

## **PAPI TO CAPI: THE OPCS EXPERIENCE**

*Jean Martin*

*Office of Population Censuses and Surveys, London, UK*

### **1. Background**

In Great Britain and in many other European countries, the main method of data collection for social surveys carried out for government departments is face-to-face interviewing. Unlike in the US, telephone interviewing is not widely used, owing mainly to a lower rate of telephone ownership and sampling difficulties. So, with the advent of computer assisted interviewing (CAI), the problem for many European survey organisations is one of moving direct from paper and pencil interviewing (PAPI) to computer assisted personal interviewing (CAPI) without necessarily any experience of computer assisted telephone interviewing (CATI). This contrasts with the US experience where CATI has been used by most survey organisations for many years and CAPI is a relative newcomer on the scene.

In Europe the pioneers in CAPI development were large government survey organisations - the Netherlands Central Bureau of Statistics, Statistics Sweden and the Office of Population Censuses and Surveys in Britain, for example. The high initial costs of development and the costs of buying laptop computers for interviewers restricted development to organisations carrying out large continuous surveys for customers prepared to set the initial costs against anticipated savings in later years.

In Britain, the Social Survey Division (SSD) of OPCS took the lead in CAPI development, starting in 1987. The initial work at OPCS was funded by the Employment Department, customer for Britain's largest government household survey by far, the Labour Force Survey (LFS). Successful development work led to a change to CAI, using a mixture of CAPI and CATI, in 1990.

At the outset of CAI development SSD had to select a software package which would in the first instance meet the requirements of the LFS, but would also be likely to prove suitable for other surveys in the future. Blaise was selected at an early stage because of its integration, flexibility, reliability and user-orientation.

From the early years it was apparent that a move to CAPI would have major organisational implications. Indeed in order to justify such a move for the LFS, a business case had to be made to show that the initial costs of development and equipment would be recovered by savings in running the survey using CAI and earlier availability of results. On the LFS the pay-back period turned out to be only two years. The cost savings came mainly from the elimination of the data keying, coding and computer edit stages that had been necessary with paper questionnaires. This of course implies a reduction in the number of staff required for the survey and thus affected the organisation as a whole. However, the move of the LFS to CAI coincided with a period of considerable growth in the amount of survey work being carried out so there was no problem in absorbing the staff released by the changes. Moreover, because of the general fluctuations in the volume of survey work, a proportion of temporary staff have always been used to cover peaks, for example students working over the summer vacation, new graduates doing temporary clerical work before finding permanent jobs. This provides the organisation with a degree of flexibility in staffing levels.

After the successful transfer of the LFS to CAI, attention turned to other surveys. Next in line were two new continuous surveys which were implemented in CAPI from the outset. Both were won in competitive tenders against the private sector and had to justify the costs of CAPI development over the life of the initial contracts which were three to four years. For the second CAPI survey, a complex income survey, the customer department commissioned a comparative evaluation of three possible CAPI packages after which Blaise was selected. (This survey is carried out jointly with another survey organisation, SCPR, who is responsible for the Blaise programming.)

Shortly after, consideration started being given to the transfer of other continuous surveys from PAPI to CAPI, for example the Family Expenditure Survey and the General Household Survey. Both are now scheduled to move to CAPI in April 1994. Thus far a business case had been made separately for each survey in turn, although it was apparent that the development costs were reducing as each survey built on its predecessors. Thus the economics of CAPI have been changing over time, reflecting the balance of survey specific costs to organisational costs of changing to CAPI.

By April 1994 it is expected that CAPI will be used for the majority of surveys — as new surveys are considered and costed, the assumption is that CAPI will be used unless there are particular reasons why PAPI would be more appropriate. Indeed recently SSD has been involved in a number of competitive tenders where the comparative costs of using CAPI and PAPI have been requested. In most cases, even allowing for development costs, CAPI is cheaper. The only continuous household survey not moving to CAPI for the time being is the National Travel Survey, which, owing to its very complex hierarchical structure, cannot be implemented in Blaise until version 3 is available.

The fact that greater automation of survey processing would have implications for organisations carrying out surveys was apparent from the outset. An early paper from the Netherlands CBS (Brakenhoff et al., 1987) described the main stages of traditional survey processing based on paper questionnaires and how this might change under an integrated computer assisted system including CAPI, CATI or CAPI, as required. The authors discuss problems for the organisation in making such a change, pointing out that survey organisations are generally structured around the stages of a PAPI survey and that there may be resistance to restructuring the organisation in order to implement technological innovation. Bateson and Hunter (1991) discuss some of the changes which moving to CAPI entails, concentrating on the role of respondents and interviewers but also describing how the job of the questionnaire designer will change.

The introduction of widespread automation at the Netherlands CBS clearly had a major impact on the whole organisation, not just the survey area; a number of papers describing different aspects of the changes have been

published in recent years, particularly about the organisational aspects of the introduction of integrated survey processing using Blaise (ég. de Heer, 1991; Bethlehem and Keller, 1991; Woolf, 1991).

It is apparent from a search of the literature that what has been said about implications for organisations of moving from PAPI to CAPI is very dependent on the functions of the organisation concerned, in particular on whether the organisation is solely a survey organisation or whether surveys are carried out within a wider statistical organisation. (See, for example, Nicholls and Matchett, 1992 for a discussion of some of the issues at the US Bureau of the Census and Jamieson et al., 1992 about Statistics Canada.)

## **2. Surveys at OPCS**

Before discussing in more detail the implications of moving from PAPI to CAPI for the organisation, some background about the surveys carried by the Social Survey Division and how the Division fits into the wider department, OPCS, is provided. The situation is very different from those at other government statistical offices such as the Bureau of the Census, Statistics Canada or Netherlands CBS, for example. Britain does not have a centralised government statistical office; OPCS is responsible for some government statistical functions but it also has another role: it administers the registration of births, marriages and deaths in England and Wales. It collects and analyses population, medical and some social statistics, and carries out the Census of Population. However, it has no responsibility for statistics in other areas. The Social Survey Division is the official government survey research agency, carrying out social, rather than business or agricultural, surveys. Most of its surveys involve face-to-face interviewing in households, but other modes of data collection are also used and information may be collected from other than household settings, for example, schools, prisons, residential homes. One survey, the International Passenger Survey (IPS), has its own field force stationed in major sea and airports.

SSD comprises about 230 headquarters staff (of whom around 55 are professional survey researchers) and some 1000 interviewers; to a large extent, it has all the staff required to carry out all stages of survey work. The

main exceptions are around 40 professional computing staff, who form a dedicated branch within the IT Division, and some support services such as data prep, data communications, document production, finance and personnel functions which are provided centrally for all Divisions of OPCS. Although SSD uses the office wide data communications network, its computing arrangements are in general different and separate from those of the rest of the office. SSD has its own Microvax for survey processing and is increasingly moving to devolved PC-based systems.

Currently about 30 surveys a year are carried out, 9 on a continuous basis. All surveys are commissioned by government departments and other public sector bodies and are charged on a full cost recovery basis. SSD has virtually no other source of funds, and is expected to pay other parts of OPCS for any services it requires and for overhead costs such as accommodation charges. In this respect it functions very like an independent non-profit making survey organisation. In the past SSD was commissioned directly to carry out many surveys; now increasingly it has to bid against other survey organisations through competitive tendering. Even surveys that SSD has been carrying out for many years will be subject to this process in future. It is therefore vital that SSD is able to offer the best value for money possible in terms of very high quality and timely survey research within acceptable budgets.

Against this background SSD has decided that moving the majority of its surveys to CAI will provide major benefits in helping maintain its position as the lead organisation carrying out large complex government surveys to high technical standards. It is apparent the main impact of a change in the mode of survey data collection will be on SSD rather than on OPCS as a whole.

### **3. Coordination of CASM development**

Early CAI development was specifically for the LFS and was handled by the team working on that survey. It thus developed independently of other

PAPI surveys. So although there was some awareness of the implications of CAPI for the organisation as a whole - particularly since several members of staff had heard Wouter Keller (from Netherlands CBS) speak about the issues on a number of occasions - it was some time before SSD really had to face up to the organisational implications of a large-scale move of all its major surveys from PAPI to CAPI.

By 1991 the need for coordinated development was recognised and one of the SSD Directors (JM) was given overall responsibility for coordinating the development of computer assisted survey methods (CASM) in SSD, and in particular for the move of continuous surveys from PAPI to CAPI. A year later a CASM Technical Coordinator was designated, as the need for coordination in such a rapidly expanding area grew. Initially, the two main problems were financial and technical. The first was how to raise enough money to provide every interviewer with a laptop and modem, and to have sufficient file servers and other equipment to handle incoming data, case management, disc duplication etc. As mentioned earlier, SSD is largely customer funded, with the exception of some core funding of basic methodological research and access to a small portion of a central OPCS budget for purchasing IT equipment which is used mainly for standard desktop PCs, printers and equipment for administrative, rather than survey specific, systems.

The technical problems included how to develop systems to handle a range of different CAI surveys, to transfer experience from one survey to another, and to start instituting organisational standards for common elements of CAPI surveys rather than letting each survey team develop methods and procedures independently.

#### **4. Financing the move to CAPI**

##### *4.1 Purchasing and charging for laptops*

In principle the purchase of a large number of laptop computers for the

LFS might have helped equip interviewers with machines to use on other surveys, but in practice this did not happen. One reason was that the sample size of the LFS was increased four-fold (which would not have been possible to the same timetable without the use of CAI), allowing an unclustered sample design and the establishment of a completely separate field force for this survey alone. A second reason was that development work on the next survey in line, a complicated income survey with a very long interview, showed that the laptop computer being used on the LFS, which had a basic 8086 processor, did not run fast enough for the new survey; a minimum of a 286 processor was required and in the event, as hardware prices fell, a 386 processor was in fact used. So the organisation was starting from scratch in terms of equipping the general field force of around 400 interviewers who carry out all face-to-face household interviews other than on the LFS.

Government departments with large research budgets can (with advance warning) handle annual fluctuations in the costs of a particular survey much more easily than those with smaller budgets. Thus the Employment Department could cope with buying all the laptops required for the LFS in Year 1 and replacing them all in Year 4. But customers for smaller surveys could not be expected to finance all the laptops needed to carry out their survey alone - on this basis it would be difficult to make a business case for using CAPI. But of course interviewers can use laptops for more than one survey, so each customer could be asked to buy a proportion of the laptops needed for a number of surveys. By this means, and with some help from central OPCS funds, it was possible to supply each of the interviewers with a laptop and modem (and have sufficient extras for spares, training, development etc).

SSD is now moving to a new system of charging surveys for the use of laptops rather than for their purchase. Assuming laptops have a life of three years, one third of the replacement cost each year needs to be raised. Dividing that cost, plus an allowance for software, support and maintenance, by the anticipated number of CAPI interviews in a year results in an overhead charge per interview for the use and replacement of the equipment.

This provides an equitable means of sharing costs over surveys proportional to the volume of interviewing and thus does not penalise small surveys.

#### *4.2 Financing other equipment*

The above arrangements apply to the equipment required by interviewers - laptops, modems, discs. But the move to CAPI using Blaise implies the use of PCs for at least some other parts of the survey process. For PAPI surveys SSD uses a Microvax for survey processing; most staff have desktop PCs which can access the Microvax and can also be used for wordprocessing, spreadsheets etc. OPCS as a whole had developed an IT strategy in 1987 which had resulted in the purchase of a large mainframe computer, primarily to handle the 1991 Census and the regular statistical work of the Office connected with population and medical statistics. SSD had to consider whether this was the most appropriate solution for survey processing, which was already carried out separately from other computing activities in OPCS, and a review concluded that it would be most economic for SSD to develop its own IT strategy, taking into account the implications of the move of surveys to CAI.

Following this comprehensive review, a strategy for survey computing was adopted which involves moving over the next few years to PC-based solutions, making use of the OPCS network which consists of a WAN connecting the OPCS sites in different parts of the country and LANs at each site. At the time it was envisaged that these would be run under UNIX but now it seems likely that a Windows approach will be more suitable, particularly since SPSS Windows has now been released, SPSS being the main tabulation and analysis package used in SSD.

The immediate problem has been how to fund the purchase of file servers and power PCs to handle data being transmitted from the field and for subsequent processing. Customers for the two new CAPI surveys were prepared to finance the equipment required for their surveys. SSD has managed to cope so far by obtaining some money from the central OPCS



budget and by getting each of a number of surveys to pay for a share of the equipment required; often one PC or file server can handle a number of surveys. A charge will be introduced for using such machines in future to cover support, maintenance and replacement, as is the case for the use of the Microvax.

## **5. From one to many — sharing expertise and experience**

SSD has been concerned both to avoid each new survey 'reinventing the wheel' and to try and establish standard methods and systems which would be used on each survey. The period of development of CAI on the LFS had resulted in a pool of staff with various useful skills and experience. Some of these trained their replacements and were then moved to work on the new CAPI surveys. To aid coordination, a CASM domain was set up on the OPCS network allowing all working on the new CAI surveys to share information easily. Regular CASM coordination meetings have been held at which staff working on the different surveys either using or planning to move to CAI report progress and plans. A weekly lunchtime CASM workshop takes place mainly for people actively involved on the CAI surveys. Participants can bring their problems (with Blaise or other aspects of CAI) or can give an account of an interesting aspect of their work to share with others. A CASM Newsletter has been started to keep everyone in SSD informed about developments. It is planned to build a library of blocks of Blaise code which are likely to be of use on a variety of surveys, such as the household grid information, employment status and job details; common elements of CAI surveys already use common code.

## **6. Changes in survey tasks**

The changes in the tasks required to carry out a survey as a consequence of moving from PAPI to CAPI are by now well known; however, they are often described in rather abstract terms which do not bring out the changes likely to occur in people's jobs. The summary provided here describes these changes in the OPCS context.

### *6.1 Planning and costing*

As in most survey organisations, senior staff are responsible for writing proposals, drawing up timetables and costing surveys; for large complex surveys these are demanding tasks which draw on many years of survey experience rather than on following set rules. One problem already encountered is that of educating customers about the implications of changing to CAPI, particularly with respect to survey timetables. At least on the more complex surveys, a longer period before starting fieldwork is required; writing edit checks has to be finished before going into the field rather than being written during or after fieldwork, and the survey instrument needs more thorough testing than a paper questionnaire before interviewing starts. The savings in time come after the fieldwork so the total length of the survey process is shorter. SSD is trying to educate its customers to specify when they want results rather than when they want fieldwork to take place, or, if the timing of the fieldwork is important, to allow sufficient time for development and testing.

Since one of the main aims in changing to CAI is to reduce costs, by eliminating some survey operations and reducing others, it is obvious that the change will have implications for how surveys are costed. The difficulty comes with tasks that are done in different ways and by different groups of people, and with costs that do not arise at all under PAPI. It will take time to build up the same level of expertise in costing that SSD has relied on over the years for PAPI surveys.

### *6.2 Writing survey instruments*

SSD uses Blaise for all its CAI work. Blaise was designed to be used by survey researchers rather than computing professionals, and SSD has viewed the preparation of questionnaires in Blaise as primarily a research task in the same way as the design of paper questionnaires is viewed. Although simple survey instruments could be written in Blaise by research staff with basic computing skills, (all are expected to use SPSS for survey analysis for example), it became apparent, as more complex instruments were tackled,

that professional training in programming was an advantage and in some instances essential. At present the aim is to operate on the basis of small interdisciplinary teams, typically one researcher and one computing professional, using their complementary skills to design the survey instrument. There is also close involvement with a member of the Field Department to provide an input about what will work best from the interviewers' point of view, and with editing staff for help with specifying edit checks to be incorporated in the interview.

This has meant changes to the jobs of both the research and computing staff. With PAPI the researchers design questionnaires with little input from computing staff while the latter are responsible for setting up the database and writing edit programs later in the survey process. With CAPI the tasks are integrated: not only are most of the edit checks incorporated into the survey instrument and carried out during the interview, the specification of the instrument is used to derive the meta-data required to describe the data for the analysis stage.

Initially those using Blaise learnt by a mixture of reference to the documentation and trial and error, supplemented by asking CBS staff for help and advice for the more serious problems. Subsequently a few staff attended Blaise courses run under the Training of European Statisticians scheme which were taught by the CBS staff involved in Blaise development. More recently SSD has used courses run on a commercial basis by SIA, the British suppliers of Blaise; when numbers justify, SIA runs in-house courses for OPCS staff.

### *6.3 Interviewer training*

On the first CAPI surveys, the cost of training interviewers to use the new method of interviewing was borne by each survey. Typically two days training was required to cover both CAPI and the specific survey. Costs started reducing somewhat as surveys were able to use some interviewers who were already trained in CAPI and therefore needed only survey-specific training, but this meant having two different versions of the training course. However, the training was taking place against a background of interviewers

all having undergone a basic 3 day training course in interviewing techniques and the use of paper questionnaires. From next year the basic training course will change to cover CAPI; this is likely to require 4 rather than 3 days. But survey specific training can probably reduce to one day on all but the most complex surveys - ie to the same level as on PAPI surveys. Thus CAPI is expected to result in a slight net increase in training costs as basic training will require an extra day compared with the situation under PAPI.

#### *6.4 Case management*

Unlike most survey organisations, SSD did not have a sophisticated computerised system for managing PAPI surveys (with the exception of the LFS). Moreover, at the time when the early CAPI development was taking place on the LFS, the facilities provided within Blaise were not extensive (although they have since been enhanced) so a system was developed for the LFS, written in Clipper. The starting point is the sample of addresses selected for the survey. The system sets up skeleton interviews for each case in Blaise, together with an administration block for interviewers to enter details of the progress of the case and its final outcome. The information for each case in an interviewer's assignment is put onto a disc sent out to each interviewer. When completed, interviews are transmitted by modem back to the office and the system strips off the outcome information to provide information for field management. Since the LFS is a panel survey the system also extracts information required to feed forward to the next wave of interviewing. There is also a facility to sort cases into the correct month of allocation even when cases for different months are transmitted together. Facilities at the laptop end check that all required information is complete and correct before transmission can take place.

This system has been enhanced for use on the newer CAPI surveys to allow for interviewers working on more than one survey at the same time. At present cases assigned to interviewers are sent out on discs but SSD has recently acquired the LIPS-SPIL system from CBS which allows two-way data transmission and provides a range of case management facilities. This

is currently being customised for SSD use and the plan is to implement it from next January.

A spin-off from all the above work is that it became apparent that it would be easy to develop a computerised case management system for PAPI surveys in Blaise. As paper questionnaires are returned to the office, the booking-in clerks log the return of each case, its outcome and any other relevant details into a Blaise instrument from which regular field reports can be derived as from the CAPI system. This will ease the transition of continuous PAPI surveys to CAPI and provide a means for integrating the remaining paper components for CAPI surveys which still require, for example, paper diaries to be completed by respondents.

#### *6.5 Office coding and editing*

On some CAPI surveys, eg the LFS, the change to CAI has completely eliminated office coding and editing, both taking place either during the interview or, in the case of coding occupation and industry, by the interviewers at home before transmitting completed cases. But on other CAPI surveys a post-interview office stage is still required. For example, on the complex income survey mentioned above, the interview includes many 'soft' checks which interviewers can override. Such checks are run again in the office and further editing may take place. In addition, other checks are carried out which were considered too complex to include in the interview and clerical imputation of some missing values takes place. It is likely that specialised coding carried out on some surveys (eg of diseases or nutrition information) will continue to require an office stage (probably using computer assisted coding) after the surveys concerned move to CAPI.

#### *6.6 Survey computing*

With PAPI the main job of survey programmers is to write programs to set up survey databases, describing the form and structure of the data, to carry out edit checks and to create derived variables. On most PAPI surveys this stage of the work is carried out in SIR. From SIR the data can be passed to SPSS (or SAS) for analysis, which is carried out by the research staff.

On CAPI surveys most of the editing is carried out using Blaise, either during the interview or later in the office. The exception is checks and other operations carried out on the whole data set rather than on individual cases - looking at distributions to check for outliers and some statistical imputations, for example. Other checks need to wait until imputations for missing values have been carried out.

New computer facilities are required to manage CAI surveys which can be referred to broadly as data handling systems. Case management facilities are required both for use by interviewers on their laptops and by field staff for monitoring and supervision. Facilities are also required to check the integrity of data, to check the structure of complex data sets etc. Many of these facilities are required by the organisation as a whole, with interviewers working on more than one CAI survey at a time, but some are survey specific. The initial idea was to develop a system for a 'typical' survey, which could then be adapted for surveys with special requirements. In practice we have not encountered anything which could be considered a typical survey and prefer to think in terms of a core system which is customised for each survey to deal with its particular requirements.

## **7. Organisational change**

Now that the main technical hurdles have been overcome and it is apparent that, although the second CAPI survey only went live in October 1992, by April 1994 CAPI will be the main mode of interviewing for most surveys carried out in OPCS, attention has turned to how to manage what is probably the most fundamental change in survey methods in the 50 year history of survey research carried out by SSD.

### *7.1 Review of organisational structure*

A fundamental review of the structure of SSD and the staff requirements for the future was carried out earlier in the year. This started with the identification of areas the review should cover (including some which had

nothing to do with CAI but which could conveniently be included in the same process). A review team, comprising the Director of SSD and the Assistant Director responsible for CASM coordination, consulted widely both within SSD and with other parts of OPCS which provide services for SSD, to get ideas about how to cope with changes anticipated as a consequence of the move to CAI, what changes should be made to the structure of SSD, and views about suggestions for both as they arose.

It is important to recognise that, even if no changes in structure were envisaged, SSD would still be faced with the problem of having to manage a reduction in the number of staff required to carry out surveys in the future; the cost savings which CAPI brings about mean that jobs will disappear. Since the rate of increase in survey work seen in recent years is not expected to continue, a net reduction in the staff of SSD must be anticipated. The majority of the proposals have now been accepted in principle; the next step is working out detailed implementation plans and managing the changes carefully over the coming months.

### *7.2 Organisation under PAPI*

Clearly at the detailed level, the organisational implications of moving from PAPI to CAPI depend on how the organisation was structured to carry out PAPI surveys and thus each survey organisation will face somewhat different problems. But there are sufficient similarities in the way in which survey organisations are structured, at least in Britain, for the experiences of one to be relevant to others.

In OPCS at present the following areas of survey work are carried out by distinct parts of the organisation:

#### *Within SSD:*

- Research (ie survey and questionnaire design, project management, analysis, reporting)
- Statistical advice on sampling, estimation, analysis etc.

- Sample implementation (most samples being drawn in the office rather than in the field)
- Interviewing and field management (allocating and despatching work to interviewers, progress monitoring, booking-in, interviewer recruitment and training)
- Coding and editing

*Dedicated branch in IT Division:*

- Survey computing (data description, database design, programming edit checks and derived variables)

*Shared service:*

- Data preparation
- Document production

Each survey has a Project Manager (a member of the research staff) who liaises with the customer about the requirements of the survey, is responsible for the survey design, the sample design (with assistance from a sampling expert if necessary), questionnaire design, overseeing the survey through all of its stages - briefing interviewers, answering field queries, agreeing coding and editing instructions, specifying derived variables and checking the data prior to analysis - and in general liaising with support staff carrying out other parts of the survey process, and finally (on most surveys) for carrying out the analysis and writing an interpretative report. This role clearly requires a considerable amount of subject matter knowledge as well as extensive knowledge of survey methods.

### *7.3 Changes with the move to CAI*

Much of the literature on the effects of moving to CAPI concentrates on the changes in the roles of the interviewers and respondents, which are clearly of prime importance. However, changes in the nature of their roles do not have major implications for the structure of a survey organisation. Far



more radical is the huge reduction in the demand for the clerical staff carrying out booking-in, correction of edit check failures etc. and for data prep and document production services. The other major changes are in the nature of the work carried out by the computing and research staff and, to a lesser extent, the field managers.

If the work of certain sections of the organisation is expected to decline considerably, we need to ask whether a smaller section carrying out the same function is viable. This in turn will be affected by whether the level of work is roughly constant or whether large fluctuations are expected; fluctuations on a small base are much more difficult to manage than on a large base.

#### *7.4 Coding and editing*

With PAPI one section is responsible for coding verbatim text answers, scrutinising paper questionnaires prior to keying, dealing with interviewer queries and comments written on the questionnaire, and correcting edit failures with reference to the paper questionnaires after the keyed data had been run through the edit program. Much of this work will disappear with CAPI, but not all; some surveys will still require a post-interview edit stage as described earlier. However, the volume of work is unlikely to be sufficient to warrant the continuation of a separate coding and editing section so a number of other options were reviewed.

One option was to combine the remaining work with that of the survey computing section with whom there has traditionally been close contact over the editing stages. At present the advice of the coding staff is a valuable input to the design of survey instruments for all CAI surveys, but particularly those with a post interview stage. However, with no post interview stage on most surveys, it will be difficult to maintain such expertise. Moreover, the two groups of staff are very different: programmers on the one hand and clerical staff and their supervisors on the other, so on balance combining these two sections does not appear to offer any significant advantages.

A second option was to attach the coding staff for each survey to the research team responsible for that survey, which in many cases includes a number of administrative support staff. This would work for continuous surveys which continue to have a post interview stage, but it is not clear how it could be managed for ad hoc surveys where there may be intensive requirement for a team of coding/editing staff for a short period of time. A roving team moving from one project to another seems like a management nightmare; a small centralised team would exacerbate the problems of peaks and troughs. And even on continuous surveys, having direct management by the survey team rather than any central management is likely to make it more difficult to move staff from one survey to another or have supervisors responsible for several surveys so this solution is not favoured.

The option which was eventually agreed involves merging the remaining coding and editing staff with the field department. Currently there is a distinction in the field department between those responsible for the interviewers — allocating and despatching work, progress chasing, authorising payment etc. — and those responsible for particular surveys, who advise the research staff on field issues, organise interviewer briefings, deal with interviewer queries, oversee response etc. It will be possible to extend the latter role to overseeing the survey through any further office coding or editing stages. Increasingly on CAPI surveys interviewers are carrying out coding in the field, so at the moment responsibility for coding is divided, depending on whether interviewers or office coders are carrying it out; this solution places responsibility for coding in one department whether it was carried out by interviewers or headquarters staff. With CAPI dealing with edit checks is mainly carried out during the interview, and thus clearly the responsibility of the field department. Since this department is responsible for the quality of the work carried out by the interviewers, having to deal with any problems in the data in the office during a post interview stage as well as in the field is a legitimate concern. If responsibility for handling any paper documents were located here as well, one team could be responsible for seeing the survey right through to the analysis stage, whatever mode of data collection were used.

Apart from any organisational changes, the fact remains that fewer staff will be required in total. However, owing to the fluctuating volume of survey work and past difficulties in recruitment, many of the present clerical staff throughout SSD are employed on a temporary basis and there is quite high turnover (due to promotions as well as resignations), so it is likely to prove possible to reduce numbers by natural wastage and not re-engaging temporary staff.

### *7.5 Survey computing*

A reduction in this area of work as a result of CAI is not yet apparent, although the nature of the work is changing rapidly. The main effect of moving to CAI is a large reduction in the type of work formerly carried out by survey programmers. What has so far replaced this type of work is that required to develop data handling facilities for CAI surveys. Thus we have not yet seen any reduction in the total amount of survey computing work required because the resources released from routine edit programming have been used to develop these facilities. However, we envisage the level of new developments to start decreasing although there will always be a need to customise systems for each new survey and there will still be involvement in programming each new survey instrument in Blaise.

The work of the computing staff will be much more integrated with that of other parts of the organisation than was required for PAPI surveys. There is an initial input passing the sampling information to the case management system to set up each new survey, customising the case management and data handling arrangements for the interviewers and field department, setting up any post interview stage, checking the integrity of the data and passing files from Blaise to other software for imputation, adding derived variables, grossing or weighting factors or anything else which is required prior to analysis. This is in marked contrast to the sequential processing arrangements for PAPI surveys where the computing stage starts after the data are keyed and ends when the files are ready for analysis.

Another factor which needs to be considered is what software to use for programming derived variables and adding any weighting or grossing fac-

tors. As mentioned earlier, for PAPI surveys SIR is used to carry out the editing and to add some derived variables. It may not be cost effective to put a survey into SIR if only a few operations are to be carried out. However, SSD has made a major investment in SIR over the years so needs to be sure that alternatives will provide better value for money for the tasks which cannot be carried out in Blaise. Alternatives will be investigated, in particular Manipula, which is designed by CBS to be used with Blaise and will be more fully integrated with Blaise under the next release.

### *7.6 Data prep*

The data prep unit in OPCS carries out work for the whole of OPCS, but SSD is by far its largest customer therefore the move to CAI has serious implications for this unit. A data entry facility will still be required for the foreseeable future: two of the continuous surveys have diaries which respondents fill in which are returned to the office to be keyed; others have self-completion forms administered during the interview and supervised self-completion surveys are carried out in schools. Although SSD does not carry out many mail surveys, they are undertaken from time to time, as are mail screening stages to identify particular subgroups for interview. Some surveys have other documents which are currently keyed by data prep. In addition other information is entered into computers which is not keyed at present by the data prep unit, for example booking-in the receipt of paper documents mentioned above.

The problem is that, apart from the diaries for the two continuous surveys, none of this work is regular; so the requirement for data prep once the major surveys have moved to CAI will be small and fluctuating. It is not yet clear whether the data prep unit will continue to be viable in a few years time. If not, one option would be to abandon traditional 'head down' fast keying in favour of computer assisted data input (CADI) using Blaise for much of this work; another would be to use external contractors when fast keying is required. The former would bring all data entry work within SSD with flexibility to schedule it among other tasks involving data input into Blaise, for example booking-in the remaining paper documents as

described above, and other administrative tasks. Finally, extension of the use of computer assisted coding to very large classifications, which is under development, may also shift the balance of advantage towards interactive data entry for diaries and other self-completion documents.

## **8. Conclusions**

The description given above shows the situation from the point of view of one large survey organisation which uses Blaise; other organisations face different issues, particularly larger organisations carrying out a greater variety of different surveys. As yet few survey organisations have gone through a complete change from PAPI to CAPI as their main mode of data collection (although many more have made the change to CATI). So it is valuable to hear about the experiences of different organisations even though their situations all differ somewhat. Many of the changes described above are still in the future; SSD is in the middle of a period of major change which will continue over the next couple of years. At this stage it is important to have a vision of where one wants to be in a few years' time and to start working out in detail what needs to be done to realise the vision.

## **References**

- Bateson N and Hunter P (1991) *The use of CAPI for official British surveys* Bulletin de Methodologie Sociologique 30 pp16-26.
- Bethlehem J G and Keller W J (1991) *The Blaise system for integrated survey processing* Survey Methodology 17 (1) pp43-56.
- Brakenhoff W J, Remmerswaal P W M and Sikkel D (1987) *Integration of computer assisted survey research* in Automation in survey processing, CBS Select 4, Netherlands Central Bureau of Statistics.
- de Heer W F (1991) *The use of handheld computers in social surveys of the Netherlands* Central Bureau of Statistics The Statistician 40 pp125-138.

Jamieson R, Coutts M, Williams B and Braslins A (1992) *The building of an integrated collection operation in Statistics Canada's regional offices: a time for change* in Proceedings of the 1992 Annual Research Conference, March 22-25 1992, Bureau of the Census.

Keller W J and Metz K J (1989) *On the impact of new data processing techniques at the Netherlands Central Bureau of Statistics* Paper presented at the Conference of European Statisticians Working Party on Electronic Data Processing (21st Session) of the Statistical Commission and Economic Commission for Europe.

Nicholls W L II and Matchett S D (1992) *CASIC issues at the Census Bureau as seen by members of outside panels* in Proceedings of the 1992 Annual Research Conference, March 22-25 1992, Bureau of the Census.

Woolf M (1991) *The implications of changing automation systems for survey practitioners* Unpublished paper, Goldsmiths' College, University of London.