine where the respondents make reporting errors and to help and instruct them where necessary. As the usual work of field officers is to collect survey data, they are not trained to unobtrusively observe respondents and to collect qualitative data on the response process. The role of the interviewer is to make sure that this kind of test data is collected as well as possible. Finally, we want to mention that videotaping of some visits was very useful for the training of the interviewers, for the analyses of the data, and for the illustration of the results of this study. Showing the SBS group and the management how respondents actually work with the questionnaire helped create support for some radical changes.

A drawback of the qualitative approach in our testing phase, was that we could not make a clear choice between the top-down of bottom-up approach. We could only conclude that in our sample the top-down approach seemed to have a large positive effect on the response burden and a small negative effect on the data quality. However, our research design did not allow for any quantitative generalizations about to what extent these effects would occur in the population. The latter kind of information would have been very helpful for a far-reaching decision such as to what the main structure of the questionnaire should be. We would have liked to assess to what extent these positive and negative effects would actually influence data quality and response burden, also in particular because the implementation of the top-down structure requires an expensive adjustment of the data editing software. If such difficult and far-reaching choices can be foreseen in a future project, it would be wise to include a quantitative experiment in the redesign research plan.

References

Hak, T. & Van Sebille, M. (2002) *Het respons proces bij bedrijfsenquêtes. Verslag van een pilot studie.*[The response process in establishment surveys. Report of a pilot study] Rotterdam / Voorburg: Erasmus Research Institute of Management / Statistics Netherlands.

Heinink, J., van Oostrum, H. & G. Snijkers (2004) *Evaluatie vragenlijsten PS2004: Focusgroepen met CBS medewerkers*. [Evaluation Questionnaires PS2004: Focus groups with SN staff]. Heerlen: Statistics Netherlands

Sudman, S., Willimack D.K., Nichols E. & Mesenbourg T. (2000), *Exploratory Research at the U.S. Census Bureau on the Survey Response Process in Large Companies*. Invited paper presented at the Second International Conference on Establishment Surveys – ICES II, Buffalo (NY).

Snijkers, G, Göttgens, R. and Luppes, M. (2003) *WaarnemingsExpertise Programma(WEP): Naar een professionele waarneming* [Data Collection Expertise Programme. Towards Professional Data Collection] Heerlen: Statistiscs Netherlands

Willimack, D.K. & Nichols E. (2001), *Building an Alternative Response Process Model for Business Surveys*. Paper presented at the AAPOR annual meeting, May 2001, Montreal.