Netherlands Official Statistics

Special Issue
Integration of household surveys
Design, advantages and methods

Volume 13, Summer 1998

Voorburg

Prinses Beatrixlaan 428
P.O. Box 4000
2270 JM Voorburg (Netherlands)

Telephone : . .31 (070) 337 38 00
Fax : . .31 (070) 387 74 29
E-mail: lhka@cbs.nl
Internet: http://www.cbs.nl

Heerlen

Kloosterweg 1
P.O. Box 4481
6401 CZ Heerlen (Netherlands)

Telephone : . .31 (045) 570 60 00
Fax : . .31 (045) 572 74 40

Key figure A-125/1998

Subscription: Dfl. 42.00 per year
Price per copy: Dfl. 20.00

ISSN 0920-2048

Postage will be charged.

Contents

Editor in chief:
Jeroen Pannekoek

Guest editors:
Bart F.M. Bakker and Jeroen W. Winkels

Coordinating editor:
Lieneke E. Hoeksma

Why integration of household surveys? - Why POLS?

Bart F.M. Bakker and Jeroen W. Winkels

Summary

Statistics Netherlands has designed an integrated system of social surveys. Both this integrated survey and the programme used to build it are called POLS. In this introductory chapter we present the main incentives for this development: to meet new demands, develop new methods to link and enrich survey data, reduce non-response, improve data quality and reduce costs. The relationships between these incentives and the other contributions in this issue are specified.

1. Introduction

In the last three years Statistics Netherlands has developed an integrated system of social surveys. It is known under its Dutch acronym POLS: *Permanent Onderzoek Leefsituatie*. As in other European countries, in the Netherlands several surveys existed to gather information on living conditions. The history of these surveys goes back to the mid seventies or early eighties. They not only varied in content, but also in the sample frame used, sample size, response, interview length, the number of persons to be interviewed within the household, the use of proxy-interviews, type of data-editing and weighting method. Some of these surveys were revised in the beginning of the nineties when Statistics Netherlands introduced computer assisted personal interviewing (CAPI). These revisions focused both on the harmonisation of questions used for demographic and socio-economic classifications and on the 'translation into CAPI' of questionnaires that had originally been developed for 'paper and pencil' interviews (PAPI).
Several developments prompted us to go a step beyond the harmonisation of survey variables and design an integrated system of social surveys. In particular, these developments had to do with new demands for statistics on living conditions, the development of new methods to link data, the need to improve response in social surveys, the policy to design more respondent-friendly questionnaires in order to improve data quality, and the need to standardise the production processes in order to improve the quality and efficiency. The introduction and implementation of POLS were organised as a Business Process Redesign programme.

2. Meeting new demands

Rather than being concerned with the distribution of aggregate variables over groups of statistical units, social research frequently focuses on relations between variables at the micro-level. This focus on micro-relations has been the driving force behind the development of ever more comprehensive household surveys; and in order to be able to analyse all social relations and thus obtain a comprehensive data set, these surveys aim to cover as many variables as possible. However, there is a growing awareness of the limits to this approach. The response burden on the sampled households becomes too heavy if too many variables are covered in a single survey. This burden forms a serious constraint on statistical agencies to meet the demands.
To some extent these increasing demands have to do with the fact that there is no consensus on a list of ‘core’ indicators in the field of social statistics. The existing lists of social indicators vary greatly in size and content, and their theoretical basis is rather weak. This makes it possible for statisticians and politicians to change their demands relatively easily with respect to the requested statistical information. Also the output databases of statistical offices seem to be hampered by the rather fragmented structure of social statistics compared with economic statistics.

The drive to retain existing indicators in a survey in order to make time series, turns it into something of a pressure-cooker: the response burden only increases. A profound way to select social indicators for a survey has to be based on the strategic use in policy and on some kind of behavioural assumptions. To move forward, Everaers (1995), for example, proposed the selection of indicators already integrated in separate conceptual frameworks which have proven their validity in the various fields of social science. In the meantime Eurostat had used a more pragmatic method to develop a list of 107 social indicators. Although this may take some pressure off the surveys in question, the list of 107 social indicators is only a first stop on the road to indicator development, and is mainly relevant in an international context (Oudhof & Everaers, 1996).

The depth of statistical information needed by policy makers also seems to change. The demand is no longer for rather 'one dimensional' statistics about the important political issues like health, crime, education, employment and income, but increasingly for information on the more complex relations between these themes. This demand reflects the policy questions arising from societal developments such as poverty, social exclusion and deprivation. Such societal problems often converge in rather small groups in society or in problem areas, like inner cities, and these developments require special attention from statisticians if they are to fulfil the politicians’ needs. In terms of 'variable language': the demand is not for the univariate or bivariate distributions of variables, but for the multivariate relations between sets of variables. These sets give information on the *accumulation of effects* that has recently become a policy issue.

Another demand by users was for adequate flexibility to add questions on certain topics or oversample specific population groups during certain periods. By improving the flexibility, it would be easier to fulfil user needs. The urge for flexibility has much to do with the serious problem that nearly always arises within the agency when users of the data ask for addition of new information. The choice for a person based sample frame (see section 4) with elementary demographic information is conducive for the flexibility of oversampling because oversampling on these demographic characteristics is rather easy. In the future this sample frame will be combined with other registers (see section 3) enhancing the possibilities of oversampling of particular population groups.
Furthermore, policy users want monitor information, i.e. time-series on the most important policy issues: ethnic minorities, for example. To fulfil this need, it is necessary to harmonise the different surveys and combine the information from them into one database. For instance, when information on ethnic minorities is taken from different surveys, the definition of different ethnic groups should be similar. In particular it should be possible to select politically relevant target groups in a similar way from different surveys. Another existing demand was for the possibility to detect information on several phenomena with a small incidence; victims of a range of offences, for example.
To fulfil both existing and new demands the sample size should also be large enough to allow the description of relatively small subgroups and phenomena with a small incidence. The POLS

questionnaire, for example, consists of a joint questionnaire for a large sample and subsequent questionnaires for subsamples. The sample size of the large sample should be large enough to make it possible to give reliable information on small subgroups and on phenomena with a small incidence. Relatively small subgroups and phenomena with a small incidence then, could only be described in POLS if all the social-cultural surveys were included. However, the estimated net response of 34,000 households in 1997 forces Statistics Netherlands to extend POLS to include other surveys to find the smallest target groups and phenomena required by policy makers.
We avoided the possibility of selecting subsamples only by means of filter questions in a joint questionnaire for a large sample and in a subsequent questionnaire for the subsample. A filter question is defined as a question for the large sample which has some correlation with the phenomenon under study in the smaller subsample. This would make it possible to give respondents in the large sample a somewhat higher chance of being in a particular subsample, depending on the information from the filter question. The difference between a filter question and a routing question is that a routing question has a chance of 0 or 1 to continue the interview with a particular part of the questionnaire, while for filter questions these chances vary between 0 and 1. We avoid this as a general principle in POLS because such a procedure would lead to subsamples which are not representative for the population. If only 'crime victims' are asked about 'feeling safe in their neighbourhood' or only 'the disabled' about their health, no population estimates about victimisation and health are possible, because no weighting procedures are available to correct for the skewness of the sample. However, screening questions are used in POLS to select persons from the (representative) subsamples for a second interview (see the contribution of Winkels and Everaers in this issue). Screening questions are like routing questions, but are used to select persons for a second interview, while routing questions are used within the same interview.

3. Enrichment of survey data

Other reasons for changing the social surveys at Statistics Netherlands stem from information technology developments. In the first place the possibilities for linking large administrative registers containing information on living conditions are still increasing. It is the policy of Statistics Netherlands to make optimal use of this administrative information in order to reduce the response burden and costs. It is also necessary to link this administrative information with survey data to fulfil the user needs. As described in Van Bochove and Everaers (1996) there are basically three methods to link micro data from different sources.
The first is *exact matching*. In this case micro data concerning the same individual, but from different sources are linked. This requires identification of the individual by a matching key that is available in both sources. Apart from using these large data files for improving statistical information, the combined register information can be used as a sample frame. Within the available registers information on persons, households, jobs, social security benefits, addresses and dwellings can be found. These exactly matched registers are suitable as a sample frame for social surveys to collect data on behaviour or attitudes that almost by definition are not included in the registers. Oversampling particular categories from the population will also be much easier by using these registers.

The second way of achieving linkage is *synthetic matching*, where common variables with a common breakdown are selected. The micro data from all data sets are grouped according to this common breakdown; and within the resulting cells micro data from different sources are combined randomly. This approach does not give any privacy problems and matching between samples is no problem either. The price to pay is that the validity of the micro to micro relations derived from the synthetically matched data depends on the degree of association between the synthetic matching variables and the other variables. Within-cell joint variation is not picked up. A final method of micro to micro linkage is by *redesigning the sources* in such a way that there will be a joint questionnaire for a large sample that provides succinct information on core variables from all original surveys; and more in-depth-questionnaires on the separate areas for smaller subsamples. This way a core micro data set is obtained directly (the equivalent of exact matching) and an in-depth synthetic data set can be created by combining the in-depth data from the various sub-samples using the joint core as a synthetic matching key. Because the variables in the core are associated with the in-depth variables of the subsamples, this approach picks up relations more easily than synthetically matched data sets that rely on just demographic variables for the matching. The POLS project is an attempt to redesign the sources in this way. Another, more ambitious project, tries to combine exact matching and synthetic matching to build a database which contains all the relevant information on persons, households, jobs, social security benefits, addresses, dwellings, etc. The POLS survey will be linked by exact matching to the existing register information. The register information is used as a weighting frame, which makes it possible to reduce non-response bias. Moreover, the information will be completed and integrated with other register and survey information. The results of this project are expected to fulfil the demand on census information for the next European census.
Further on in the present publication, Jeroen Winkels and Pieter Everaers will present the design of the integrated socio-cultural survey, and Marion Hofmans will address the possibilities of using this design to reduce variances. Making use of core questions in the joint questionnaire for the large sample, a weighting procedure is developed to estimate population totals of target variables in the subsamples.

4. Reducing non-response

Non-response is a serious problem in household surveys, and compared with other countries it is relatively high in the Netherlands (De Heer, 1996). In the field of the socio-cultural surveys the level of response varies between 50% and 60%, and thus probably causes serious bias which cannot always be neutralised by smart methods of weighting. With respect to the four yearly National Election Study, for example, doubts have been expressed in the Netherlands about what the figures from this survey really tell us about political alienation, when so many people do not participate.

Statistics Netherlands has taken several initiatives to reduce non-response. Many of these initiatives have to do with how surveys are presented to the public (e.g. introduction letters), the monitoring of interviewers and the optimisation of using different data collection methods. These initiatives are always taken under the restriction that the overall respondent burden should be as low as possible.

One of the more important measures to reduce non-response was the use of a person based sample frame. Formerly, addresses were sampled and at one address one or more persons were interviewed. Statistics Netherlands now has the possibility to take samples from a file based on the fully automated population register of the Netherlands. This register contains information about nearly all persons and households in the Netherlands. This is very useful to combat non-response: an introductory letter can be sent to the person to be interviewed and can be personally addressed by the interviewers. Furthermore, it is possible to reduce the consequences of non-response, as some characteristics of the non-cooperators (sex, age, nationality) are known and can be used for reweighing.

Rob Vousten and Wim de Heer describe the efforts to reduce non-response in POLS elsewhere in this publication. They examine the effects of monitoring the interviewers, the use of mixed methods, the treatment of refusals and the use of a person-based sample frame.

5. Improving data quality

The quality of raw data can be judged by criteria such as relevance, timeliness, completeness, validity and answer reliability. If these criteria are to be met satisfactorily, commitment from the interviewers, respondents and users of the survey outcomes are all very important. Until the early nineties there were no serious studies at Statistics Netherlands to determine the survey data quality or to test the commitment of the persons concerned. The interviewers, however, were quite critical of the quality of the questionnaires, in particular the comprehensibility for interviewers and respondents.
Since then, the policy of Statistics Netherlands has been to aim at more respondent-friendly questionnaires, in order to improve data quality. A Questionnaire Design Research Centre (QDRC) was set up, offering facilities and methods for designing and performing pretests as part of survey design and development (Akkerboom, 1996). The POLS questionnaire was designed in co-operation between the internal users of the survey outcomes and the QDRC.

One aspect of the respondent-friendliness of a questionnaire is the minimal burden on respondents. This condition can be viewed from both a micro and a macro-perspective. 'Micro' means that the mean interview-time for a person or within a household should not exceed 45 minutes. 'Macro' means that the total time of all the modules should be reduced by POLS.

To improve data quality, it is also very important to standardise and document the processes of data collection, data processing and data analysis. The new methods of linking data only work in practice when both hardware and software are modernised. The POLS design would in fact be impossible without the new version of the Blaise system for computer assisted interviewing. Blaise 3 meets the demands for specifying dependency relations between different questionnaires (the so called POLS Plug and Play Principle).

To improve data quality no proxy interviews were permitted in POLS for people over eleven years of age, unless the empirical relation with other characteristics is well known. This was decided beforehand, because proxy information is generally less reliable and because the burden on respondents of proxy interviewing is relatively high.

In their article elsewhere in this issue, Hans Akkerboom, Francine Dehue and Ger Snijkers present a general model of questionnaire development and the implementation of the model in the POLS project. Wim Faessen and Frans Kerssemakers describe the standardised production processes of POLS. They present the data collection with the use of laptop computers and the Blaise software, the process of data analysis and documentation.

6. Organisation in progress

In the Netherlands as in other countries the demand for statistical information is strongly on the increase while at the same time the willingness to participate in surveys is diminishing. These bifurcating trends lead to budgetary problems and challenge statisticians to respond with long-term strategies which combine statistical ambition and logistic economies. Van Bochove (1996) outlines how Statistics Netherlands started a process of innovation based on two widely recognised assets: data collection technology and integrated data frameworks.
This process of innovation has a parallel process in organisational reform. For the development and implementation of the integrated household survey (POLS), Statistics Netherlands started a redesign program, along the lines of a Business Process Redesign project. The organisation of the work, the culture, the content of the surveys and the information technology are all elements of this redesign. Special attention is given to the form and quality of the internal communication in the different phases of the restructuring process: from a simple idea, via formulation of the idea, the definition of the program and the implementation of the program in several well-defined and inter-related projects.
The standardisation of the production processes also has important consequences for the quality and efficiency of these processes . By standardising and adjusting the production processes of the social surveys, it is expected that the costs of data collection will be reduced. In their article, further on, Pieter Everaers and Hans Wajon give an overview of the organisational changes that have been introduced or stimulated by the POLS project.

References

Akkerboom, J.C., 1996, Labor für die Entwicklung und den Test von Erhebungsinstrumenten, In: *Statistisches Bundesambt: Pretest un Weiterentwicklung von Fragebogen*, Schriftenreihe Spektrum Bundesststistik, Band 9 (Stuttgart: Metzler-Poeschel)

De Heer, W., 1996, *International response trends. Developments and results of an International Survey* (London: Paper prepared for the second ASC International Conference 1-13 September 1996).

Everaers, P.C.J., 1995, *The integration of household surveys*. (Mondorf-les-Bains: Paper prepared for the seminar on 'The Future of Social Statistics').

Oudhof, K.and P.C.J. Everaers, 1996, Social indicators for Europe, In: *Netherlands Official Statistics*, vol. 11, winter 1996, pp. 17-20

Van Bochove, C.A. and P.C.J. Everaers, 1996, Micro-macro and micro-micro linkage in social statistics, In: *Netherlands Official Statistics,* vol. 11, winter 1996, pp. 5-16

Van Bochove, C.A., 1996, From assembly line to electronic highway junction: a twin-track transformation of the statistical process, In: *Netherlands Official Statistics*, vol. 11, summer 1996, pp. 5-36

Design of an integrated survey in the Netherlands. The case of POLS

Jeroen W. Winkels and Pieter C.J. Everaers

Summary

In this article we give an overview of the design of POLS as it was implemented in the first year (1997). We shall focus on the contents of POLS in general and on how the six formerly separate surveys have been integrated into one design. Somewhat more detail is given to what we see as the core of integration, the joint questionnaire. Furthermore we pay attention to the implementation of the various modes to gather the data within various subsamples and on some new oversampling features.

1. Introduction

In the previous article Bakker and Winkels elaborated on the main incentives of Statistics Netherlands in redesigning an important part of its households surveys. This contribution looks at some of these incentives in more detail. We describe how the design of POLS meets newly emerging user demands, concerning sample size and the clustering of the various thematic modules. The joint questionnaire plays a central role as one of the new tools to link and enrich survey data. The reduction of non-response especially comes to the fore when the sample frame and the mix of modes to gather data from households are described. First we present the design of POLS in general and the joint questionnaire in particular. Furthermore we pay attention to the implementation of the various methods to gather the data within various subsamples and on the oversampling features. The contribution ends with some remarks on future developments with respect to the design of POLS.

2. The crux of POLS: one frame and N modules

The design of POLS is based on a modular structure with a common frame. The common frame consists of a joint sample frame and a joint questionnaire. In principle there is no limit to the number of modules in time, so panel studies can be included. In its first year of data collection (1997) there were eight different modules (excluding the joint questionnaire). This will also be the case in 1998. Every module has its own characteristics (see Figure 1) and consists of various (thematic) blocks of questions.

Figure 1. Integrated System of Surveys on Living Conditions (POLS) (numbers refer to net sample in 1997)

module 1

Joint questionnaire N = 34,437	part 1	CAPI/CATI

module 2

Joint questionnaire N = 34,437	part 2,3,4	CAPI/CATI

module 3

Health	Justice		Accidents & time-use (partial)	Well-being	Youth	Housing
	Environment & time-use	Participation				
CAPI PAPI 10,897	CAPI PAPI 5,554	CAPI 3,776	CAPI PAPI 4,335	CAPI 3,279	CAPI 2,749	CAPI PAPI 1,639

module 4

Accidents	CATI	2,888

The *sample frame* of POLS is based on a new development: the possibility for Statistics Netherlands to use the fully automated population registers of all the municipalities in the Netherlands to compose person or address based sample frames. POLS uses a person based sample frame. Only people living in institutions or shelters and those who have already been interviewed by Statistics Netherlands in the last year are excluded. More details about the relationships between sample design and the various interview shells are given in Hofmans' contribution to this issue.

The central feature of POLS is that the total sample is confronted with a *joint questionnaire*. Each interview starts with this joint part, consisting of modules 1 and 2 (see below). After the joint questionnaire several modules 3 are presented to mutually exclusive subsamples 3. However, the order of the questions during the interview may differ, in order to achieve a respondent friendly questionnaire. For example, the questions on income that make up part of the joint questionnaire (module 1) are always asked at the end of the interview, when module 3 has been finished.

Sometimes sampled persons are asked if they want to join in a follow-up study. These follow-ups form the fourth module of POLS. A fourth module thus means that another interview will be held (face-to-face or by telephone) within a certain time period after the first interview.

The joint questionnaire was designed under the restriction that both telephone (CATI) and face-to-face (CAPI or PAPI) interviewing should be possible. The CATI part was used as a standard routine within POLS to reach certain types of non-respondents from the first fieldwork stage. At the end of this telephone interview the respondents are asked if they are willing to give a face-to-face interview (to gather information requested in module 3). This mixed mode procedure led to an increase in response of approximately 2%. The overall response rate in the first year of POLS was 61.9%. (more details about response are given in the article by Vousten and De Heer in this issue). As not all respondents who fill in the (joint) CATI questionnaire want to be contacted again, the total number of respondents with a joint questionnaire is not identical to the sum of the respondents in the various sub-samples. In 1997 6.5% of the respondents (34,437 in total) filled in only the joint questionnaire. The response rates in module 3 vary between 55.2 (housing) and 59.4 (health and working conditions). Figure 1 presents an overview of the design of POLS in 1997.

An interview then should ideally consist of three modules: the joint questionnaire (modules 1 and 2) and one of the in-depth thematic third modules. Module 4 was designed for members of the sample who passed certain screening procedures. In 1997 the screening procedure looked for certain types of accidents and injuries within the households of the sample persons. This resulted in 2,888 persons being asked to respond to a telephone questionnaire about the accident(s). As each of these respondents had also completed a joint questionnaire and one of the third modules, a lot is known about these accidents victims. Each of the modules consists of a number of question blocks. More details are given below.

3. The core of integration: a joint questionnaire

The joint questionnaire was gradually developed in 1996 by a team of experts on both socio-cultural household surveys and data collection. Much research was done to build this core of POLS. The article by Akkerboom, Dehue and Snijkers further on in this issue gives information on the process to improve the quality of the joint questionnaire. Theoretically the joint questionnaire consists of four parts, which are described below. These four parts can be seen as the implementation of tools to:

- harmonise statistical results;
- start the measurement of phenomena with low incidence;
- improve the precision of (some) population estimators regarding living conditions;
- be flexible with respect to actual user needs.

Joint questionnaire part 1: harmonised classification variables

This first module contains all the questions to be asked of every person in the sample. The questions use the harmonised classification and survey variables of Statistics Netherlands, both the demographic ones (age, sex, marital status, nationality, place of birth) and the socio-economic ones (education, socio-economic position, net household income and main source of income). The demographic questions within POLS make use of the so-called 'forward feeding' routines in CAPI: the name, date of birth, sex and marital status as given by the population registers and used in the sampling procedure are checked in the course of the interview.

Joint questionnaire part 2: core variables in the socio-cultural domain

This module contains the core variables of the socio-cultural surveys. From all the variables in the socio-cultural domain, the ones with the highest correlation with the other variables in the domain are chosen as the core variables. The reason this is named module 2 (of course, the interviewer is not aware of these terms) is because in a later stage core variables from other surveys to be incorporated in POLS will have to be added. This module is therefore less stable in content than module 1. Because these questions are also presented to the total sample, they allow the quarterly publication of important indicators, the annual publication at a low regional level or the annual publication for relatively small population categories. Moreover, this part of the questionnaire plays a central role in the implementation of the idea of consistent weighting that makes use of the core variables (see the contribution by Hofmans in this issue).

Joint questionnaire part 3: screening questions in the socio-cultural domain

This part varies from year to year. Although the large sample allows finding many particular groups, it is not possible to include too many screening questions in the joint questionnaire as the maximum interview length would then be exceeded. In 1997 the screening questions focused on accidents and injuries. The information gathered was used to ask a specific subsample to join in a follow-up study (module 4, see below).

Joint questionnaire part 4: variable block of questions

POLS has implemented the principle that a small part of the joint questionnaire should be reserved for topical themes and/or themes that require a large sample because of low incidence. This part can be varied twice a year. In 1997 and 1998 it was used for a block of questions on victimisation.

4. Renewing surveys: in-depth subsamples and a mix of data collection techniques

The third and fourth module consist of theme specific questions. However, themes are combined in such a way that multi-dimensional indicators can be calculated. For example, the questions on health and the questions on working conditions, that

were part of two separate surveys until 1996, are now combined into one module 3. Naturally, the contents of these modules 3 will change to meet the needs of the main users of the POLS data. Data on some topics have to be collected continuously, each and every year, while other data require continuous interviewing throughout the year, but not necessarily every year. A third type of data needs a certain period of time, etc. Seven modules 3, for which non-overlapping samples are used, were implemented in 1997. As the mean interview length for the joint questionnaire is fixed at 15 minutes, module 3 may generally take another 30 minutes.

In both the joint questionnaire and the modules the screening option is used to follow up target groups. These follow-ups are the fourth module in POLS and can also be used to incorporate a panel survey. In 1997 the fourth module consisted of a 15-minute CATI interview on accidents and injuries. The contents and the combination of techniques to collect the data of the various modules are described in more detail below.

Module 3: Health and working conditions

This module uses two data collection methods: CAPI and PAPI. The paper and pencil method is primarily chosen because of the more personal questions on health, for example about 'burnt out' syndrome.
This is a continuos module (January-December) and describes the complete population, using proxy respondents (parents) for the sampled persons between 0 and 11 years old. Some of the questions are asked at the request of the Ministry of Social Affairs in aid of monitor information on working conditions. In addition, the combination of health indicators with the fairly extensive question set on working conditions creates a new data source, which may meet important user demands.

Module 3: Justice and Environment

This module also uses two data collection methods: CAPI and PAPI. The PAPI is chosen because of the questions on time-use: a so-called 'yesterday interview', consisting of table-like questions which cannot easily be 'translated' to a CAPI method. This module is also continuous, every year and concerns the population aged 12 and older.

Module 3: Justice and Participation

This module uses only CAPI. Some of the questions (on justice) are identical to the former module 3 to fulfil sample-size requirements about themes that have to do with victims of crime. In other words: the idea of a joint questionnaire is applied a second time in POLS to a combination of two (out of seven) subsamples.
In so doing the precision of the population estimators on crime victims improves substantially.

Module 3: Youngsters

This module - presented to respondents aged between 12 and 30 youth and covering attitudes with regard to education, parents, sexuality, etc. - is also a continuous one. It uses only CAPI, but in a slightly different way than described above. Because of the age of the sampled persons it is possible to let respondents complete parts of the questionnaire themselves on a laptop computer. We expect the answers to these blocks of questions (on sex, drugs and crime) to be of a better quality with this method. The module is scheduled three times per ten years and was asked from March 1997 to December 1997.

Module 3: Well-being

This module only uses CAPI and consists mainly of questions asked in Statistics Netherlands surveys since 1974. This module describes the population aged 18 and over and was developed in cooperation with the Social and Cultural Planning Office (SCP). Using the data of this module in combination with the joint questionnaire, the SCP will be able to produce the so-called Well-being Index to sum up the living conditions of the Dutch population in one figure. The module is scheduled for three times per ten years and was implemented from March 1997 to December 1997.

Module 3: Accidents and injuries (and time use)

This module uses both CAPI and PAPI. The PAPI part, covering time use (population aged 12 and over), is identical to that in the 'Justice and environment' module and is only used for the interviews held on Fridays. By doing this we can meet the restriction on the equal distribution of time use data throughout the week. Again this is an example of combining subsamples within an integrated survey. We foresee that such a use of identical question blocks in different subsamples, making optimal use of Blaise-routines (see the contribution by Faessen and Kerssemakers further on in this issue), will become a more common feature of integrated (systems of) surveys. However, the other side of this coin may not be forgotten: the weighting procedures become more complex (see Hofmans elsewhere in this issue).

The accidents and injuries module is part of a new research project at Statistics Netherlands, set up at the request of the Ministry of Health. The remainder of the data requirements for this project are met in the joint questionnaire (screening questions) and a follow up study in a fourth module (see below).

Module 3: Housing

This module uses CAPI, PAPI and CATI and marks several new developments within Statistics Netherlands. In cooperation with the Ministry of Housing and Environment, the four-yearly Housing Demand Survey (HDS) has been changed into a continuous survey (see Everaers and Winkels, 1997). As can be seen in Figure 1, this new style HDS (population aged 18 and over) is referred to as the Housing module. The small sample size in 1997 has to do with the fact that this project only started in November 1997. From 1998 onwards the Housing module aims at a net sample size of 15,000 people, three times as large as the other modules together, because of the implementation of over-sampling features. The following section examines this oversampling in more detail; the process brings together several new developments:

a) the implementation of a POLS module 3 by CATI, consisting of a summary (core indicators only) of another module 3 on housing by CAPI.
b) the full implementation of the data collection from a subsample by a market research institute on behalf of Statistics Netherlands.

Module 4: a follow-up study on accidents

Based on screening questions in the joint questionnaire, a telephone survey on accidents and injuries was conducted in 1997. The interviews lasted about 15 minutes and were held within three weeks after completion of module 3.

5. The reduction of the response burden: over-sampling features in POLS

POLS makes it relatively easy to oversample one or more modules. This can reduce macro response burden, as it makes a separate survey superfluous. Oversampling was introduced on a small scale in POLS in 1997. In some municipalities there was an oversampling (net 300 persons) for the health and working conditions module (including the joint questionnaire) in relatively poor parts of the city. This extra sample was based on previous research that linked welfare indicators to clusters of addresses.
Also in 1997 preparations were made for a much larger oversampling: in 1998 some 65,000 extra people will be interviewed about their housing situation. This oversample is part CAPI (the largest part) and part CATI.

a) as explained above the Housing Demand Survey (HDS) has been changed into a continuous housing survey. From 1998 onwards the HDS will be held on a yearly basis, so a wide variety of developments can be monitored yearly at national and provincial level. To enable comparisons with the HDS 1993/1994 and former studies, the total sample in 1997/1998 will be extended to 60,000 persons. This number of interviews guarantees that the same level of regional detail will be obtained in POLS compared with the series HDS 1977-1994. The oversampling of 45,000 net interviews will be carried out by a market research institute, as the number of extra interviews is too large for the fieldwork department of Statistics Netherlands. However, the data processing however will be done by Statistics Netherlands and will be fully integrated in the procedures developed for the joint questionnaire and the housing module.
b) Both the Ministry of Housing and Environment and Statistics Netherlands offer the possibility to municipalities and/or provinces to add new records to the POLS housing subsample, in order to realise more regional detail. An area that previously would have remained unseen within the HDS data set will then become visible. The lower limit used for oversampling is an area with approximately 10,000 inhabitants. There are two types of oversampling: using the complete questionnaire or using a CATI questionnaire with the key questions recommended by the Department of Housing and the Association of Dutch Municipalities. All these extra interviews include the joint questionnaire, which not only benefits the users of housing data. Because of the contents of module 2 (see above), more information on socio-cultural phenomena becomes available at a regional level, but in a standard form and by definition linked to the national figures.

6. Some future developments of the POLS design

It should be stated that the design of POLS is not yet definite. Several pilots projects remain with respect to:

- data collection techniques (including a PAPI joint questionnaire, an exploration of data collection via the World Wide Web, an evaluation of the improvement of response figures by using CATI, etc.);
- the structure of the questionnaire (for example using new features of Blaise to read open-ended questions efficiently);
- the potential reduction of sample sizes due to the results of consistent weighting;
- the inclusion of panels and other surveys or parts of surveys of Statistics Netherlands which have not yet been integrated;
- the ways of monitoring both response and interviewers.

One future development will have serious consequences for both the questionnaire and data processing: more and more data from computer registrations can be linked by Statistics Netherlands to produce statistics. This means that parts of the POLS questionnaire can then be skipped (for example the questions on health insurance, housing characteristics, etc.), and that the data processing becomes more register-driven (instead of survey-driven). Although obviously not all the problems involved have been foreseen, the implementation of POLS in 1997 has tried to anticipate these new developments.

References

Everaers, P.C.J. and Winkels, J.W., 1997, The new housing demand survey, In: *Netherlands Journal of Housing and the Built Environment,* Vol. 12, no 3, pp. 325-333

Innovative weighting in POLS. Making use of core questions

Marion G. Hofmans

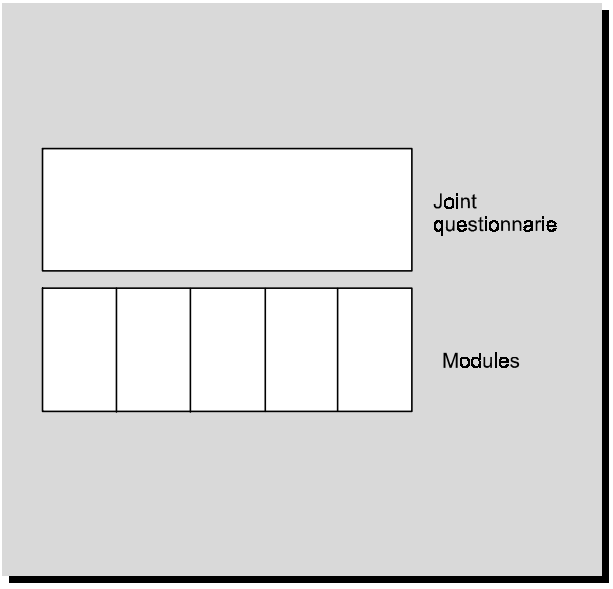
Summary

POLS is a continuous survey of various aspects of living conditions. It integrates several individual and household sample surveys covering subjects like health, justice and time use. All surveys are linked by a joint questionnaire containing general questions about demographic and socio-economic features. The joint questionnaire also contains core questions concerning the topics of the integrated surveys. These core questions play an important part in the estimating method for population totals of POLS target variables. They can achieve consistency between the integrated surveys. In some cases variances of estimated totals may decrease substantially because of the core questions and, conversely, the sample sizes of the integrated surveys may be reduced.

1. Introduction

POLS, the Dutch integrated system of social surveys is a continuous survey among the Dutch population on various aspects of living conditions. The project integrates several individual and household sample surveys, and covers aspects like health, working conditions, justice, time use and accidents.
The POLS questionnaire consists of a joint questionnaire and several subsequent modules. The joint questionnaire acts as a common frame for the modules which each represent one of the surveys to be integrated. The joint questionnaire contains harmonised questions for demographic and socio-economic classifications. It also contains core questions intended to provide fundamental information on social indicators. These social indicators concern the particular topics of the modules. Every person in the POLS sample (joint or basic sample) will be asked the questions of the joint questionnaire. The basic sample is divided into several module samples and the respondents in these samples will be asked the questions of the module they are assigned to. Figure 1 gives a schematic overview of the design of POLS. More details are given in the contribution by Winkels and Everaers in this issue.

Figure 1. Schematic overview of POLS design

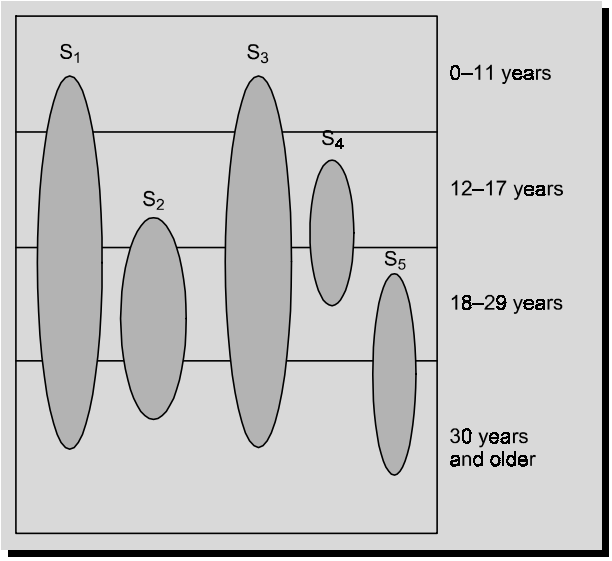


The core questions in the joint questionnaire allow a detailed publication of important social indicators, because all persons of the basic sample are asked these questions. As these core variables can also be crossed with variables from the modules, the modules should be consistent with each other on the main social indicators and possibly with respect to other information in the joint questionnaire. To achieve this consistency the technique of consistent weighting is used to estimate population totals of the survey results. More specifically, first the basic sample is weighted, and this subsequently serves as a weighting frame for the samples of the modules. Using core variables and other common variables from the joint questionnaire as auxiliary information in the weighting process yields consistency between the modules. An additional advantage is that in some cases this weighting method considerably reduces the variance of an estimated population parameter. The method may also serve to reduce biases caused by non-response.
This contribution discusses the weighting method to estimate population totals of POLS target variables, and shows some methodological advantages of the method. In section 2 first the sample design of POLS will be described. In section 3 the weighting method will be given, emphasising the aspect of consistency when using core variables and other common variables. Section 4 will go into the effects of the integration of surveys with common variables. It will be shown that, owing to the use of core variables in the weighting method, in some cases a substantial reduction of variance can be reached and, conversely, sample sizes could be reduced. Lastly, section 5 contains the conclusions of this article.

2. The sample design of POLS

POLS is a continuous survey. Every month a sample is composed by drawing a self-weighted sample for each module and subsequently joining these samples to get the basic sample. The module samples are person-based and drawn from the Dutch population register.
All modules in POLS have different target populations. The module on health and working conditions, for instance, concerns the whole population. The module on justice, participation and environment, on the other hand, concerns only people aged twelve and older. Figure 2 gives a schematic overview of the sample design of POLS in 1997. For most of 1997 the design roughly consisted of five modules. The first module was on health and working conditions (sample S_1), the second on justice, participation and environment (S_2), the third concerned accidents (S_3), the fourth living conditions of young people (S_4) and, lastly, the fifth covered well-being (S_5). The second module (S_2) had one target group, but contained two partly overlapping questionnaires.
The target populations of the modules were respectively all persons from age zero, persons aged twelve and over, all persons from age zero, persons aged of twelve to thirty years and persons aged eighteen and over. People living in institutions or shelters are excluded from the POLS samples.

Figure 2. Schematic overview sample design POLS 1997



The differences between the target populations have implications for the first order inclusion probabilities. All first order inclusion probabilities are equal for the individual modules because of the self-weighting character. However, the first order inclusion probability of a person with respect to the basic sample depends on the number of target populations of the modules the person belongs to. The more target populations the person belongs to, the higher the inclusion probability will be (Hofmans, 1997). In POLS (1997) eighteen to thirty year-olds had the largest inclusion probability, assuming all sampling fractions of the module samples to be nearly the same. The under-twelve's had the smallest inclusion probability.

3. Weighting method of POLS

3.1 Weighting method for the basic sample

The well-known π-estimator (see e.g. Särndal et al., 1992) can be used. to estimate population totals of target variables of the joint questionnaire. The weight of a person in the basic sample is then equal to the inverse of his or her inclusion probability. This means that every person in the sample represents as many persons as this inclusion weight indicates. Compared with the π-estimator a more precise estimator of the total of a target variable can be achieved by using auxiliary variables to estimate the total. These auxiliary variables should then correlate to the target variable, and the population totals of the auxiliary variables must be known or precisely estimated from external sources. Auxiliary information can be used in the estimating procedure of a target variable by means of the general regression estimator (Särndal et al., 1992). Applying the regression estimator yields a weight for every person in the sample. Every regression weight is equal to the inclusion weight multiplied by a correction weight. Therefore the regression estimator can be seen as an adjustment of the π-estimator.
Applying the general regression estimator with auxiliary variables which correlate highly to target variable may decrease the variance of the estimated total of this target variable. Also biases caused by non-response may be reduced. Furthermore, consistency is achieved with the known population totals of the auxiliary variables, i.e. estimating population totals of the auxiliary variables from the sample yields the known population totals of these variables.

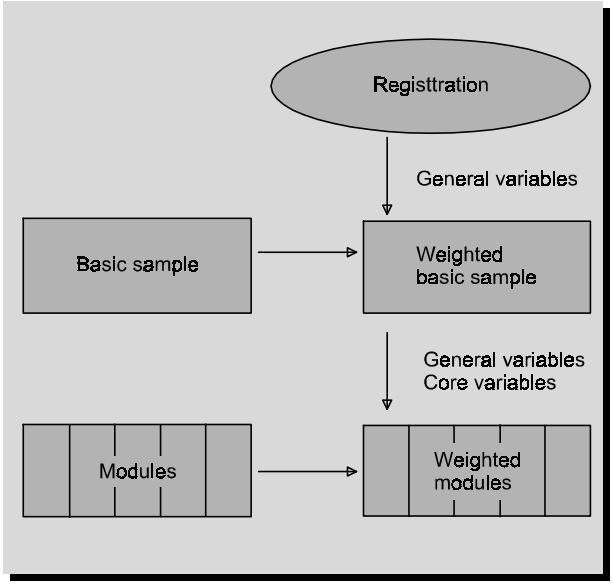
The method of incomplete multiple (linear) weighting (see Bethlehem & Keller, 1987) is applied to the basic sample of POLS. This method is a special case of the general regression estimator. Only categorical variables are used to weight the basic sample. To be more specific, demographic features like sex, age and marital status as well as regional variables like province and degree of urbanisation are used. The population totals of these variables are obtained from the Dutch population register. In the future it may be possible to combine other registrations with the population register so that more auxiliary information will be available to weight the basic sample. This may improve the estimating method of the basic sample and subsequently that of the module samples.

3.2 Weighting method for the module samples

Compared with the π-estimator the variance of the regression estimator often decreases by only a few per cent when using auxiliary variables like demographic features and regional variables. Using the weighted basic sample as auxiliary information for the weighting of a module sample may reduce the variance of estimated totals considerably more. After all, in addition to the mentioned variables core variables can also be used to weight the module sample. Although the population totals of the core variables are initially unknown, they can be estimated after weighting the basic sample. The core variables tend to have a relatively high correlation with the target variables of the modules because they are focused on the particular topics of the modules. The higher this correlation is, the less the variance of the estimated total of the target variable will be.

Using core variables to weight the module samples can be performed by applying the adjusted regression estimator (see Renssen & Nieuwenbroek, 1997). This kind of weighting is very similar to applying the general regression estimator to two-stage sample surveys (see e.g. Särndal et al., 1992). Apart from the variance reduction, the adjusted regression estimator also brings about consistency between the various modules as well as consistency between the modules and registrations (external sources) with respect to the common variables used in the weighting method. These variables concern general variables like sex and age, which are used for weighting the basic sample, as well as core variables. To avoid too much fluctuation in the weights of a module sample, because of insufficient observations, the

Figure 3. Weighting procedure of POLS



number of categories of the auxiliary variables for weighting the module samples should be limited.

Figure 3 shows the weighting procedure of POLS. Summing up, first the basic sample is weighted using auxiliary variables of which the population totals are known from the Dutch population register (general variables). Secondly, every module sample is weighted using the weighted basic sample to obtain the population information. Together with general variables, other auxiliary variables - for which the population totals are estimated from the basic sample - are used to weight the module samples (core variables). These core variables relate to the topic of the module. The method of incomplete multiple weighting is applied to weight the basic sample and the module samples.

4. The effects of integration

The weighting method of POLS may be particularly effective for the modules. Variances of some estimated totals may be substantially reduced when core variables are used. Conversely this means that sample sizes may be reduced, provided that after this reduction the same weighting method can still be carried out and thus the variance reduction will be comparable with the reduction reached before. In general the reduction of variances (and conversely sample sizes) when weighting with core variables compared with weighting without core variables depends on two aspects. First, it depends on the extent of correlation of the used core variables to the target variable; and secondly, the variance of the estimated total of a target variable depends on the number of module samples that are combined to estimate the total (Hofmans, 1997). This number depends on the target population of the variable. If two or more module samples are used to estimate the total there will probably be a higher variance reduction than if only one module sample is used. After all, if a combination of module samples is used, more auxiliary information (i.e. information on more respondents) is used to estimate the total. For example, target variables on living conditions of eighteen to thirty year-olds will relatively have the smallest variance because the use of auxiliary information here is the most extensive.

4.1 Reduction of variance and sample size: an example

Let us illustrate the effect of integration by means of an example, based on the POLS questionnaire of the second quarter of 1997. Only three modules are used in the example, i.e. those on health and working conditions; justice, participation and environment; and well-being. The basic sample is the combination of these module samples. Taking into account the target variables to be considered, only persons aged eighteen and over are part of the example. The basic sample then consists of approximately four thousand persons. The sample of the module on health and working conditions (S_1) consists of approximately two thousand persons, that on justice, participation and environment (S_2) of approximately a thousand persons and the sample of the module on well-being (S_5) also of approximately a thousand persons.
From the health and working conditions module, the target variable ‘smokes a lot’ is considered. This dummy variable represents whether a respondent smokes ten or more cigarettes a day. From the module on justice, participation and environment the target variable ‘member of a public library’ is considered, and from the module on living conditions the target variable ‘goes to the opera’ is considered.
To estimate populations totals of the target variables first the basic sample is weighted. Therefore in this example the model

sex × age

is used (for notation, see Bethlehem & Keller, 1987). In fact the basic sample is weighted by post-stratification using ten post-strata; age is divided into five categories and sex into two. Subsequently, the module samples are weighted. This is done in two ways in order to show the effect of integration when weighting with or without core variables. First, only sex and age are used to weight the module samples. Just as in the basic sample, this is done by post-stratification using ten post-strata. Secondly, in addition to sex and age core variables are used to weight the module samples (see figure 3). For the module on health and working conditions the core variable ‘smoking’ is used. This variable is divided into three categories (‘smokes’, ‘smokes sometimes’ and ‘does not smoke’). The model

(sex × age) + smoking,

a model for incomplete multiple weighting, is carried out to weight sample S_1. For the module on justice, participation and environment the model

(sex × age) + visits museums

is carried out. The core variable ‘visits museums’ indicates whether a respondent occasionally visits a museum. This model is also carried out for the module on well-being.
Table 1 shows the reduction of variance when using core variables (variance after integration) compared with not using core variables in the weighting model (variance before integration). For the variable ‘smokes a lot’ there is a considerable reduction of variance and for the variable ‘goes to the opera’ too, a reduction of variance is achieved. Considerable reductions in variances of the estimated totals of the target variables may be achieved particularly when target variables are a refinement of the auxiliary variables used for weighting. For the variable ‘member of a public library’ there is almost no variance reduction. Apparently, there is little or no correlation between this variable and the auxiliary variables. Nevertheless, the auxiliary (core) variables will always be used for weighting the sample S_5 for reasons of consistency.

Table 1
Reduction of variance (in %) as a result of integration

	Before integration	After integration
Smokes a lot (S_1)	100	71
Member of a public library (S_2)	100	97
Goes to the opera (S_5)	100	89

Roughly speaking, reduction of variance is inversely proportional to reduction of the sample size. For integrated sample surveys (modules), linked by a joint questionnaire, reduction of one sample size has an indirect effect on the other sample sizes. After all, the size of the basic sample decreases and subsequently, leaving aside whether the original weighting models can still be used, this is at the expense of the precision of estimated totals of target variables of the other modules. If target variables of the modules are to be estimated with a certain degree of precision, the sizes of the module samples will therefore have to be determined simultaneously (Hofmans, 1997). Table 2 shows the results of simultaneously determined sample sizes based on the precision of the target variables reached before integration. For the sake of comparability, the sample sizes corresponding to the variance before integration are set at one hundred. For sample S_1

the reduction of sample size differs from the reduction of variance. This is caused by the proportional change between the individual sample sizes. Note that the size of the basic sample is reduced by approximately twenty percent.

Table 2
Reduction of sample size as a result of integration (% in brackets)

	Before integration		After integration	
Smokes a lot (S_1)	1,980	(100)	1,318	(67)
Member of a public library (S_2)	948	(100)	920	(97)
Goes to the opera (S_5)	1,021	(100)	912	(89)

5. Conclusions

This article has discussed some methodological advantages of the integration of several individual and household surveys in POLS. The integration in POLS consists of a linkage of the surveys by a joint questionnaire. Through this linkage the surveys (modules) have several variables in common: general variables like demographic and regional features, but also core variables, providing fundamental information on the particular topics of the modules. The core variables play an important part in the weighting method of POLS.

POLS applies the method of consistent weighting, using the adjusted general regression estimator. This method is very similar to the application of the general regression estimator in two-stage sample surveys. The weighting method yields consistency between the modules and the joint questionnaire with respect to the core variables and other common variables used in the weighting method. An additional advantage of this kind of weighting is that in some cases a reduction of variance of an estimated population parameter may be obtained. To achieve this, the variables used in the weighting method should correlate to the target variable concerning the population parameter. This reduction in variance can be considerable in cases where the population parameter is a refinement of a (core) variable used in the weighting method.

In roughly the opposite direction from the reduction of variance, there is the reduction of sample size. As a result of integration, sample sizes of the modules and hence of the joint questionnaire may be reduced. One condition for the reduction of sample sizes is that it should still be possible to carry out the original weighting method in order to obtain the same variance reduction as before.
Although up to now the main reasons for carrying out the particular weighting method of POLS were to achieve consistency and reduce variance, the method may also serve to reduce biases caused by non-response. First, in the future it may be possible to join other registrations with the Dutch population register from which the sample was drawn. This will make more information available on non-respondents, which may help to reduce biases caused by non-response. Secondly, in any case information from the joint questionnaire is available for people who did respond to the joint questionnaire, but not to the module they were assigned to. This information can be used for further improvement of the weighting method of the module samples. Further research on non-response in POLS will be carried out in the near future.

References

Bethlehem, J.G. and W.J. Keller, 1987, Linear weighting of sample survey data, In: *Journal of Official Statistics,* Vol. 3, No. 2, pp. 141-153 (Statistics Sweden)

Hofmans, M.G., 1997, *Reduction of variance and sample size by integration of sample surveys* (Heerlen: Statistics Netherlands, external report, BPA. No. 4982-97-RSM) (in Dutch)

Renssen, R.H. and N.J. Nieuwenbroek, 1997, Aligning estimates for common variables in two or more sample surveys, In: *Journal of the American Statistical Association* 92 (437), pp. 368-374

Särndal, C.E., B. Swensson and J. Wretman, 1992, *Model Assisted Survey Sampling,* (New York: Springer-Verlag)

Reducing non-response: the POLS fieldwork design

Rob Vousten and Wim de Heer

Summary

The start of a new integrated survey on living conditions (POLS) was accompanied by the introduction of a new fieldwork design. Elements of this new design are refusal conversion, the use of mixed modes and a personal approach of intended respondents, using a person based sample frame. The new fieldwork design lifts the response rate from 55% to 63%. A relatively small part of the increased response can be contributed to the use of mixed modes, the largest part must be contributed to the personal approach and to the training and monitoring of the interviewer corps.

1. Response trends

Non-response is a serious problem in social statistics and social surveys. It makes survey estimates questionable, because of a potential and hardly measurable bias.
It is a largely unknown factor and it is unclear to what extent post-survey adjustment can deal with non-response bias. In recent years there has been a growing concern about this problem. In 1990 an International Working group on Household Survey Non-response was founded as a setting to exchange experiences, expertise and findings.
The working group initiated an International Survey on Non-response to compare results in different countries, taking into account differences with respect to sample and survey design, fieldwork strategy and fieldwork organisation (De Heer and Israels, 1992; Maas and De Heer, 1995; De Heer, 1996). From these results it appeared that response rates for the Netherlands are very low compared with other countries.
For instance, the Dutch Labour Force Survey achieved a response rate of 60% in 1995, while the response rate for Labour Force Surveys in other European countries varied between 82% and 93% in the same year (De Heer, 1996).
The main question is whether a low response rate implies unreliable results.
As the non-response error is a function of the non-response rate and the difference between average scores among non-respondents and respondents, increasing response rates do not always reduce the non-response error (Groves, 1989). Improving response rates alone might not be enough. Reducing the non-response bias must be the ultimate goal. Nevertheless, when a new survey is designed it is useful to aim for an optimal response result.

The main, widely used non-response categories are refusals and non-contacts, which account for some 80% of non-response. Other categories are *language barrier*, *person unable to respond* (handicapped) and *person untraceable* (in hospital, abroad etc.). Refusal rates in the Netherlands are extremely high: 26% compared with 2% to 13% in other countries (De Heer, 1996). Efforts aimed at improving response will therefore require special attention for the treatment of refusals. Response can be affected by a number of factors that can be manipulated to some extent: the sample and survey design, the fieldwork strategy and characteristics of the fieldwork organisation. One factor that cannot be easily influenced is the survey climate. This refers to the 'response burden', which is increasing, largely owing to the growth of call centres and tele-marketing. In addition, the public debate on privacy issues will also influence the survey climate. Most of these aspects were taken into account in the development of the new POLS fieldwork strategy. The new design should focus on optimising the conditions to get the best possible response results.

2. Fieldwork strategy

In order to design and test fieldwork strategies it was necessary to formulate expectations with respect to the behaviour of intended respondents. According to Groves and Cialdini (1991), the decision to participate in a survey is not the result of a rational-analytic decision process, but rather of a 'heuristic' process. The decision is influenced or based on circumstantial grounds that might be related to the moment of the first contact. This means that the interviewer-respondent interaction during the first contact is essential. Interviewers who are able to use all the signals transmitted by an intended respondent, seem to be the most successful. Some intended respondents may refuse because of a 'no time' reason, but if approached again on a different day at a different time, they might comply with the request. By contacting initial refusals at a different time and under different circumstances, at least a part of this group will be persuaded to participate. In most countries where this kind of 'refusal conversion' is used for survey fieldwork, it pays off.
A fieldwork strategy for POLS was developed on the basis of these principles. A combination of data collection methods ('modes') seemed plausible as not all population groups can be reached easily using only one mode. During the first tests, attention was focused on how to combine different modes of data collection and how to establish the best conditions for the interviewer-respondent interaction. The modes were Computer Assisted Personal Interviewing (CAPI) and Computer Assisted Telephone Interviewing (CATI). In a quantitative test, two combinations were examined: CAPI followed by CATI, and CATI followed by CAPI.[1)] One problem that arose was that the POLS questionnaire was too long for CATI only; only part of it (the joint questionnaire) could be used by telephone. The main objective of the test was to determine the most successful of two mixed mode procedures: start with CATI and ask the respondent to participate in a CAPI follow-up, or start with CAPI and then try to convert refusals and reach non-contacts using CATI. The test revealed that the CAPI/CATI combination generates the highest response rate: 67% versus 42% for the CATI/CAPI combination. Another lesson from the test seemed to be that a combination of different modes can lead to extra losses, because changing the mode also gives the initial respondent a new opportunity to refuse.

3. Fieldwork implementation

At Statistics Netherlands, the Household Surveys department is responsible for the data collection for all person and household surveys. The implementation of POLS in the operation of the fieldwork department required some adjustment of existing processes and procedures and the development of new processes and procedures. New processes were developed and implemented for:

- sampling procedure
- interviewer instruction
- mixed mode
- fieldwork monitoring
- response monitoring

3.1. Sampling procedure

POLS is the first large and continuous survey of Statistics Netherlands in which a person based sample frame is used in combination with computer assisted interviewing. The sample frame is based on the relatively new national database of all registered inhabitants of all Dutch municipalities, which contains approximately 15 million persons. The sample frame used for one year of POLS samples consists of 10% of the full base, in such a way that a particular person will be included in the frame once every 10 years. From this frame the actual POLS sample is drawn in 12 monthly portions. While POLS uses a person based sample frame, samples for other large surveys of Statistics Netherlands (such as the Labour Force Survey) are drawn from a register of addresses. To reduce the respondent burden, Statistics Netherlands does not approach the same sample unit more than once in two years.

Therefore, the person and address samples are screened to eliminate doubles: a person living at an address already in the address sample is removed from the POLS sample.

As POLS is aimed at the non-institutionalised population of the Netherlands, persons living at addresses of registered institutions (as recorded in an 'institution base') are also removed from the POLS sample.

The last step in the preparation of the sample is the addition of telephone numbers, to be used for a possible re-approach by means of CATI. Telephone numbers of only 75% of the sample can be found in the telephone register; the other 25% have an unlisted number or no telephone at all.

3.2. Interviewer instruction

The new POLS questionnaire contained many novelties for the interviewers: the use of Blaise 3 and a new laptop management system were the most important. [2)] Also, some fieldwork related aims of the POLS design (such as the reduction of non-response) had to be explained to the interviewers. During a period of six months, the complete interviewer corps of Statistics Netherlands (approximately 500) was trained for interviewing according to the POLS design. In groups of fifteen, the interviewers were given a two-day instruction course. New interviewer manuals were written to provide the necessary information, not only for Blaise 3 and the new laptop system, but also for the joint questionnaire and for each of the particular modules incorporated in POLS.

3.3. Mixed mode implementation

As pointed out above, the mixed mode design for POLS uses a mix of CAPI and CATI. Respondents are first approached by a field interviewer for a personal interview (CAPI). In the case of non-response the CAPI interviewer will classify the non-response according to a set of categories. Non-contacts and refusals are then re-approached by means of CATI. The CATI interviewer will try to administer the joint questionnaire and will ask the respondent's permission for a visit by a field interviewer. The field interviewer will then complete the interview using CAPI. The reason for the division between joint questionnaire and module has to do with the interview length.

The total length of a POLS interview is approximately 45 minutes (15 minutes for the joint questionnaire, 30 minutes for the particular module). This is considered too long for a complete telephone interview, as it would increase the risk of interruption. The POLS fieldwork thus consists of three cycles: CAPI-1, CATI and CAPI-2. The duration of the three cycles is two months: one month for CAPI-1, one week for CATI, and the remaining part of the second month for CAPI-2. Every month part of the POLS sample is distributed among the interviewers, resulting in two parallel processes for different months: while CAPI-1 is being carried out for a certain month, CATI and CAPI-2 for the previous month are still running. The implementation of POLS in the fieldwork organisation was assigned to a project team within the POLS programme. Apart from setting up the fieldwork, this team was responsible for instructing and committing the organisational units in which POLS was introduced, gradually replacing the 'old' surveys.

3.4 Fieldwork monitoring

The development of a new CAPI management system for laptops made it possible to add some new features to obtain more detailed information on the fieldwork activities of each interviewer. Interviewers were instructed to perform data communication with the office at least twice a week. At any given time, the following aspects can thus be monitored for each interviewer:

- datacom activities
- the number of visits made to a particular address
- the result of each visit (non-contact, appointment, refusal etc.)
- the date and time of each visit
- the 'status' of each sample unit (response/non-response)

The information is used to generate listings of interviewers who need to be reminded of a latent or manifest backlog in their work. A first 'paper' reminder is sent out halfway through every month and is followed by one or more reminders by telephone. This procedure is assumed to have a positive impact on the overall response, although the exact contribution is hard to determine. At the moment the fieldwork information is not used to a full extent; some of the collected information remains unused. The reminder procedures could be intensified even further, if more of the available information were to be used.

3.5 Response monitoring

The increased complexity of the POLS fieldwork organisation made it necessary to develop a new monitoring system to determine the relative contribution of each of the three cycles to the overall response.

To this end some dedicated software had to be developed. The response monitoring should be able to give an accurate indication of the number of sample units to be distributed in the field, and the response and non-response rates. A number of non-response categories are used within POLS. For some of the categories such as language barrier, person unable to respond (handicapped) and person untraceable, a re-approach would serve no purpose. The non-response categories which are re-approached can be reduced to two categories: refusals and non-contacts.

The response monitoring system should be able to give an insight into the contribution to the overall response, but also into which initial non-response types are most likely to respond in a re-approach. Although the POLS fieldwork started in January 1997, it took until the end of that year before an integrated monitoring system was available. Until then, monitoring information had to be collected from various improvised systems, preventing a complete overall view of the process.

4. Some fieldwork results

4.1 Response

Response rates for POLS are considerably higher than they were for the former component surveys of POLS. In 1996 the 'old' surveys obtained an average response level of 55%. The overall response rate for POLS over the first three quarters of 1997 was 63%.[3] Within POLS, the health module seems to generate a response rate above this average. There are indications that the willingness to participate has some relation to the content of the survey. Although the introduction letter is the same for all POLS modules, it is thought that the interviewer (who knows which module has to be administered to a certain respondent) may influence the response rate by giving information on the content. This might work in favour of the health survey.
The response rates reported below refer to the complete POLS survey over the first three quarters of 1997.

The first approach in the field yields a response rate of 61%. About 56% of the non-response is eligible for re-approach by means of CATI[4]; the response rate here is 29%. Half of these respondents (52%) give permission for a further re-approach by means of CAPI to complete the interview. However, only 59% of those giving permission actually respond to the field interviewer. All in all, the mixed mode design lifts the overall response rate from 61% to 63%.
This means that only 9% of the non-response wave approached by CATI leads to a complete response. A positive side effect of the CATI cycle is an additional number of interviews in which the joint questionnaire is completed. This brings the response rate for the joint questionnaire up to 67%.
Compared with the 1996 results, the introduction of POLS in 1997 lifts the response rate from 55% to 63%. About two percent points of the additional response can be attributed to the mixed mode design; the largest part must be attributed to the use of a person based sample frame and to the training and monitoring of the interviewer corps.

4.2. Additional response from the mixed mode design

The additional two percent points resulting from the mixed mode design come from respondents who were initially assigned 'refusal' or 'non-contact'. Refusals outnumber non-contacts by three to one. Table 2 shows the CATI and CAPI-2 response rates for refusals and non-contacts.

Table 2 shows that 25% of the initial refusals, and 41% of the initial non-contacts respond in the CATI cycle. Of the initial refusals who respond in the CATI cycle, only 43% give permission to be approached for the CAPI-2 cycle, compared with 69% of the initial non-contacts. Apparently, CATI converted refusals are still relatively reluctant to let a field interviewer into their home. Once the field interviewer shows up, the response rate is nearly equal for both groups (around 59%).
As pointed out earlier, this implies that only 9% of the initial non-response is converted to respond in the CAPI-2 cycle. Further inspection shows that the additional response predominantly comes from the 'non-contacts'. Only 6% of the initial refusals are persuaded to participate and convert into complete responses; the conversion rate for the non-contacts is 17%. However, within the additional response, there are slightly more refusals than non-contacts (53% versus 47%).
The most important conclusion that can be derived from tables 1 and 2 is not so much that it is possible to convert refusals or to look hard to find respondents, but that splitting the interview process leads to a relatively great loss of potential responses: less than one third of the CATI responses lead to a complete CAPI-2 response.

Table 2
Response of initial refusals and non-contacts in the CATI/CAPI-2 cycles

	refusals	non-contacts	
CATI			
total	6,062	1,982	8,044
response	1,486	821	2,307
CAPI-2			
total	636	563	1,199
response	370	332	702

4.3. Using a person based sample frame in the field

The use of a person based sample frame may cause some difficulties in reaching all the persons in the sample. The address at which a person is going to be contacted is taken from the municipal population registers. Some groups will be hard to trace: the 'highly mobiles', the 'ghosts' and the 'unregistered'.
A small part of the sample (the highly mobiles) may have moved to a different address by the time the field interviewer makes her first approach. Also, the population register may not be completely up-to-date. The ghosts consist of the very small part of the population who do not actually reside at the registered address. The unregistered are unknown to the population register and will be bypassed in a person based sample frame.
It appears that 1.9% of the persons in the POLS sample cannot

Table 1
Response results for the three POLS cycles (Jan–Sep 1997) in % of field sample

cycle 1 (joint quest. + module)	field sample	cycle 2 (joint quest.)	inflow CATI	cycle 3 (module)	inflow CAPI-2
first approach	100 (n=36,957)				
response CAPI-1	61				
non-response CAPI-1	39	re-approach	21.7		
		response CATI	6.3	re-approach	3.2
				response CAPI-2	1.9

be found by the field interviewers at the registered address. During 1997 no special effort was made to trace these sample units. The highly mobiles, however, may be important targets for some surveys, such as a housing survey. The start of a housing survey in 1998 requires the introduction of a procedure for tracing the highly mobiles. The field interviewer will try to determine the new address, or at least the new place of residence, so the sample person can be relocated by a new query to the population register. In addition to some difficulties, the use of a person based sample frame may also have some positive response effects, although the extent to which this is the case will be hard to determine. One effect may be an increased willingness on the part of the respondent to participate in the survey, instigated by a personally addressed introduction letter. An indication can be found in the results of a small survey that was carried out among the interviewer corps at the end of 1997.
According to three quarters of the 443 responding interviewers, the fact that the introduction letter is addressed to a person instead of only an address facilitates the first approach. The same proportion of the corps claims to use the additional information from a person based frame (such as age) to determine the best strategy for the first approach (such as the time of visit). Other effects in the field are related to the fact that more effort has to be put into contacting the right person from the sample. As no proxy interviews are allowed, this results in an increase in the number of visits to an address and an increase in the number of appointments.

5. Discussion

The current fieldwork design can be improved in several ways. One way is to increase the proportion of the non-response that is eligible for a CATI re-approach. To this end, a revision of the non-response categories will be necessary. For instance, an existing residual category (‘others’) gathers non-response that is now not, but possibly could be, re-approached.
The CATI cycle causes some difficulties for the telephone interviewers who have to deal with aggression from unwilling respondents, particularly from the initial refusals. A further training of the interviewers to handle these delicate situations might prove useful.
A relatively large proportion of the CATI response does not lead to a complete response in the CAPI-2 cycle. This may be caused by the division of the interview into two parts: a CATI part and a CAPI follow-up, several days later. An undisrupted interview may result in a higher response rate. If the interview length prevents the complete use of CATI, an option may be to perform the re-approach by means of CAPI, preferably by a different interviewer than the first time around. Another advantage of this method would be that people without a registered telephone number can also be re-approached. In addition, further improvements can be made to maximise the response at the first approach. Plans are being developed for extensive training of interviewers in ‘doorstep interaction’. Also, the need for a more dedicated fieldwork monitoring system is felt as an absolute necessity.
Of course there is a price to be paid for the additional response generated by the current mixed mode design: the cost of a ‘mixed-mode’record is approximately 15% higher than that of a normal record, which makes the extra, hard-won cases relatively expensive. Further research into non-response bias will be necessary to make a well-considered decision on whether the extra response is worth the extra costs.

References

De Heer, W. and A. Israels, 1992, Response trends in Europe. In: *ASA 1992 Proceedings of the Section on Survey Research Methods.* (Alexandria, VA.: American Statistical Association), pp. 92-101.

De Heer, W., 1996, *International response trends: description and explanation.* (Rome: Paper presented at the 7th Workshop on Household Survey Non-response, 2-4 October 1996, ISTAT).

Groves, R.M., 1989, *Survey errors and survey costs* (New York: Wiley)

Groves, R.M. and R.B. Cialdini, 1991, Towards a useful theory of survey participation. In: *ASA 1991 Proceedings of the Section on Survey Research Methods.* (Alexandria, VA.: American Statistical Association), pp. 88-97

Maas, C. and W. de Heer, 1995, Response developments and the fieldwork strategy. *Bulletin de Méthodologie Sociologique*, vol. 48, pp. 36-51.

Notes

1) See the contribution by Akkerboom, Dehue & Snijkers in this issue for more information on the qualitative and quantitative test procedures.
2) See the contribution by Faessen & Kerssemakers in this issue for more information on the data processing of POLS
3) As the response rate is generally lower in the last quarter of the year, the final response rate for 1997 may be slightly below this.
4) Only refusals and non-contacts for which a telephone number could be found are eligible for the second (CATI) cycle.

Qualitative tests for data collection (re-)design. The case of POLS

Hans C. Akkerboom, Francine M.J. Dehue and Ger J.M.E. Snijkers

Summary

Before the introduction of POLS, the interviewers were critical of the quality of the questionnaires, in particular with regard to their comprehensibility for both interviewers and respondents. The policy of Statistics Netherlands is to aim at more respondent-friendly questionnaires in order to improve data quality. Therefore, a Questionnaire Design Resource Centre (QDRC) was set up, to offer facilities and methods for designing and performing pretests as part of survey design and development. The POLS-questionnaire was designed in co-operation between the internal users of the survey outcomes and the QDRC. We present a general model of questionnaire development and the implementation of the model in the POLS-project. In particular we discuss the use of different forms of meta-interviewing for testing and improving survey questionnaires. We conclude that Computer Assisted Qualitative Interviewing is very useful in this respect.

1. Composing pretesting programmes according to a model of data collection development

The quality of raw data can be judged by criteria such as relevance, timeliness, completeness, validity and answer reliability. To meet these criteria satisfactorily, commitment of the interviewers, respondents and users of the survey outcomes is important. Since the early nineties, the policy of Statistics Netherlands has been to aim at more respondent-friendly questionnaires, in order to improve data quality. Therefore, a Questionnaire Design Resource Centre (QDRC) was set up, to offer facilities and methods for designing and performing pretest as part of survey design and development (Akkerboom, 1996). In these pretests, the focus is on collecting respondent-related 'meta-information' in addition to the data needed for the survey in question. Generally, such meta-information is information on what respondents think, do and feel while answering the survey questions, and how respondents fare while answering the questions and transmitting the answers: 'What conflicts of interpretation can arise? Under what circumstances is the respondent willing and able to provide data with reasonable effort and cost?'.

Respondent-related meta-information is collected in interviews, differing in the number of participating respondents, in the structure of the interview and in the type of information to be gathered. A focus group, for example, is a one-to-many (one interviewer, many respondents) unstructured, exploratory or evaluative interview with a topic list as agenda. Comparable with focus groups are one-to-one unstructured open interviews. A meta-interview is a one-to-one ordinary interview supplemented with meta-interviewing techniques like vignette, sorting and paraphrasing tasks, with interpretational, reference set, explanatory or expanding probes, or with check questions, confidence or sensitivity ratings, debriefing questions or reaction coding. Cognitive stimuli can also be included, explicitly related to a detailed cognitive model for the question-and-answer process, e.g. concurrent and retrospective thinking aloud prompts.

In performing any pretest, the QDRC uses standard protocols, that thoroughly describe which design issue should be addressed when, by which method and by which techniques. They also describe the organisational aspects of the pretest program, the training and supervision of the meta-interviewers and the communication with and presentation of the results to the client. The pretest of POLS was also designed according to the standard protocols. The ultimate aim of the pretest was to investigate the proposed prototype questionnaire and differing data collection procedures to improve the data collection before the survey officially began. The methods used in this pretest and the results of these methods are described below.

2. The POLS prototype joint questionnaire

A team of experts - in subject matter, questionnaire design and fieldwork - developed a proposal for a prototype of the joint questionnaire. Other experts were consulted for a review of the proposal. The proposed prototype of the joint questionnaire consists of a sequence of nine coherent question sets (see table 2). Each set is composed of a number of indicators, representing the relevant variables. These indicators in turn are expressed by one or a few successive items. Apart from main items, one may have (sub-)items for filtering, items that serve to check or elaborate one or more preceding questions, items that expand on a subset of preceding answer categories, items which request explicit consent or a proxy informant, etc. The items constitute the level on which problems are most easily detected. By contrast, the indicators constitute the level on which survey output quality is affected. Hence, when considering amendments to a problem detected on item level, the corresponding indicator has to be reviewed as a whole.

An example of a question set is the one on time use, consisting of the indicators 'main activity', 'number of hours worked', and 'time used for housekeeping'. The latter indicator consists of two items:

- main item: 'How many hours a week do you yourself usually spend on housekeeping activities?'
- sub-item (if answer=don't know): 'Would you know whether this amounts to:→
    1. 4 hours or less a week,
    2. less than 12 hours,
    3. less than 30 hours,
    4. or 30 hours or more a week?'

The question set on time use is crucial for the background variable 'employment status' (employed, unemployed, or inactive), and is also used as a routing criterion for the next question sets (on paid work and job search). Time used for housekeeping is meant to be a double-check on main activity.

All question sets and indicators in the joint questionnaire were derived from existing surveys. They are listed (in their final order) in the left-hand column of Table 2.

3. POLS Qualitative Content Test

The QDRC was asked to investigate the following issues in the qualitative content test:

1) general respondent commitment (attitude and interest concerning POLS);
2) the order of question sets in the joint questionnaire which respondents preferred most (subject to a few conditions of cohesion and routing;
3) respondents' opinion on transitions (between question sets, between joint questionnaire and particular module);
4) the comparability of CATI and CAPI with respect to the quality of the short and simple joint questionnaire items;
5) the average interviewing time, required to be 15 minutes or less.

Table 1
The qualitative test for the POLS joint questionnaire

	Steps	Topics	Methods & techniques	Test size
	Qualitative Content test			
1.	Focus groups (joint questionnaire)	Respondent commitment: appeal, burden, interest Order of question sets Transitions	Ordinary CAPI-interviews, followed by focus groups (2 QDRC focus group leaders)	v=2x6, i=5 (balance by age/sex)
2.	CAQI-tests (questionnaires in 4 variants of question-set order in joint questionnaire)	Order of question sets Expected problems with validity, precision of answers, respondent burden Unexpected problems Transitions	Expert re-appraisal CAQI 1-1 meta-interviews with focused probes: interpretational, reference set, explanatory or expanding probes, check questions etc.	e=1 v=16 (4x4) (balance by age/sex/ educational level)
3.	Joint questionnaire CATI-interviews followed by CAQI-tests	CATI and CAPI comparison for joint questionnaire Redesign of problematic items Preparation for Step 3: mixedmode protocol/ meta-interview	CATI-interviews at laboratory, followed by (a) focused probes about joint questionnaire, and (b) CAQI meta-interview of sm1	v=15
4.	Joint questionnaire CATI-interviews followed by CAQI-tests	CATI and CAPI comparison for joint questionnaire Second particular module Preparation for Step 3: mixed-mode protocol/ meta-interview	Expert reappraisal CATI-interviews at home, next day at laboratory	e=3 v=15
	Qualitative Operational test			
5.	CAQI-field test	Choice between mixed-mode scenario's Fieldwork procedures Unresolved problems and changes.	CAQI 1-1 meta-interviews with respondent and interviewer debriefing questions, reaction coding and focused probes.	n = 688 i_{CAPI} = 21 i_{CATI} = 13

v=consulted volunteer respondent, i=interviewer, e=expert, n=sample respondents

The joint questionnaire team did not expect the QDRC to test individual items (validity, acceptance, etc.). Table 1 presents the set-up of the qualitative content test.
The focus groups (Step 1) followed an individual 'ordinary' interview since particular questionnaire items had to be discussed. In the focus groups, the issues addressed by the QDRC were the first three required ones mentioned above. Although respondents' attitudes towards the whole POLS interview were generally positive, the focus groups led to serious doubts about the position of two question sets. Respondents expressed a kind of life cycle attitude towards the question set on education: instead of a position at the end of the joint questionnaire, following a number of topics on living conditions, they preferred a position just before the question set on paid work or on job search, or even before the time use question set preceding the latter two. Several respondents considered the general overview about being a victim of a criminal offence too personal for a position between the joint questionnaire and the particular module, and preferred it at the end of the whole POLS questionnaire.
These Step 1 findings led us to administer four questionnaire variants, to four respondents each, in the first series of meta-interviews in Step 2. The variants differed in question-set order only, according to two possible positions for education and training and two possible positions for victim of a criminal offence.
Another issue addressed in the meta-interviews in Step 2 as well as in the meta-interviews in Step 3 and Step 4 resulted from an risk assessment of the joint questionnaire by a QDRC expert. This assessment consisted of a careful confrontation of the item formulations with the stated statistical aims. The measurement error risks were categorised according to an eclectic list of risk factors (which is not as detailed as the cognitive coding scheme of Lessler and Forsythe (1995).) The eclectic list is presented in Table 3 in the section 'Joint questionnaire Test Results According to an Eclectic Classification of Measurement Error Risks' below. 'Risk hypotheses' were formulated for 22 of the 46 indicators in the joint questionnaire. These hypotheses concerned potential problems with question formulation, answer format, and validity in general (i.e. the degree to which the data that can be collected fit the intended use of an indicator). For any given risk hypothesis, a corresponding 'meta-interviewing task' was defined, almost exclusively in the form of a focused probe, cf. the methods column of Table 1.

Step 2 of the qualitative content test was carried out while using the technique of Computer-Assisted Qualitative Interviewing (CAQI), which was also used for parts of Step 3 and 4. This means that the meta-interviewing tasks were interwoven with or added to the questionnaire for POLS, resulting into a so-called 'meta-questionnaire'. As a general result of all meta-interviews (16 in Step 2, 15 in Steps 3 and 4) the preferences of respondents for the position of the two question sets mentioned above corresponded with the preferences of respondents in the focus groups. The joint questionnaire items turned out to be comparable

in CATI and CAPI settings, and to take an average interviewing time of 15 minutes.
Many of the risk hypotheses were confirmed in Steps 2 and 3 of the qualitative content test, but unexpected problems turned up as well. Some indicators revealed more than one problem. As a result, one indicator was eliminated from the joint questionnaire, while twelve others were reformulated by three experts and tested in Step 4. These new formulations provided solutions for some problems but again revealed new ones. Some reformulated indicators remained problematic if the reformulation turned out to amend only one of several problems. A summary of expected and detected problem indicators is given in Table 2. Note that for 18 out of 22 indicators (82%), expectations about problems were confirmed. Other unexpected problems were detected for twelve indicators, that is for 40% of all indicators found to be problematic.

Table 2
Total numbers of indicators and the expected, confirmed and detected problematic ones in each question set

Question set	indicators	expected	confirmed	unexpected
1. Household	5	0	0	1
2. Time use	3	2	2	0
3. Education	2	1	1	1
4. Paid work	6	2	1	1
5. Job search	4	2	1	0
6. Housing	6	2	2	2
7. Health	5	4	4	1
8. Leisure time	12	8	6	4
9. Household income	3	1	1	2
Totals	46	22	18	12

Consider what happened to the indicator time used for housekeeping, mentioned above. The respondents participating in the focus groups of Step 1 indicated that the key concept ‘housekeeping’ was too general. To investigate this, the interpretational probe ‘What does [the term] housekeeping activities mean to you [in this question]?’ was used in Steps 2 and 3. There was indeed a great, and partly unexpected, variation in activities mentioned. For example, some men included bookkeeping activities at home, and some housewives included the making of an appointment with the garage for car maintenance. Moreover, the indicator ‘time used for housekeeping’ did not prove a valuable check on ‘main activity’, because the former did not discriminate between the categories of the latter. In general, respondents experienced some difficulties specifying the number of hours a week spent on housekeeping activities. For a further test in Step 4, the following alternative was used:
Main item: ‘The next question is about housekeeping activities, like tidying up and cleaning, washing and ironing, shopping, cooking and doing the dishes, and caring for plants and pets. How many hours a week do you yourself usually spend on such housekeeping activities?’
A few reference set probes were used to test this item: ‘Do you indeed do all these housekeeping activities? If not, which ones don’t you do? Are there any other housekeeping activities you spend time on?’, followed by the expanding probe ‘If so, how many hours?’ This alternative was an improvement: for two-thirds of the respondents, all activities mentioned as an example were relevant and were considered to be exhaustive.
Another example concerns an item on the indicator *dwelling size*. The original question in Steps 1 to 3 was: ‘How many *rooms* are there in this dwelling?’ An interviewer instruction gave a precise and comprehensive description of rooms *not* to be included. Unfortunately, however, this description was to be given only if the respondent needed clarification. In Step 4 we tested an alternative version in which the rooms to be counted are explicitly named in the item: ‘How many living rooms, bedrooms and working rooms are there in your house?’. The test was done by an explanatory probe: ‘How did you get your answer?’. The interpretation of ‘rooms’ appeared to be more uniform, but the ambiguity of one of the descriptive terms presented a new problem: ‘Does ‘working room’ include a study or a kitchen?’.
The above examples concern the cognitive domains tapped by meta-interviewing. A non-cognitive problem was experienced with the item ‘Do you feel lonely sometimes?’, to be used as a subjective indicator of social participation. This item raised serious affective objections. The focus groups (Step 1) already showed that the item is perceived as a rather sensitive one. In the meta-interviews (Steps 2 and 3) the respondents, though willing to give an answer, felt uncomfortable by the abrupt transition to the next item: ‘How many hours a week do you usually watch television?’ Respondents simply expect a sensitive topic like loneliness to be addressed more fully than by just one item. A similar example is the item asking for the main income source of the main breadwinner. This item may be completely clear, but might nevertheless somewhat offend the breadwinner’s partner (who may do paid work as well).

4. Joint questionnaire test results according to an eclectic classification of measurement error risks

For the 22 joint questionnaire indicators involved in the initial risk assessment of the qualitative content test, measurement error risks were classified according to Table 3. This scheme does not apply to indicators but to separate items. Some indicators consist of more than one item and for some items more than one measurement error risk was expected. The number of expectations thus exceeds the number of joint questionnaire indicators (22). The scheme does not include risks due to item context effects, acquiescence, and the like. (Order effects of answer categories fall under “LF-structure”.) It is an eclectic scheme based on various sources in the literature (e.g. Foddy, 1993) and it is not meant to replace sophisticated coding schemes used in (cognitive) expert appraisal, cf. Lessler and Forsythe (1995). Our scheme should rather serve as a list of questionnaire design risk factors, which may be used by designers of questionnaire prototypes for a practical discussion on hypotheses to be (pre)tested, and for classifying pretest findings and discussing possible actions to amend problems.

Table 3 shows that while many expectations were confirmed, many unexpected problems arose as well. For example, 12 items were expected to have an unclear or ambiguous meaning (QM). This was confirmed for eight of them, but five other items unexpectedly turned out to show similar flaws. In total, on item level, two-thirds (20 out of 31) of the expected problems were indeed found. Of the total of 71 problems found, 51 (72%) were unexpected. If we recall from Table 2 that, on indicator level, 82% of expected risks were confirmed, and 40% of detected risks were unexpected, we may conclude that the formulation of risk hypotheses does not obviate the need for an actual test, which is necessary both for the explanation of the nature of the risk and to trigger the respondent to provide meta-information spontaneously.

5. POLS qualitative operational test

After completion of the qualitative content test the emphasis shifted from the contents of the questionnaire to fieldwork procedures. The first goal of the qualitative operational test (Step 5 in Table 1) was to compare two mixed-mode scenarios

Table 3
An eclectic classification of measurement error risks on item level, and their prevalence in the qualitative content test (Step 2)

Problem label	Description	C/E	U/D
AS: Applicability/Suitability	Item is not realistic enough:	1/1	2/3
	– Non-existent or inaccessible data	-/-	1/1
	– Hypothetical or fictitious data	-/-	1/1
	– data referring to someone else	1/1	0/1
QM: Question Meaning	Item has unclear/ambiguous/ unintended meaning	8/12	5/13
	– Misleading or unclear instructions	4/4	3/7
	– Reference set (frame) not sufficiently specified	1/2	1/2
	– Ambiguity or vagueness in item meaning	1/1	1/2
	– Unintended, though univocal, item meaning	2/5	0/2
KC: Key Concept Meaning	Key concept	7/10	6/13
	– has unclear or ambiguous meaning	-/-	5/12
	– is unknown or unnoticed	7/10	1/1
CD: Cognitive Difficulty	High cognitive burden to respondent:	3/3	6/9
	– Difficult recall/recognition	-/-	0/0
	– Difficult deduction (estimation, guess)	3/3	1/4
	– Difficult judgement	-/-	
	– Complicated answer format		4/4
TD: Technical Difficulty	Complicated item conditions/presentations	-/-	1/1
	– Too many key concepts or clauses	-/-	0/0
	– Double negation/other syntax complexity	-/-	0/0
	– Implicit assumptions	-/-	1/1
	– Unclear presentation (layout/intonation)	-/-	0/0
LF: Logical Flaw	Formulation or routing logically incorrect	1/2	7/18
	– Question-answer (Q/A) discrepancy	0/1	7/7
	– Incorrect/incomplete Q/A structure	-/-	6/6
	– Conflict with previous answers	1/1	2/3
	– Conflict with language rules	-/-	2/2
MA: Motivation/ Affection	Undesirable item	0/1	14/14
	– Goal of item unclear or insufficient	0/1	11/11
	– Item too intrusive or too personal	-/-	3/3
SN: Social Norms	Unbalanced, directive or non-neutral item	0/1	0/0
	– Non-neutral concept or directive task	-/-	0/0
	– Unbalanced or non-neutral answer	-/-	0/0
	– Risk of social desirability	0/1	0/0

C/E = the ratio number of Confirmed / number of Expected problems, U/D = the ratio number of Unexpected / number of Detected problems

(CAPI/CATI and CATI/CAPI) under realistic fieldwork conditions in order to make a choice for one of the two on empirical grounds. For this purpose about 700 respondents were sampled. Respondents who participated in the CAPI/CATI scenario partook in one single CAPI interview in which both the joint questionnaire and a particular module were administered. For initial non-respondents (after several trials) a CATI interview for the joint questionnaire was planned, followed by a CAPI interview for the module. As for the CATI/CAPI scenario, a CATI interview was conducted to administer the joint questionnaire, followed by a CAPI interview for the module. Initial non-respondents were re-contacted for one single CAPI interview including both joint questionnaire and module.

The criteria for the choice of a scenario were response rate, operational aspects and respondents' and interviewers' opinions on the scenario. Therefore, debriefing questions for respondents and for interviewers were added at the end of the questionnaire as an evaluative part of the interview. Moreover, interviewer debriefing sessions were organised after the fieldwork had been completed.

A second goal of this operational test was to investigate some problems which were (partially) unresolved in the qualitative content test. Nine indicators were selected for further investigation. Eight of these had been reformulated and tested in Step 4 of the qualitative content test. For two indicators, two evaluation questions were added to the questionnaire: an expanding probe and a (closed) check question. These evaluation questions for sample respondents in the qualitative operational test were quite similar to the focused probes for volunteer respondents in the meta-interviews in the qualitative content test. To test the other seven indicators, respondent reactions were coded during the interviews. By adding the debriefing questions, probes, check questions and reaction coding to the questionnaire, the interview became a meta-interview using CAQI. These interviews were conducted by CAPI and CATI interviewers especially trained for the purpose.

For the CAPI/CATI scenario the net sample size was 338, for the CATI/CAPI scenario 298, excluding 52 addresses without telephone or with an unlisted number, which were contacted using CAPI. Table 4 gives the various response results. In terms of

Table 4
Response on mixed-mode scenarios

Mixed mode scenario	Net sample size	Joint questionnaire		Joint questionnaire and particular module
		CAPI (incl. particular module)	CATI	
CAPI-CATI	338	220 (65%)	14 (4%)	222 (66%)
CATI-CAPI	298	40 (13%)	120 (40%)	113 (38%)
CATI-CAPI (unlisted)	52	30 (58%)	0 (0%)	30 (58%)
Total	688	290	134	365

The figures in italics refer to the initial non-respondents who were recontacted for the joint questionnaire.

response rates, the CAPI/CATI scenario performed better than CATI/CAPI. For a complete set of data (joint questionnaire and particular module), 66% response was reached with the CAPI/CATI scenario, and 38% with CATI/CAPI. Also for the joint questionnaire only, the CAPI/CATI scenario resulted in higher response rates: 69% compared with 53% in the CATI/CAPI scenario. To summarise, the figures in Table 4 indicate that (1) gathering all data in one single interview (CAPI joint questionnaire + particular module) resulted in higher total response rates than gathering data by two interviews (CATI joint questionnaire, CAPI particular module), and (2) response rates were higher when respondents were initially contacted on the doorstep (65%) than when initially contacted by phone (40%).

However, in interpreting these figures it is important to keep in mind that this was a first field test of POLS. The test brought to light many logistical problems, which especially affected the response rates of the CATI/CAPI scenario adversely. These problems were: (1) a very tight time schedule, making it impossible to re-contact all initial non-respondents for the joint questionnaire and respondents for the particular module, (2) a lack of experience with mixed mode designs among interviewers (concerning the attunement between CATI and CAPI interviews) and among automation staff (concerning data-transmission and processing).

The debriefing questions at the end of each interview concerned the length of the interview, the interest of the respondent in the questionnaire, the understanding of the questions, the respondent's cooperation during the interview, the interviewer's opinion on the flow of the interview, and the burden of the interview task.

The results on these questions do not differ greatly between the various scenarios and between interviewers and respondents. Hence, they are not very conclusive as to the choice of a scenario. This corresponds nicely with the findings of the qualitative content test, which showed no differences between the mode used for conducting the joint questionnaire.

In the interviewer debriefing sessions, the interviewers expressed a clear preference for the CAPI/CATI scenario. They came up with three reasons. First of all, the possibilities of refusal conversion with non-respondents were better in this scenario. In the field the same interviewer would visit non-respondents again, while in the CATI procedure non-respondents could seldom be re-contacted by the same interviewer. Moreover, in the field there was less time pressure during the introduction than on the phone. Secondly, in the CATI/CAPI scenario it was always necessary to re-contact the respondents for the particular module. This resulted in additional non-response, even if respondents had agreed to participate in a follow-up CAPI interview at the end of the joint questionnaire CATI interview. Furthermore, the process of switching from one mode to the other worked out better in the CAPI/CATI scenario. In summary, response rates, operational aspects and interviewers' opinions all argued in favour of the CAPI/CATI scenario.

The Blaise questionnaire prompted the interviewers to code respondent reactions immediately after the survey questions, applying to the seven indicators. The computer screen listed seven codes: respondent asks for repetition of questions, asks for clarifications of questions, answers while question is still being read, expresses doubt between answering categories, does not know what category to choose, changes the answer, becomes irritated. Interviewers could enter all appropriate codes, or, if none of the listed or other reactions appeared, none at all. The results from the reaction coding corroborated earlier findings: due to the reformulations, some problems were solved but some remained.

One indicator that was reformulated after the qualitative content test was the indicator on *'Disabilities'*. In the new version this is a set of three long questions (on disabilities at home, at work, or in other situations), posed immediately after each other. The wording of these three questions is very similar, differing only on the location to which it applies, indicated at the very end of each question. Reactions were registered for 20% of the respondents. More than half of these did not wait till the end of the last question, but answered while the interviewer was still reading. This indicates that the question is redundant and that respondents would like to go on with the interview, especially when they are not at all disabled.

With two indicators, evaluation questions were added to the Blaise questionnaire. An expanding probe was added to the alternative indicator on *'Time used for housekeeping'*, as used in the qualitative content test. The probe referred to the overall number of hours asked for in this indicator, in which the question includes examples of relevant housekeeping activities. At the end of the joint questionnaire, long after the indicator itself, the probe elicited respondents to specify the number of hours spent on each of the housekeeping activities mentioned as an example. Results suggested an underestimation of the overall number of hours spent on housekeeping activities. For about 80% of the 424 respondents the overall number of hours was less than the sum of the hours spent on each of the activities, with a difference of two hours each day for about 50% of the respondents. An explanation may be that some activities included as examples are overlooked, or that the overall number is based on an unspecified, rough guess. Nevertheless, the alternative formulation was felt to be an improvement of the original one, also considering the results of Step 4.

Another probe, in the form of a closed check question, referred to the answer to the reformulated indicator on *dwelling size*, also used in the qualitative content test. Immediately after the question, it was checked whether respondents had included the kitchen, the bathroom and/or the cellar. About 20% of the 424 respondents mentioned that they had indeed included rooms other than the ones named in the question. The check question also evoked a spontaneous revision of the original answer (9%), even though some respondents (6%) did not include other rooms than those mentioned in the question. These results indicate that (in accordance with the results of Step 4) a uniform interpretation of the concepts referring to the intended rooms is hard to obtain: a kitchen may be viewed as a working room, and a large kitchen and a large cellar as a living room.

6. Discussion.

The POLS pre-testing programme has been rather extensive. This article has not elaborated on all the modules and question sets added to POLS compared with the old 'separate' surveys: a mini time-use survey, a new survey on accidents and the implementation of the housing demand survey (formerly PAPI). All these questionnaires were pretested in sub-steps, according to the procedure used in the pre-test of the joint questionnaire, and with a focus on their qualitative content and/or fieldwork

procedures. In testing the qualitative content of POLS, cognitive stimuli were omitted. In particular, an explicit general stimulus to 'think aloud' was consciously avoided, for reasons of efficiency. Such an unfocused procedure was suspected to take up too much time and respondent concentration, possibly at the cost of the meta-information that we could obtain for a test of the specific risk hypotheses. The focused probes we actually used turned out to trigger the respondents to expand freely (think aloud) on the indicators and items that the probes referred to, which resulted in a considerable amount of useful information to improve the questionnaires.

The comprehensive pretesting programme for POLS would not have been possible without CAQI as a pretesting technique perfectly adapted to the computerised data collection design. For both qualitative POLS tests (content and operational), CAQI turned out to be a convenient way to develop and use meta-interviewing protocols, that is the script for applying the meta-interviewing tasks during or after the ordinary interview. CAQI appears to be particularly rewarding to obtain a moderate standardisation of the meta-interviews and a quick and easy transition from one pretesting step to another. Items and their ordering can be adapted in the course of testing, as can the various meta-interviewing techniques to be used. Moreover, the CAQI meta-questionnaire makes it easy for meta-interviewers to become acquainted with the test protocol, and to focus on the communication with the respondent. In general we may conclude that CAQI helps to improve the quality of the meta-information. A more extended discussion on CAQI is presented in Snijkers (1997) and Snijkers et al. (1996).

References

Akkerboom, J.C., 1996. Labor für die Entwicklung und den Test von Erhebungsinstrumenten. In: *Statisches Bundesamt: Pretest und Weiterentwicklung von Fragebogen*, Schriftenreihe Spektrum Bundesstatistik, Band 9, (Stuttgart: Metzler-Poeschel)

Akkerboom, J. and F. Dehue, 1997, The Dutch Model of Data Collection Development for Official Surveys. In: *International Journal of Public Opinion Research*, vol. 9, no. 2, pp. 126-145

Foddy, W., 1993, *Constructing questions for Interviews and Questionnaires. Theory and Practice in Social Research* (Cambridge: Cambridge University Press)

Lessler, J.T. and Forsyth, B.H., 1995, A coding system for appraising questionnaires, In: Schwarz and Sudman (Eds.): *Answering questions.* (San Francisco: Jossey-Bass)

Snijkers, G.J.M.E., J.J. Akkerboom, I.F.E. Kuijpers and E.D. de Leeuw, 1996, *Computer-Assisted Qualitative Interviewing: An Intermediate Technology of Quality Assessment.* (San Antonio, Texas, USA: Paper presented at the INTERCASIC '96 Conference, Dec. 11-14)

Snijkers, G.J.M.E., 1997, Computer-Assisted Qualitative Interviewing: A Method for Cognitive Pretesting of Computerized Questionnaires. In: *Bulletin de Méthodologie Sociologique,* Nr. 55, pp. 93-107.

Standardisation of computer-aided survey collection

Wim B.M. Faessen and Frans A.M. Kerssemakers

Summary

POLS occupies a major place in the development of computer-aided survey collection at Statistics Netherlands. This chapter describes how in this development the integration of question modules from different surveys was accomplished and linked up with management systems. The data processing system has been modernised and a new documentation system set up.

1. Introduction

Survey collection as considered here consists of three different main steps: data collection, data processing and documentation. Redesigning the social surveys into the integrated survey of POLS deeply effected parts of these processes.
Data collection had already been modernised before POLS by changing from paper and pencil to computer assisted interviewing. The program language developed and used at Statistics Netherlands is Blaise (see Section 2). Accompanying the introduction of Blaise was a call for standardisation of questionnaires (Section 3). A new generation of Blaise - Blaise III - was used from the very start of POLS; it is very suited to the integrated survey system as offers an integrated control centre for survey processing (Section 4). Although much has still to be done to fully exploit Blaise, its value has already been proven (Section 5).
In POLSsubsequent data processing is conceived of as the total of standard actions on the collected data, resulting in a common statistical base from which user-defined statistical products can be built. It is organised in such a way that the data can be analysed in all stages. For this purpose a user-friendly information system has been developed to select variables and micro data. The main stages are: *initialisation*, *execution* (checking data, deriving variables, weighting) and *output*. An important restriction on the process is 'non-increasing complexity'. The process has been partly redesigned and largely automated (Section 6).
As analysis requires knowledge about the process, the data and the contents of a survey, new documentation standards were developed for POLS (Section 7).

2. The introduction of computer-aided survey collection

From 1984 up to about 1992 Statistics Netherlands introduced fairly radical changes in its process of data collection from persons and households, implementing methods of computer-aided surveying. Following the introduction of small handheld computers and laptops suited for multi-survey purposes, a home-made system was developed for computer-assisted interviewing. The combination of interviewing and data entry opened up new possibilities for interactively checking and editing data in an early stage of the collection process, preferably while still in contact with the respondent. With paper forms, data entry and data editing could also be put in the hands of one specialist and thereby decentralised, thus replacing the traditional macro cycles of computer checking and manual correction by micro cycles in which data processing takes place interactively with the help of an intelligent program (Bethlehem, 1987). All this led to the aim of integrating existing tools for survey data processing into a single system in which the data concerned are described only once in a uniform way. The language for describing data in a structured way was called Blaise (based on the programming language Pascal), as was the system for controlling different applications within a single survey based on a common data description (such as questionnaire machines, tools for manipulating data or generating set-ups for statistical packages). It was decided that from 1992 all interviewing should be computer-assisted and the home-made Blaise system be used for all surveys.
Computer-assisted interviewing at Statistics Netherlands started with the large-scale application of telephone interviewing in January 1984. Face-to-face interviewing with a computerised questionnaire was first used at the end of 1984 in order to examine whether this kind of computer-assisted personal interviewing (CAPI) was acceptable for both respondents and interviewers. For telephone surveys a ready-made CATI package was leased. But for CAPI a home-made and memory-saving language, called QUEST, had to be developed. When preparations started in 1985 for an extensive, continuous Labour Force Survey one of the main objectives was a quick delivery of outcomes. Both this and the large sample size made the use of handheld computers for data collection worth trying. The whole CAPI system for just the Labour Force Survey came into full operation in January 1987. The collected data were sent back by telephone by modem.
Blaise was first used in 1987 for data entry and interactive data editing of paper forms. This application has since been called CADI (computer-assisted data input). But once Statistics Netherlands could afford some 300 laptop computers (Toshiba T1000) that could handle different questionnaires (on diskette), a rapid shift occurred to computer-assisted face-to-face or personal interviewing (CAPI). By 1989 three surveys on living conditions had been converted from paper to CAPI, namely those on crime, health and quality of life. At the same time these three surveys were made continuous, running throughout the year. Telephone surveys, which had been computerised from the start in 1984, started using the CATI-machine of Blaise in 1990. In fact, the shift from CADI towards the interviewing mode (CAPI or CATI) mainly implied a change in the user-interface, achieved by splitting the screen into two parts for the interviewer. The CADI-screen for data entry moved to the bottom half and is now called the answer page. The top half is the question text area. Here, in most cases, one question is displayed at a time, together with the answer categories. By presenting a page with input variables in the bottom half, Blaise offers an often preferred page-based interface (Pierzchala, 1997). Concurrently the Blaise tool Manipula was being developed for common ways of manipulating data and files. Both a tabulation (Abacus) and a weighting (Bascula) program were added. Meanwhile Blaise had become a well-known and widely used package internationally, with its own user meetings in 1992, 1993, 1995 and 1997. Altogether Statistics Netherlands now uses Blaise for monthly gross samples of about 40,000 addresses or persons, for some twenty different surveys a year (CAPI and CATI).

Accompanying questionnaire-based Blaise-applications for interviewing, data manipulation and tabulation, major automated or semi-automated systems have been built to perform tasks like data communication with the interviewer and additional special coding of occupation, firm and education at the central office. Much effort was put into developing management systems on the laptop computer for processing data files (back-up, compression, encryption) and handling cases (addresses or persons). An important principle from the viewpoint of fieldwork control by both the central office and the interviewer was that every sample element should be traceable and its status known at any time. It soon became clear that interviewers needed their own computer-

assisted management system to support them in data communication, case-handling, switching from one survey to another, and their reporting on visits and results, concurrently checking on completeness and consistency. The result was a dedicated menu-driven system, named LIPS (Laptop Information system for Personal Surveys), which came into use in 1990 and has been in full-scale operation since 1992. On the other hand, the CATI facilities for call scheduling, making appointments, creating management information etc. were fully integrated in the Blaise system itself. Both the call data and the appointment data were made part of the questionnaire. So here all Blaise tools for data manipulation could be used. (Hofman et al., 1991).

3. A call for standardisation

A properly working interviewing machine will follow the routing instructions as specified in the Blaise-questionnaire and not allow out-of-range codes. The interviewer can no longer skip questions by mistake or answer the wrong questions. So the data will be in accordance with the pre-specified rules laid down in the Blaise questionnaire. Besides, data entry is done by the interviewer on the spot, so the researcher has computer files with the original data registered by the interviewer at his immediate disposal. An important consequence of obtaining syntactically correct data from the Blaise machine was that the bulk of manual editing of paper forms sent to the central office was no longer necessary. Computerised questionnaires make data inconsistencies far less likely, reducing the need for data editing to such an extent that it was decided to place the remaining data editing in the field altogether. (An exception was, and still is, made for the coding of firm, occupation and education by classification specialists.) With the elimination of this second round, remaining inconsistencies within the range offered by the machine have since been tackled by structuring the questionnaire so as to prevent errors from happening, and by building both ‘hard’ and ‘soft’ checks at crucial points or where errors are likely to be made. Thus, attention was focused on the interviewing situation and the questionnaire to a much larger extent than when inconsistencies would have been spotted in the extra round of data editing anyway. However, the shift from post-editing data towards designing and specifying a questionnaire in which many kinds of error-prone situations are already anticipated and coped with, requires a great deal of effort. Using Blaise optimally in order to enhance data quality often meant, for instance, that one-question concepts were elaborated into a large number of simpler, more self-evident tailor-made questions. The resulting quite complex structures really put Blaise to the test, as was the case with the revised Labour Force Survey (now in use since 1992).

Inherent in the process of tackling the greater structural complexity of question forms and the problem of increasingly required developmental resources, was the general and lasting pressure towards standard solutions. A household roster, for instance, in which the members of a household and their interrelations are listed and checked can easily become a rather complex question module. Once developed and successfully applied, there is hardly any reason why other surveys should not profit from the efforts made, the experiences gained, and the accompanying training given to the interviewers. This tendency was strengthened by the fact that there were only a few specialist writers. They naturally wanted to re-use solutions already arrived at. Later on it was formally decided to prescribe standard solutions for a number of demographic (sex, age, marital status, nationality, place of birth) and socio-economic variables (education, firm, occupation, status-in-employment etc.), sometimes allowing both a minimum and a more extended version. So, when POLS came up, a number of standard question modules for the so-called harmonised classification variables of the basic questionnaire were already in use, among other things for surveys to be integrated in POLS like those on crime, health and quality of life. This does not mean, however, that all these modules were already crystallised into well-tested and fully equipped solutions, satisfactorily embedded in the working process and integrated in the analysis of a survey. Some were still controversial, insufficiently elaborated or not fully communicated yet. But, at least, surveys were on their way to standardisation.

4. Integrating surveys with Blaise III

Since 1996 a new generation of Blaise has gradually been implemented at Statistics Netherlands: Blaise III. Apart from enhancing facilities that were already available in the old Blaise, important new features have been added; to name some important ones: more flexible layout-facilities, multi-language use, possibilities for calling question blocks independently of the position in the routing or inserting one’s own dedicated program, and the introduction of parameters for making question modules relatively independent of the context of a particular questionnaire. But the core of the new Blaise came from the need to integrate separate tools for data processing which were being developed *ad hoc* as computer-aided survey collection was expanding. They had to be linked up through the Blaise system itself. Consequently Blaise III has been designed to offer a well- integrated and encompassing control centre for survey processing. The starting point is a model in which information about data is specified in machine-readable form. Here the meaning of the data in a particular field is mostly represented by a question, which - by applying the interviewing machine -can be displayed on a computer-screen. But it can also be a description of a variable from a secondary source, like a tax register or a company’s account. Generally there are also some rules to be obeyed for obtaining the actual data, such as instructions about which information should be obtained for different target groups. In a questionnaire this is called the ‘routing’. To enhance the quality of the data, checks can also be added. Once the data model has been specified, applying CAPI or CATI could provide the actual data. But the model can also be used for data manipulation with Manipula. The information in the model is called meta data and is available throughout the Blaise system for all kinds of applications. Meta data can also be translated for use with another program outside Blaise, like SAS or Oracle. To do this Blaise uses a tool by the name of Cameleon.

Lastly, the addition of a tool called Maniplus to Blaise III should be mentioned. Maniplus makes it possible to combine all kinds of applications to perform a series of related actions in a survey. It is essentially a self-starting Manipula which can activate questionnaires, invoke procedures for file handling or pass control to another Maniplus application. But it can also be made user-controlled by interrupting the execution at certain points, for example to let the interviewer first take a decision from a menu or dialogue box before going on. Thus, both the management system for the interviewer and the CAPI questionnaire can be integrated in one Blaise control system.

However, what was especially important for POLS was the possibility to combine different questionnaires in one interview. For the very design of POLS (see Winkels and Everaers in this issue) implies a joint basic questionnaire for all non-overlapping subsamples of POLS, whereas the number of subject-specific questionnaires which are to be combined in POLS is in principle indefinite. As new subjects are included and others removed, management of the building and maintenance process requires that the specific questionnaires should be built in a modular way, making them independent from each other as much as possible. The integration concept rests entirely on the joint basic questionnaire. The latter could have been duplicated in different data models, one for every subject-specific questionnaire. It would have been possible to put these data models under the umbrella of a single case management system and have them handled by Maniplus such that POLS would still look like a single survey to

the interviewer. But the basic questionnaire in particular is meant to be relatively stable and independent from the variety of subject-oriented wishes and developments. And if the data models are to be recognisable entities linked to a particular subject, changing the basic module should leave them unchanged, if possible. Subject-matter specialists who are responsible for a particular data model should not be bothered with things they did not initiate and that may be irrelevant to them. It was therefore decided to have a separate data model for the basic questionnaire on the integrative level alongside a set of data models for which, as before, independent and non-overlapping samples are drawn.

First, the joint basic module is asked. This also includes themes that require a large sample (e.g. crime victims or accident rates). When this data model has been completed a Manipula set-up takes care of writing the data that may possibly be needed in any one of the specific data models in a separate 'external' data model. Through a dialogue box in Maniplus, the interviewer can now open the particular data model to which the sample person was assigned. After concluding this part of the interview some questions still have to be asked about the household of the sample person. Here data is collected from the person with the highest income, such as income itself, education and occupation. As this may involve a shift of respondent, these questions, which actually belong to the basic module, are asked near the end of the interview. They are put in a separate data model, which is only auxiliary because the data are subsequently inserted by Manipula (block moving) in the files of the original data model. Thus, the latter will finally contain all data from the basic questionnaire. For a particular case, only the data from the first model and from the applied subject-specific model are sent back by telephone. So, for the actual questionnaires POLS uses a *1-n-1 structure of data models*. The two constants represent the data models for the generally applied basic questionnaire. The variable n represents the data models for the specific modules, from which one is chosen per interview. Afterwards extra data models can be used, as they are, for follow-up studies among screened cases (i.e. module 4).

The interviewer can always return to preceding data models to change already given answers. However, as this may effect routing, checks and computations in subsequent models, Maniplus forces the interviewer to open these models again so that potential changes will be processed automatically. This may be inconvenient if it occurs frequently. In practice it can be dealt with by choosing the right, relatively independent modules.
The management system on the laptop which is still used by Statistics Netherlands in connection with the old Blaise is called LIPS. Written in Pascal, it is a dedicated system for some well-defined repeating tasks such as choosing addresses, making interviewer reports and treating sample elements. Although robust, the system is not easy to adapt or extend. Once the decision was taken to use Blaise III for POLS it was clear that a new and preferably more flexible system was needed.

Inspired by the possibilities of Blaise III and to improve accessibility it was decided to build the new management system entirely with the tools of Blaise III, especially with Maniplus. So, the modular design of the POLS questionnaires, mentioned before, was also used for the new LIPS. Being a management system, it was designed to be as independent from a particular survey as possible. From a Maniplus set-up, called LIPS.MAN, separate Blaise questionnaires are started to handle addresses or to make an interviewer report. Then control may be passed to another Maniplus set-up, called POLS.MAN, which steers the survey-specific questionnaires to be used in the actual interview. In a way one could define LIPS now as only taking care of the steering of the respective questionnaires and the presentation of the main entities: survey, address and sample element. That is how the core of the system is made resistant to changes. All survey-dependent codes are put in separate components which are much easier to adapt or replace. Management can now be geared to the demands of a particular survey or even be changed for a certain period. General and specific parts have been separated. Adding a new subject can simply be accomplished by writing a new data model and specifying in the criteria to call it in Maniplus. In POLS this is called the Plug-and-Play principle. Moreover, at any point during the interview only the set-up that is in control and the particular data model that is used are loaded into the internal memory of the laptop-computer (which was extended from 4 Mb to 12 Mb).
The required set-up, however, is not trivial. Specialist knowledge is indispensable. Exchange of data between models has to be arranged and a lot of careful checking is needed in order to guarantee a correct and smooth proceeding of interviewer activities. Nonetheless, the reduced complexity of the modular design clearly outweighs the complexity caused by the introduction of a separate level of control.

5. Converting to Blaise III

Somewhat contrary to expectations there were hardly any problems when the questionnaires of the old surveys had to be converted to Blaise III at the beginning of 1996. For this purpose Blaise III proved to be remarkably stable. The POLS-questionnaires themselves became simpler and easier to handle. Most discussions were about operating the user-interface and about the layout of the screen in particular. Generally, most writers think that specifying questionnaires has also become easier. On the other hand, Blaise III offers a lot of new opportunities which have to be mastered first. The use of block parameters, for instance, to optimise a modular design like the one in POLS, takes time. And to fully exploit the new possibilities source codes have to be checked almost line for line. But doing so is in no way a prerequisite for using Blaise III. In POLS old and new live together. Overall, the progress is a result of gradual enhancement. Many old parts are waiting to be redesigned by subject-matter specialists. In the meantime some aspects are being improved already, for instance, by renewing the date and time functions.

Perhaps because of the ample resources of Blaise III, a noticeable shift has already occurred from discussions about Blaise to the more fundamental question of how questionnaires should be organised so that maximum advantage can be taken of the data models once they are specified. The goals that are at stake here include subsequent data processing and analysis, maintenance and documentation of questionnaires, and comparability or exchangeability of parts between surveys. These are all paramount issues for a complex integrated design like POLS.

Statistics Netherlands strongly fosters interactive coding of open answers during the interview, both for reasons of efficiency (reduced coding staff) and quality (contact with the respondent). To achieve this quite a few concepts in POLS need a many-branched, hierarchical tree of questions (e.g. the part of the body that was mainly injured). Here the facility in Blaise III that leads the interviewer through a hierarchy of successive sets of alternatives has proven to be extremely helpful. It sometimes replaces complete modules of self-specified questions. If in addition the nested code lists are put in a type library, this helps to organise the questionnaire.

In CATI only the basic questionnaire is asked (i.e. modules 1 and 2). Except for two files, for the introduction and the 'goodbye' respectively, which are included separately, the data model used with CATI is the same as the one used with CAPI. Contrary to the old Blaise, the arrangements for using CATI can now be made completely external to the data model. For more details about the use of Blaise in POLS see Heuvelmans et al, 1997.

6. Data processing based on the principle of non-increasing complexity

Subsequent data processing has been organised as a multi-stage process. The main stages are *initialisation*, *execution* (checking data, construction of composite variables, weighting) and the *output* stage. An important restriction on this process can be formulated as 'non-increasing complexity'. The processing is organised in such a way that the data can be analysed in all stages. All stages and sub-stages are fully automated and a user-friendly information system has been developed to select variables and micro data (from the several POLS-modules) for internal use.
When they have completed the interview and corresponding data entry, the interviewers send the data to the central office by telephone. In the data collection department the records are first sorted by survey period (that is the period in which the data should have been collected), and per survey period they are divided into responses and non-responses, depending on the answer to a particular question in the interview. Accordingly interviewers are paid more or less automatically. Then the response records go to the classification unit where string variables on education, firm and occupation are classified. The resulting codes on the classification variables, which are derived in a standard way, are subsequently inserted in the Blaise records from the interview. Also added are automatically derived variables on household composition, socio-economic position and other variables derived from officially prescribed question modules. Together, the derived variables and the response files form the input files for the department of Socio-Cultural Household Surveys (SCHS), the initiator of POLS.
The processing of POLS data within the SCHS department comprises all remaining standard processing. This includes all actions pertaining to the received data that are of a general nature, not dedicated to a specific product for a particular customer. Among other things these include: data checking and data correction, deriving standard variables and weighting of data. Not included are derivations of specific variables, the coding of open answers, data imputation and data protection, as these actions depend on the specific product created and for whom it is intended. In short, data processing is conceived here as the total of actions on the raw data that bring into existence a common statistical base from which specific (customer-oriented) statistical products can be built.
This sharp distinction gives lends an important property to the whole course of data processing: *reproducibility*. At every moment all actions, from receiving the data up to the statistical base, can be repeated with the same result. Because of this the process of error correction in particular becomes more reliable. A detected error in the data is always corrected at the level of the originally received data. Then the entire data processing for the involved survey period is executed again. By doing so the successive error corrections do not cause an increasingly complex course of processing. At the same time there is a considerable reduction in the risk of new errors occurring as a direct consequence of solving old ones. The described procedure is also followed in the case of errors in the processing itself.
The data processing of POLS (within SCHS) is organised in three different stages:

Initialisation of data processing:
Data collection in the field is based on so-called fieldwork periods during which a certain portion of the sample has to be finished by the interviewers. Data processing, however, is based on the month in which the actual interview took place. So some rearrangement of the data between files is necessary. A second rearrangement concerns a partitioning of the variables into separate files: two files respectively for the fixed variables from the basic questionnaire and for those that are changing, and in addition a number of subject-specific files. Records from the latter files that are not considered as response regarding the involved subject are removed. Both response data and non-response data are saved, however, in a separate file for fieldwork analysis.

Execution of standard processing:
Processing takes place for each kind of file (fixed, variable, subject-specific) separately. So actions on the different files do not need to be synchronised. If there are new products in the output of a certain step (e.g. weighting factors) these are stored in a separate file. Such a modular data storage makes process management easier. For instance, if an error in the weighting program is solved, subsequent re-execution only produces new data in the file that contains the weighting factors. The other files remain unchanged and analyses on these files remain as valid as before.

Regular processing consists of the following steps:
- creating monthly history files (re-execution always starts from these files);
- checking and correcting data;
- establishing regional characteristics;
- deriving new variables from the data;
- creating files on a different time base (quarterly, half-yearly, yearly);
- weighting of sample units.

Creation of the common statistical base:
After closing a certain interview period and completing all processing steps, and after authorising the final results, data storage is chosen in such a way that it is ready for producing a wide range of statistical products. At this stage every module (fixed, variable and subject-specific) is assigned its own SPSS system file which contains all variables from that module and also the weighting factors for the sample cases.

One of the consequences of a mixed-mode approach (e.g. from CAPI to CATI to CAPI.) is that finishing the assigned cases in the field can take several months. The subsequent data processing can therefore not be completed for quite a long time. But because early detection of errors or mistakes during data processing is considered to be important a need was felt for easy access to the data at different stages of processing. To enable these kinds of analyses a so-called data selection system was developed, based on a meta data file which describes:
- all variables from each version of POLS, with their statistical name, answer type and whether attributes like 'don't know', 'refusal' or 'empty' are allowed;
- the relations between variables and questionnaire: which questions from the questionnaire belong to which statistical variables, and vice versa;
- the definition of the modules from the basic questionnaire and the subject-specific modules: to what version of POLS do they belong, and what is their validity period (starting and expiring date);
- the kinds of files (question variables, derived variables, weighting factors etc.) and the periods from which they are available for analysis;
- the structure of the respective files (represented by data models);
- the relations between question modules, files, data models, variables and values.

Making a selection involves interactively performing successive actions resulting in:
- an ASCII file with the values of the selected variables;
- an SPSS set-up with the statistical names of the selected variables, their answer range and value labels, and their position in the ASCII-record.

For analysis, the desired instructions then have to be added to the set-up, and the files from which the selection takes place have to be linked, using the case identification as matching key.

7. Documentation of data models and the POLS information system

For those who are able and willing to read Blaise, the data models of Blaise III can be specified in such a way that they are almost entirely self-documenting. Users should at least be able to read the block structure without being distracted by technicalities or large parts with checks and computations. Yet, even users with enough other programming experience often lack confidence when they have to use source codes for documentation purposes. We doubt whether this has much to do with Blaise in particular; that is to say, probably all computerised questionnaires of some complexity, a large size and with a lot of functionality will raise problems of documentation, not specifically Blaise.
In the meantime ideas are evolving within Statistics Netherlands for an automated aid, utilising the meta data that can be read from the data model and translated into another format by the Blaise tool *Cameleon*. What is needed is a set-up that selects meta data and interprets it as language codes with respect to the layout, and then a program to represent text in a common format so that it can be tailored interactively to the user's own wishes. WordBasic commands could be used for controlling Word. But first the main users have to define the functionality.
In short, Blaise-questionnaires can and should be made self-documenting for those who are used to reading or writing Blaise. In the absence of an automated - or even semi-automated - solution for those who cannot, POLS has its own hand-made documentation on paper. The problem with this is not so much producing it, but the maintenance, keeping it up-to-date. Therefore a prototype of an information system has been developed to be used as a tool for reading all the relevant POLS-documents, including both questionnaires in an end-user readable format and the 'self-documented' Blaise sources. This will probably help to keep them well matched.
An advanced data collection and data processing system deserves an advanced information system. Such a system should contain not only the updated questionnaires and data files, but also the backgrounds of a survey and other relevant information. Besides, it should also be able to link the information. To achieve this for POLS it was decided that the information system should be an automated one, with printing on demand as an extra service. A directory structure was installed which gives an overview of the relevant aspects of the survey and the data, including background information. Each directory includes relevant articles or other pieces of documentation (listings, questionnaires, etc.). When information changes because of new files or modules, new questions or publications, the directories are updated with the new parts. The main directories are:

- *Background and contents.* This directory contains articles on the history of the development of POLS and the logistical processes. Information on the design, on publications and (expected) changes is also incorporated in this directory.
- *Sample.* All information on the sample, sample size, target populations, sample frame, stratification, sampling methods and weights is described here.
- *Fieldwork report.* Relevant information on interviewer instructions, questionnaire guides, interview length, survey and item response and non-response and comments on the interviews by both interviewers and respondents are put in here.
- *Questionnaires and raw data.* This contains the different question modules. For each questionnaire there is a plain description in words and a routing scheme, a record description of the Blaise file and a listing of the variable and value names, and, of course, the Blaise source itself. Moreover, changes in the questionnaires throughout the years are recorded. Also described are the special coding process and the way standard variables are derived, and information is given on the data editing process.
- *Quality and reliability:* Information on the confidence margins and other data quality measures is contained here.
- *Output files for external use:* Before data files are made available for third parties outside Statistics Netherlands, they have to be protected against abuse and possible disclosure of individual data. The method of data enclosure and the composition of the file are shown here.

There are additional directories for more technical matters like specifications of missing values, imputations, weighting etc.
What is special about the POLS information system is the linking of the files and the searching system. It resembles the Internet. By clicking the mouse of the PC it is possible to get access to files and get related information on the files or variables that are shown. The information system was built using HTML and is intended to be used by our staff as well as by outsiders. Parts of it are to be made accessible through the Internet.

References

Bethlehem, J.G., 1987, The Data Editing Research Project. In: *Select 4,*
Automation of Survey Processing (Voorburg: CBS).

Heuvelmans, F., F. Kerssemakers and J. Winkels., 1997, Integrating Surveys with Blaise III. In: *Proceedings of the Fourth International Blaise Users Conference*, (Paris: INSEE)

Hofman, L.P.M.B., and W.J. Keller, 1991, Design and Management of Computer Assisted Interviews in The Netherlands. In: *Computer Assisted Survey Information Collection: International Progress*, American Statistical Association, 1991, Joint Statistical Meetings, Atlanta, Georgia.

Pierzchala, M., 1997, Optimal Screen Design in Blaise. In: *Proceedings of the Fourth International Blaise Users Conference* (Paris: INSEE)

The POLS programme as an example of Business Process Redesign: focus on the communication as a success factor.

Pieter C.J. Everaers and Hans A.F.M. Wajon [1)]

Summary

For the development and implementation of the integrated household survey (POLS), Statistics Netherlands started a redesign programme, along the lines of a Business Process Redesign project. The organisation of the work, the culture, the content of the surveys and the information technology are all elements of this redesign. The main results of the programme are the integrated household survey and the organisation of the work in the departments most involved in this survey. Special attention is given to the form and quality of the internal communication in the different phases of the restructuring process: from a simple idea, via formulation of the idea, the definition of the programme and the implementation of the programme in several well-defined and inter-related projects.

1. Introduction

In 1994 Statistics Netherlands switched over to its new organisation form. The radical reorganisation process preceding this switchover - known under the name of TEMPO - offered opportunities to redefine products and restructure statistical processes and procedures. The objectives of this restructuring were set down in a master plan (Van Bochove, 1996). In the organisation's business plan, these developments in the statistical processes are related to investments in staff training and job mobility, and to the structure of the processes assumed to be necessary to renew the organisation of Statistics Netherlands. With respect to the statistical information on individuals and households an increase in the integration of survey data and in the use of register data was foreseen.
This article describes the organisation of the restructuring process of the socio-cultural household surveys. The organisation of the work, the culture, the content of the surveys and the information technology, are all elements of this fundamental restructuring. The restructuring is worked out along the lines of a Business Process Redesign (BPR) project (see Hammer, 1993). The main results of this BPR project are the *integrated survey POLS*, and the organisation of work in the departments most involved in these surveys.

The implementation of the redesign process along the lines of BPR was new for Statistics Netherlands. In the first phase, an official document for the POLS programme was written. The structure of the POLS programme and the results may serve as an example for processes of change in complex organisations within statistical bureaux as well as other government departments. The programme has resulted in insights in the high level of commitment of employees to such a process of change. However, it also showed the inevitable conflicts between the hierarchical layers and the horizontal project organisation, and the effect of the changing level of commitment of principal actors. The organisational and managerial aspects of the redesign process are described in more detail in this article to show the immense importance of communication aspects in a BPR project. Indeed, communication proves to be the main success factor in BPR projects.

2. The origins of the POLS programme

Socio-cultural surveys generally gather information on themes like health, housing, crime, political and social participation, time use and attitudes on matters like criminal, environmental and sexual behaviour. Most of the former socio-cultural surveys at Statistics Netherlands had been set up in the early seventies and made use of Paper and Pencil Interviewing techniques (PAPI), with prestructured questionnaires and precoded answers. Most of the surveys result in outcomes valid for the total population, some in outcomes for particular subgroups. The fieldwork was done continuously, or periodically; some surveys were annual, some incidental and others held every three or four years. Each survey had its own weighting procedures, sampling techniques, etc.
The first origin of the POLS programme was the integration of these surveys. The arguments for this integration can be summarised as follows:

- The lack of consistency between the results of the socio-cultural surveys, the methods to collect the data, the weighting methods used, the cleaning and publication procedures, the wording and routing of the questions.
- The lack of related information (micro data) on core variables on a wide range of themes for a large sample and on a low regional level.

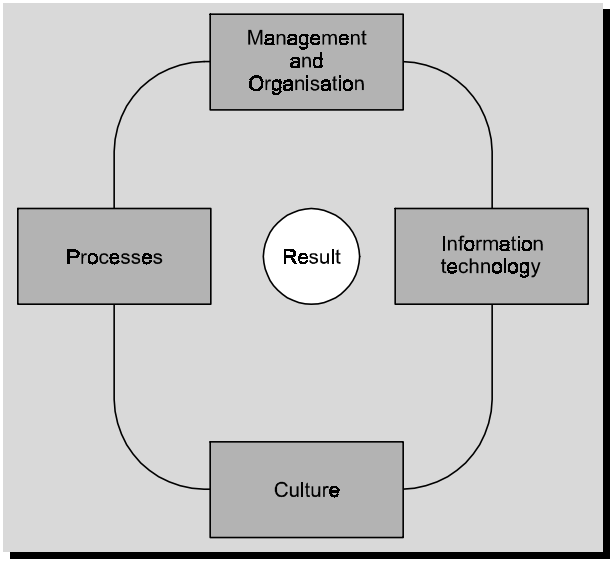
Before the TEMPO reorganisation, most of the statistical production processes were organised hierarchically. The former socio-cultural surveys were conducted by four separate subdepartments. Each survey was separately designed by staff from the conceptual department in co-operation with the fieldwork department. The division of tasks in the subdepartments was between researchers responsible for concepts, publications and internal and external relations, and specialists responsible for databases and tabulation. The size of these subdepartments (seven to eight staff) and the periodicity of the surveys resulted in a low level of specialisation: every employee was assumed to oversee the whole statistical process.
The idea of integrating the surveys into one survey was first suggested in the late eighties (Van Tuinen, 1995). However, the structure of the organisation of Statistics Netherlands, with four subdepartments, functionally divorced within four departments, each conducting several socio-cultural surveys, was a main obstructing factor. The optimum span of control was in conflict with the existence of small subdepartments.
The TEMPO operation resulted in a project-based organisation of divisions, departments and production groups and horizontally organised project groups. As a consequence, the four initially separate socio-cultural subdepartments were united. The department of Socio-Cultural Household Surveys became one of the departments of the newly formed Socio-Cultural Statistics Division. Although it was considered necessary to unite these subdepartments, it was also a severe risk for the POLS programme.
The variation in culture was immense and the success of the POLS programme depended on the successful psychological integration of different personal styles.

The second origin of the POLS programme was the 'mental' integration of these subdepartments. The integration of the instrument, the integrated survey, was considered the best tool for this mental integration. It also became clear that the mental integration was a tool for the instrumental integration of the surveys. The two integration processes are sides of the same coin; they are necessary preconditions for each other.

At the end of the eighties the PAPI questionnaires were replaced by computer assisted personal interviewing techniques (CAPI). At the beginning of the nineties the internal memory of the hand-held computers used for interviewing was still relatively small and did not allow the combined operation of several surveys. Modern CAPI systems (hardware en software like Blaise 3), however, facilitate the combined use of larger surveys. One of the pillars of the TEMPO reorganisation was the awareness that the actual structure of the organisation hindered the fulfilment of user needs and the implementation of new technology. After the integration of the formerly divided subdepartments, these obstacles disappeared. *The third origin of the design of the POLS programme was the need to organise work more efficiently* with respect to the new information technology on offer.

The fourth origin of the POLS programme was a consequence of the changes in culture, started by the TEMPO reorganisation. In 1994 Statistics Netherlands seriously started promoting project based work procedures and prudently started to implement a matrix organisation. For such organisational changes to be accepted widely within the organisation, pilot projects needed to be set up to prove the added value for statistical work of these organisational models. The management of the divisions involved in the POLS programme considered this programme to be an important tool for changing the management structure and the management culture.

Figure 1. The four elements in Business Process Redesign projects



3. The POLS programme as a Business Process Redesign project

The origins of the integration process as described above show a strong similarity with the four items in the literature mostly considered to be the main elements of a Business Process Redesign project (BPR) (see Hammer, 1993). BPR processes are used for the management of fundamental reorganisations of firms, production processes, etc. The similarity of these four origins (content of the surveys, work culture, information technology and the organisation of the work processes) with the main elements of the POLS programme was recognised at the start of the working out of the programme plan. BPR was then chosen as the organising and driving principle for the restructuring programme of the socio-cultural surveys.

A programme as large as POLS cannot be described in isolation from the main business objectives. These objectives, formulated in Statistics Netherlands Business Plan, as well as the objectives of the separate POLS surveys, would have to be operationalised in the programme

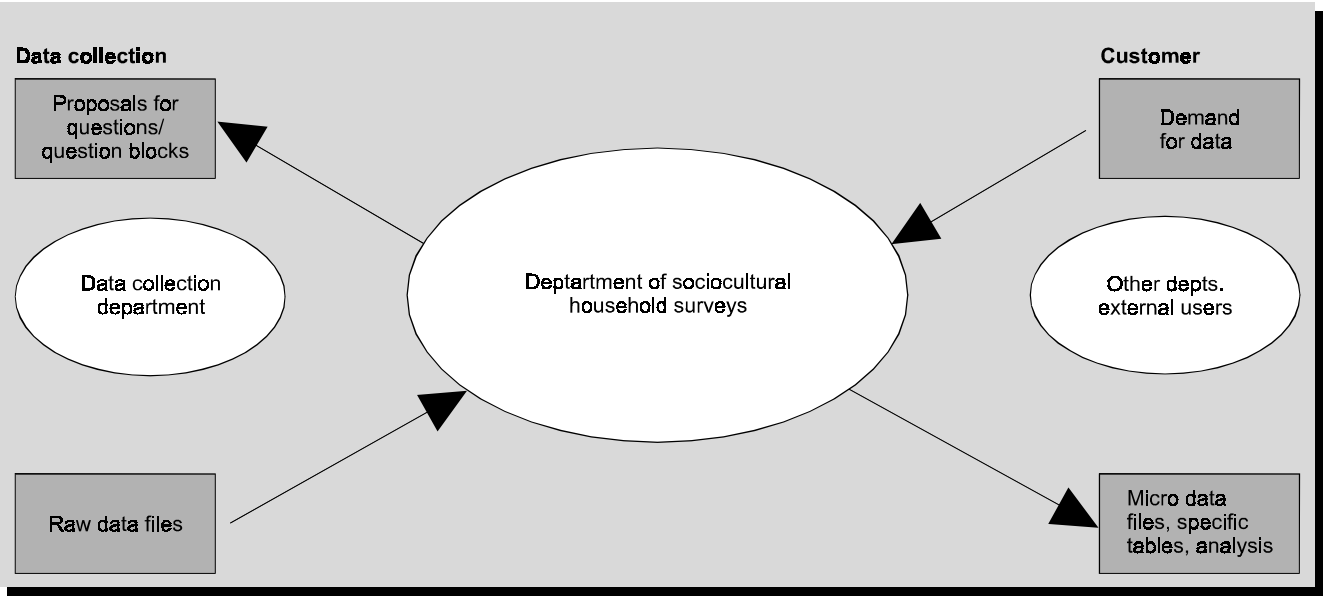
The main objective of the POLS programme was defined as:
the redesign of the separate particular subdepartments with traditional statistical procedures into a modern all-round research unit according to the objectives of TEMPO and using modern integrated survey instruments.

The objectives of the projects are characterised by:

- their relation to the instruments and to the organisation, and
- their relation to one or several of the business objectives.

As several divisions and departments are incorporated in the production chain of statistics, clearly communication is an essential element of the objectives. Not only the method of communication is important, but also the character of the member of staff who has to communicate in the chain of activities.

Figure 2. Chain of processes is survey research at Statistics Netherlands



Therefore learning to work according to project structures is one of the objectives of the programme. The realisation of the instrumental objectives is considered to be an essential condition for the realisation of the organisational objectives. Figure 3 illustrates the relations between the business objectives and the objectives of the programme.
The objectives of the projects are based on the objectives of the programme. Special attention is paid to the relations between the Business Plan and the programme objectives, and between the project objectives and the programme objectives. The objectives define the type and weight of the projects, their relations and their boundaries. Again, this is essential for the planning of the programme, the individual projects, their relations, and their input and output.
In the POLS-BPR programme some 25 projects were defined, involving over 100 employees from fifteen different departments. Ten surveys were integrated into one new design and four subdepartments were united into one department. The preparation of the programme took the first six months of 1995, excluding the first stage of scouting. The implementation officially started in September 1995 and ended formally with the delivery of the instrument to the directors concerned in June 1997. The costs of the former surveys (staff, material, information technology) amounted to ten million guilders on a yearly basis.
The costs of the POLS programme were estimated at 44 man-years and approximately half a million guilders for material expenses.

The directors of the Socio-Cultural Statistics and Data Collection divisions were the principals of the programme. The director of the Research and Development division was also strongly involved with the programme. They all agreed with the programme on the condition that the costs had to be regained in about four years. In the end, the total costs of the programme were below the estimated.
Moreover, the new design increased use by external users formerly not related to Statistics Netherlands. We therefore expect the initial investments to be regained in the agreed period. One might say that Statistics Netherlands has not only gained financially, but also in the area of image and professionalism.

4. A BPR programme in theory, and the POLS programme in detail

4.1. The phases of the restructuring process

Like other processes of change a BPR project grows stepwise, beginning with a simple idea. In general the BPR process can be divided in three main phases: the formalisation of the idea; the definition of the programme; and the implementation of the programme.
In *the formalisation phase* the idea 'grows' and commitment is gained by discussing the concept. The first stage in this phase consists of the preliminary development of the idea in the mind of a small inner circle of employees. The subsequent consultation of internal and external experts and customers is an important step: the idea is then launched in a larger group and the management becomes involved. From that moment, too, opposing movements

Figure 3. Objectives of the POLS programme and Statistics Netherlands Business Plan

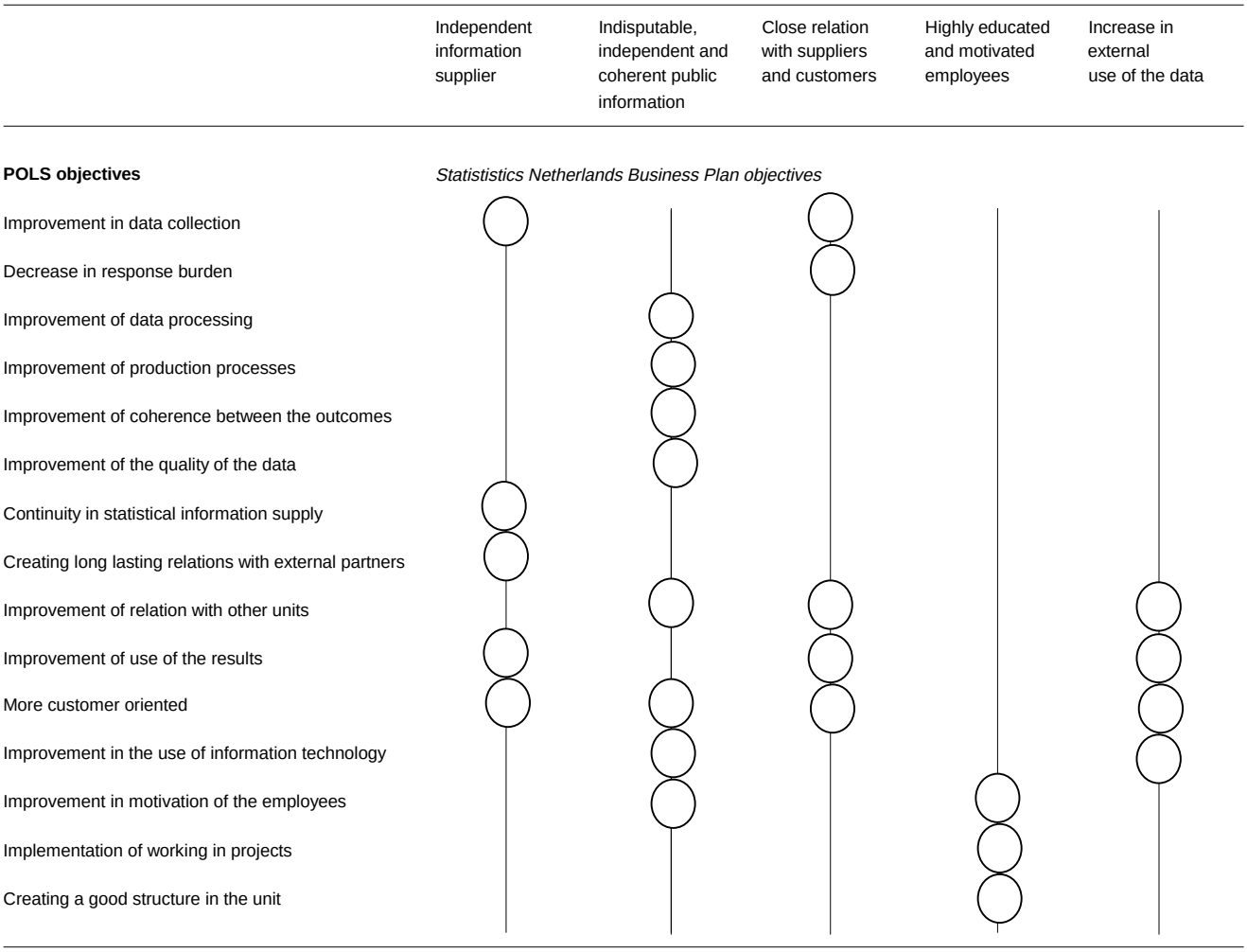
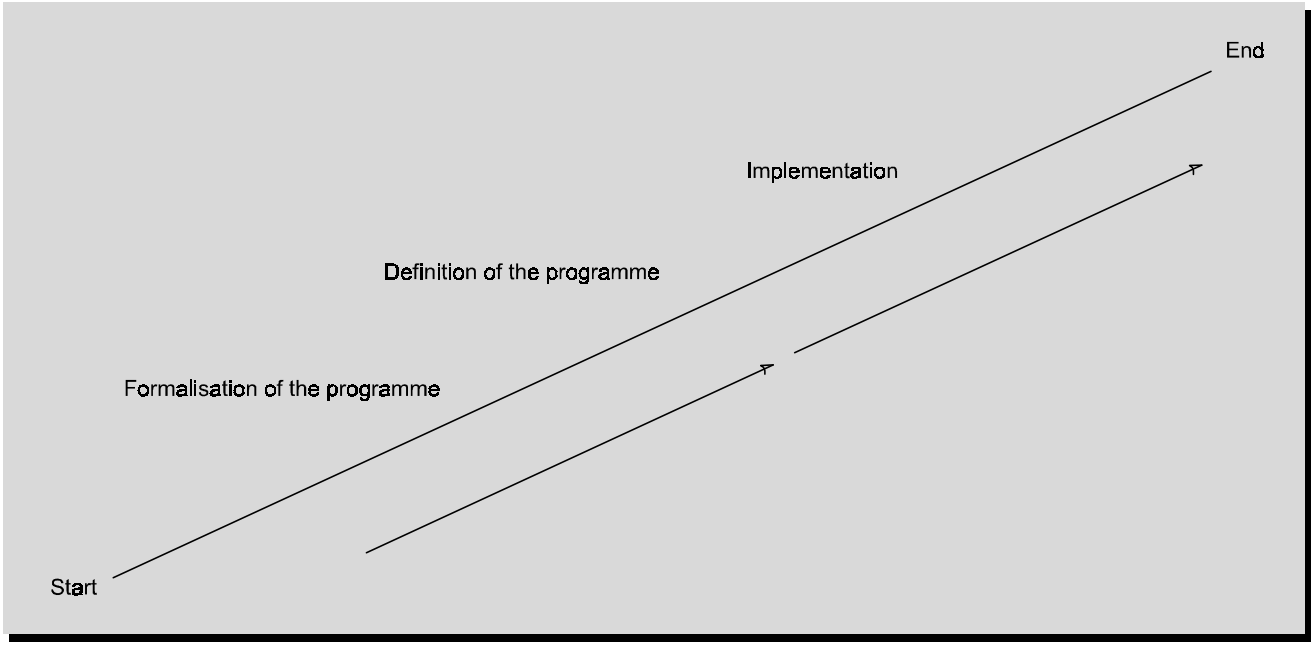


Figure 4. The concept of plateaux in Business Process Redesign projects



might develop. This period can be described as an incubation period. In principle this phase also has a 'commercial' function: the idea has to be sold. To achieve this, one has to search for the arguments in favour of the idea. The possible benefits are essential elements in the discussions. The contours of the project become visible. So an important element in this phase is communication of the idea to management, customers, employees and important external relations.

The translation of the idea into a programme, a set of projects, is the next phase: *the definition phase***.** In this phase the objectives are defined more precisely, the structure of a project organisation, the desired quality of the work and of the outcomes are set out in a programme document. The formulation of quantifiable objectives is an important exercise. These objectives have to be derivatives from general business objectives. This so-called programme phase also defines the management structure of the BPR programme and gives the margins of the changes in the programme that might be necessary during the rest of the programme. In this phase employees are asked to fulfil key functions in the management of the programme or as project managers. Knowledge about project organisations is necessary in this phase.

The discussion with management and potential project managers begins before the official approval of the programme is received and a start is made on the project. The official approval by the principals is the right moment for an official kick-off. *The implementation phase* starts with the kick-off meeting. Project managers are asked to compose their project plans and to indicate the objectives of their projects and the match of these objectives with the programme objectives as formulated by the programme management. The programme management will approve the individual project plans and help the project managers to find adequate members for their project teams.

In large BPR projects the implementation may consist of several steps. Because the exact description of the end product was not possible at the beginning, not the end product but the objectives of the projects are formulated. After each step these objectives are adjusted, and when necessary project plans added or skipped. These moments of rearrangements and revisions can be seen as resting plateaux in the climb up the mountain to the defined set of objectives. On each plateau all the elements are reconsidered.

In the POLS programme five of these plateaux can be identified: (1) at the start, reconsidering the idea (2) the presentation of the first results of some brainstorming at the end of the formalisation phase of the programme; and then three plateaux in the implementation phase (3), the revision of the objectives after preliminary tests and literature studies, (4) the start of the running of the survey parallel to the existing old surveys, and (5) the definition of the values indicating the performance of the organisation and the survey in daily practice.

The programme's project plan gives information on the capacity necessary to work out the plans. The final point of the BPR programme is the transfer of the project result to the principals of the programme. In the POLS programme the principals were the directors of the three most committed divisions. In principle they are responsible for the implementation of the results in the production environment. At their request, in the POLS programme this was also done by the POLS programme management.

4.2 The structure of the project and communication

The characteristics of BPR programmes are described in detail in textbooks on management. They focus on the structure of the project organisation, the programme objectives, on the separate responsibilities of the principals, programme and project managers and project members. The communication in the BPR programme is considered the lubricant: the key to its success. There are two aspects to communication in this respect: the communication network and the quality of the communication.

The communication network is a dimension in the structure of the organisation. It defines how the principals are informed, how project leaders communicate with the programme manager and *vice versa,* how the documents are archived, etc. The POLS programme is relatively complex and large. In order to facilitate a smooth communication the project structure of the POLS programme was kept as simple as possible. The programme is structured horizontally according to the organisational units at Statistics Netherlands. Because three divisions were directly connected with the project, the complete commitment of the directors to the project was the most important. The POLS

processes influence the work programme and the work culture in their divisions. All three of them agreed to be the principals of the POLS programme. Usually the overall management of a BPR programme is done by a steering group. However, in the POLS programme only one and later two officers were made responsible for the overall management. The principals gave these programme managers the responsibility for the realisation of the objectives of the total programme. The programme managers communicate directly with the principals; they meet formally at least every three months, and on an ad hoc basis whenever necessary.
The programme managers lead the group of project managers directly. The project managers are totally responsible for their project. Both the programme managers were responsible for the project managers, although they divided the tasks to for the sake of efficiency. In this structure the communication paths are kept as short as possible. The absence of a steering group facilitates fast decisions on essential elements. However, more emphasis was put on the informal information network of the directors to other directors who were not as involved.
Every two weeks a plenary meeting of programme and project management was held. All project managers had bilateral meetings with the programme management at least every four or five weeks. Although the large number of project managers facilitates a clear division of tasks, the burden on the programme managers to control the relations between the projects is relatively large. To prevent overlap or white spots, some separate - transversal - project groups were installed to co-ordinate the end products of more elementary projects. One project group was made responsible for the development of new ideas and the keeping warm of temporarily irrelevant elements; a second group was responsible for monitoring the expenses (hours, travel costs, costs of meetings, etc.) and a third group monitored the quality of the communication. A small staff of three was added to the programme management, to organise meetings, keep an eye on the budgets and archive the documentation. Alongside this straightforward project organisation, two levels of co-ordination were formalised.

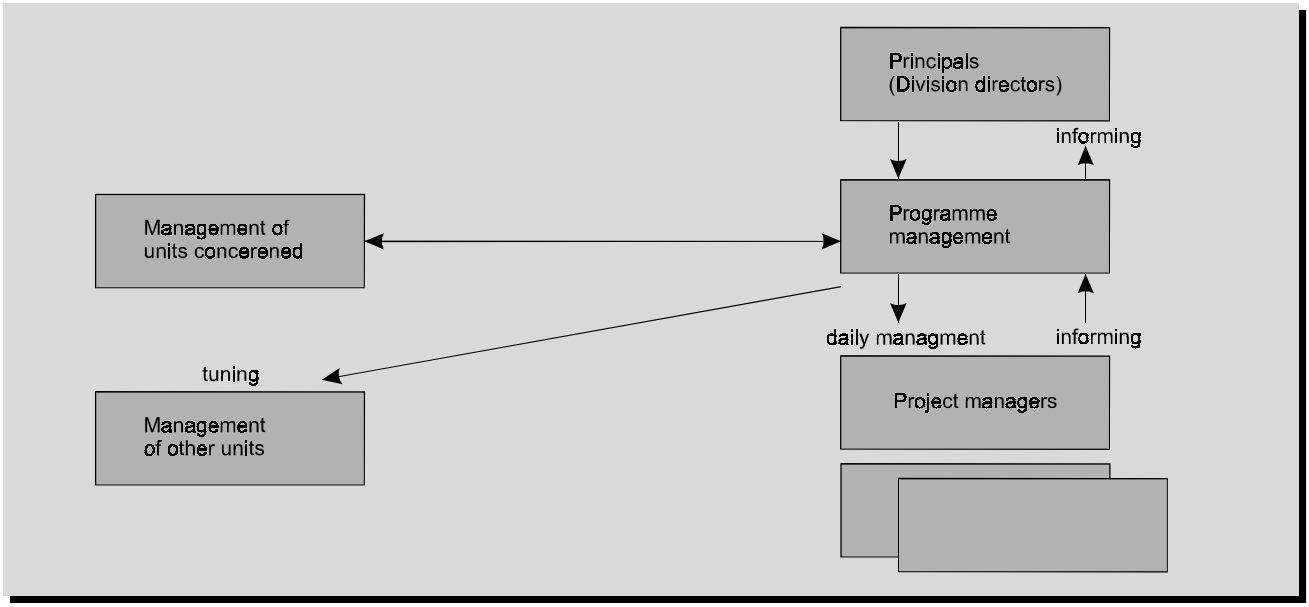
1. a monthly meeting with the heads of units (human resource management and responsible for the regular production processes) most engaged with the realisation of the programme and the implementation.
2. a meeting every two weeks with the three project managers responsible for the key projects on the main route of the programme.

5. The quality of the communication in the POLS programme

5.1. Characteristics of communication in the POLS programme

As stated above, communication is essential for the success of a BPR process. The communication structure as defined in the structure of the project organisation and the agreements reached at regular meetings are the conditions for the well functioning of the communication between the project groups and the staff concerned. However, the quality of the communication is not only conditioned by these aspects. The attitude of the staff involved, as well as the communication with those not involved are just as important. This quality of communication has a direct influence on the image of and commitment to the programme. A very positive commitment at the start might change during the project because of disappointing results or apparent opposition. Therefore one important task of the programme management is to monitor and influence the content and quality of communication. In the POLS programme this was done intensively. The quality and form of the communication may be related to the results of the different phases as described above. This give an insight not only into the way problems were tackled, but also into the main success factors for the separate phases.
As in many BPR processes, newsletters and presentations were used for both the regular external and internal (members of the programme) communication. However, in the POLS programme, a special project group was set up to check on the internal communication. This project group assisted the programme management in monitoring the communication within and between the other project groups and the quality of the meetings. They analysed the interrelations between project members, the quality of participation at meetings and the commitment of the project members to the project. They advised the project managers and reported regularly to the programme management. The quality of the communication with principals and the rest of the organisation was more difficult to steer, while it is these directions of communication in particular that can make or break a project. We distinguish five related qualitative elements considered to characterise the communication in the POLS programme:

Figure 5. The internal organisation of the POLS programme



(a) Commitment to the project as well as the solidarity between the employees.
(b) Level of intellectual dedication with respect to frankness for the innovative elements, and with respect to the content and procedures of the programme.
(c) Freedom or level of mandate at each level of the organisation.
(d) Promotion of the idea, the programme, the process, etc.
(e) Sense of responsibility at each level of the organisation.

For an insight into the proceeding of the POLS programme and the results the opinion of the programme management and some project leaders on these elements are described for each of the three phases.

5.2. Communication in the phase of formalisation

Only a few employees were involved in the start of the project. The idea of survey integration was launched during the first meetings of the potential manager of the new subdepartment with some of the staff. At this stage commitment was only very low. The presentation of the idea to the division director resulted in a very low profile agreement to describe the project in more detail and to start a small project group to scout the feasibility without any responsibility for future work. The five members of this project group had senior experience in the contents of the surveys and had an open mind for innovations. They were selected on their expertise but also on their capacity to built a good team in a very short period. Their product was later used as a basis for the presentations to promote the idea to the whole organisation. Even after three or four months, only those working in the small project group, the manager of the department and the director of the division were intellectually dedicated to the project. There was no further promotion of the idea. This period lasted from September 1994 until January 1995. After consulting external experts in redesign processes on the organisation of the project and customers about what statistical information they wanted, it was decided to launch the idea at an internal Statistics Netherlands meeting in which potential redesign plans of divisions were to be sketched. After this seminar the idea started to grow. The period January 1995 to May 1995 was used to promote the idea internally within the Bureau, to the customers and to international experts (Everaers, 1995a, 1996). After a short period of extensive support in April 1995, an opposition began to make itself felt within Statistics Netherlands. It focused on the high IT ambition of the project, the reorganisation of work in the departments, and the reluctance to stop the existing surveys because of the assumed breaks in series. The intellectual quality of the employees involved was still very high and an intensive high-level methodological debate started, perhaps as a consequence of the opposing powers and their sense of responsibility. Remarkably, at that stage the later principals of the programme had a low level of commitment. They considered the project still embryonic. Their attitude to the later programme manager - giving him and his project the freedom to proceed, to stop or to temporarily cancel the programme - gave a secure feeling. Risks were avoided by all parties involved, the division director only facilitated working hours, the project group could drop the idea and no promises were made. This position took the wind out of the opposition's sails. To sum up, during the phase of formalisation of the idea the *commitment* of those involved was high, that of Statistics Netherlands in general relatively low, the *mandate* was broad and the *intellectual dedication* extremely high, as was the *sense of responsibility*. The *promotion* of the idea was successful.

5.3. Communication in the phase of definition

The translation of an idea to an operational programme of projects puts high pressure on the level of innovative thinking. The culture within Statistics Netherlands was not directed at a redesign process like POLS. Prevailing attitudes did not include giving explicit written instructions for a programme of projects, and there was no experience in facilitating a BPR programme. This was an important reason for the relatively long period it took for the definition phase to become really successful.

In the first instance this phase was badly directed. In fact two culture changes were needed: the writing of the programme document, which was a new experience for Statistics Netherlands; and the requirement of the explicit commitment to the programme within the organisation. As a result the programme fell into a dip for a fairly long period, between April/May 1995 to September 1995. The opposing forces and the insight that other departments connected via the chain of work procedures would have to become more intensively involved slowed down progress. The commitment of the neighbouring departments had to be gained. Each hierarchical layer in these departments had to be convinced of the positive impact of the programme.

The addition to the programme management of an external consultant with senior experience in communication and redesign projects, turned out to be a very fruitful move. It became clear that the translation of an idea into an operational plan requires a specific set of qualities. The subject expertise is no longer the most important element, experience in starting redesign processes, promotion of the idea and the translation into a realistic programme demand high-level communicative skills. The two programme managers succeeded in writing a programme and constituting the project organisation described earlier in this article, in less than six weeks

In the formalisation phase the four elements of the BPR process were repeatedly presented in meetings and lectures. However, the translation into feasible projects was more complex. The focus in the period May 1995 to September 1995, was too narrowly directed towards the instrumental projects. The added value of the external consultant was his proposal to split the organisational and instrumental projects. The structure of the programme and its similarity to the BPR project then became more clear. This was communicated to the departments concerned. The agreement with the principals resulted in their explicit commitment to the project. From that moment they were responsible and had to deal with the *promotion* of the project in their management layer.

Summarising, this relatively short phase was characterised by a decreasing level of *commitment,* a steady high *quality of intellectual dedication*. The *mandate* for the project became more clear and the *sense of responsibility* became higher. There was only limited *promotion* of the project during this period.

5.4. Communication in the phase of implementation

The kick-off meeting for the implementation phase was given much attention in a newsletter and in the in-house magazine. POLS was the only redesign project not initiated by a department in the instrumental divisions, but by a statistical division. The project also received much attention from the later authors of the Statistics Netherlands Master Plan and Business Plan. The start of the selection of project managers and project members resulted in voluntary and uninvited offers from employees. There was an extremely positive commitment within the organisation. The POLS programme was considered to be the first actual and practical translation of the intentions and plans of the TEMPO reorganisation with respect to the structure and content of the work. The motivation and quality of those working on POLS was very high, as they considered this project improved their status. Of course this also led to negative reactions from those not engaged in POLS. The opposition within these groups was countered by distributing more newsletters and the organisation of several seminars informing employees about the project. The implementation phase lasted for about 18 months, from September 1995 to April 1997.

During this phase all kinds of communication had to be used to keep the project successful. There were moments the programme encountered severe problems. Intensive communication with principals and managers of departments was needed. In the second half of the implementation phase of the programme, changes were needed in the programme management. The external consultant was replaced by a manager with senior experience in information technology, as the emphasis during the latter stages moved from the organisation of the redesign processes to the organisation of the work of IT directed project groups. Other matters proceeded more smoothly.

There were at least three moments at which the programme management had to change communication tactics because commitment decreased or specific problems negatively influenced the motivation.
Written agreements had been made with the management of the departments involved about how many hours project staff were allowed for the project, and the management was informed regularly about the actual hours worked. Nevertheless, there were several serious conflicts between the line management and the project organisation. After the first enthusiastic months, the communication with the human resource managers of the departments demanded a lot of attention from the programme management. The hierarchical and project position of the employees was a severe point of struggle on at least four occasions. The burden on motivation, intellectual quality and the innovative attitude of the employees was high; too high for some of the employees and project managers. Communication by the programme management had to be directed to these problem areas. However, the fact that the communication structure of the programme, including the communication with the human resource management and the conflict handling were described in advance and in great detail in the POLS programme plan, eased the formal solution of these problems.
The special project dedicated to the communication turned out to be successful. This project group enabled project leaders and the programme management, when necessary, to intervene in projects or conflicts at an early stage.
The accessibility of the principals was an important point of attention. At some moments the communication with the principals failed and ad hoc procedures had to be used to ensure progress. Mandate was not a problem, although the role in the programme of the principals became more difficult as time went by. New challenges asked for their attention.
The positive commitment of the principals and their active participation was a precondition for the functioning of some of the projects. This commitment was discussed several times during the meetings of the project managers. The positive attitude to this link between management and project had to be positively communicated.
The motivation of the programme management, the principals' skills in facilitating and judging the results, and the enthusiasm of employees in key functions was essential for the success of the programme. This enthusiasm was reflected in an important feature with respect to the work culture in many of the projects in the programme. The method of working of the programme management in this phase was of steering and facilitating along very general lines. There was a clear distinction between the expertise of the programme management with respect to details in the techniques and content, and that of the employees. In the first months of the implementation phase the programme management was actively involved in the work progress, giving advice etc. However, during this phase the projects grow into self-steering groups. When a question arose, the programme management increasingly only had to ask the staff under a project manager how he would solve the problem, to help the problem be solved. During the last six months of the implementation phase the programme organisation can be characterised as a learning organisation. The new POLS survey was officially presented to the principals during a seminar on the results of the POLS programme in June 1997 in a meeting attended by about 100 employees. To sum up the characteristics of communication during this period, *commitment* changed almost weekly, related to progress or no progress in certain projects. The *intellectual level* and the innovative attitude decreased, also because the work became more routine. The *mandate* and freedom, however, and the *sense of responsibility* remained high. Because of the large number of employees working on it, the programme was a *promotion* instrument in its own right

Figure 6. The quality of communication by five characteristics of communication

Aspects of communication	Phases in the programme		
	Formal-isation	Definition phase	Implemen-tation
Commitment to the project	high	decreasing	variable
Intellectual dedication	extremely high	steady	decreasing
Freedom, level of mandate	high	formalised	high
Promotion	successful	limited	intensive
Sense of responsibility	moderate	increasing	high

6. Conclusion

POLS is the name of the new integrated quality of life survey of Statistics Netherlands. It was also the name of the set of projects needed to develop this survey: the POLS programme. The restructuring of the former surveys in this quality of life domain was organised along the lines of a Business Programme Redesign (BPR). An essential element in BPR processes is a restructuring approach which incorporates plateaux: moments of rest to refine or redefine the objectives of the programme. At the start of the programme it was not possible to describe exactly the final results of the programme, these could only be estimated as a consequence of the complexity of the issue under consideration. BPR implies a fundamental stepwise restructuring of all the aspects of the production process, the information technology, the content of the surveys, the work culture and the organisation. The POLS programme consisted of five stages, each characterised by a specific set of projects and work to be done.
The lubricant in a BPR programme the size of POLS (involving over 100 members of staff and lasting for an overall period of more than two years) is the structure and the quality of the communication. The commitment of the principals and staff is prerequisite for success, together with a mandate of the programme management and a sense of responsibility and expertise in all layers of the organisation involved.
The preparation of the BPR programme, and especially the design of the communication structure turned out to be an essential key factor for the successful completion of the programme. The prudent and no risk commitment in the early stages of the project prevented an early attack by opposing forces (there was apparently nothing to attack), and the intensive promotion of the idea in the second stage generated the commitment needed. These were essential conditions for the proceeding of the project. The synergetic combination in the

programme management of subject expertise, survey organisation skills and specialist knowledge of communication within a restructuring process was a very successful one. Lastly, the fact that the programme was organised immediately after the TEMPO reorganisation, causing the enthusiastic involvement of the staff concerned, might be the main factor of success of this project.

References

Everaers, P.C.J., 1996, The Integration of Household Surveys. In: *The Future of Social Statistics. Eurostat, Proceedings of the Mondorf seminar 1995*, (Luxembourg: Eurostat)

Hammer, M., 1993, *Reengineering the Corporation: a Manifesto for Business Revolution* (London, Nicholas Brealey)

Van Bochove, C.J., 1996, From assembly line to electronic highway junction: a twin-track transformation of the statistical process, In: *Netherlands Official Statistics*, vol. 11, summer 1996, pp. 5-36

Van Tuinen, H., 1995, Main Approaches to Social Statistics. In: *Statistical Journal Economic Commission for Europe*, Vol. 12, pp. 379-39

Notes

1) Pieter Everaers and Hans Wajon were the managers of the POLS programme. In a later stage Hans Wajon was replaced by John Dieteren of the Division of Research and Development. Pieter Everaers was department manager of the department of Socio-Cultural Household Surveys in the period 1994-1997; he is currently working as a consultant on new developments in Social Statistics at Eurostat, the statistical office of the European Community. Hans Wajon is managing consultant of FAIR Information Services B.V. and is currently working as interim manager for the Municipality of Utrecht.

Index