

3. Computer Assisted Personal Interviewing (CAPI)

3.1 Overview

What is CAPI?

Computer Assisted Interviewing (CAPI) is a computer assisted data collection method for replacing paper-and-pen methods of survey data collection and usually conducted at the home or business of the respondent using a portable personal computer such as a notebook. As the technology advances to provide lighter computers with longer battery life and user friendly software, CAPI will be used more often, especially for quick turnaround surveys.

How it works?

With a portable computer (generally a notebook), CAPI allows interviewers to conduct face-to-face interview using the computer. After the interviews, the interviewers send the data to a central computer, either by data communication or by sending a data disk using regular mail. CAPI can also include CASI (Computer Assisted Self-Interview) session where the interviewer hands over the computer to the respondent for a short period, but he/she remains available for instructions and assistance.

The role of Interviewer

The role of interviewer is a major factor in a CAPI interview, because the methodological survey-literature stresses the importance of good (nonverbal) communication and harmony between interviewers and respondents. Inexperienced interviewers may direct much of their attention to keeping the computer running and getting the answers correctly. If using the computer weakens the relation between interviewer and respondent, the interview will not be conducted optimally, and in consequence the data quality may suffer.

Therefore, an experienced interviewer can rely on the computer for routings and complex questions sequence, and therefore pay more attention to respondents and the social processes involved in interviewing. Sometimes, for instance in asking sensitive questions, less eye-contact is an advantage; experienced interviewers can use the presence of a computer to their advantages by directing their attentions to the screens when asking sensitive questions.

3.2 System Requirement

CAPI Architecture

As shown on the Diagram, CAPI generally has four components:

1. Sample Selection
2. Office Management Systems
3. Transmission
4. Remote Device and
5. Survey Software

Sample Selection

CAPI requirements lead to the development of a new sample and workload formation system. Workload formation is the allocation of households to interviewers. The sample selection facilities are designed to provide the basis for later standardized estimation components.

Office Management Systems

An Office Management System (OMS) is required to handle the non-field aspects of CAPI management. Processing required before dispatch and after receipt requires a range of office facilities, including instrument development, initial notebook configuration, resolution of edits, additional editing, reallocation of incomplete interviews, backup and recovery, etc.

Transmission

There is a variety of different ways in which the data collected in the field can be returned to the processing area. If electronic communications are difficult or unreliable, it may most efficient to download the data onto a floppy disk, and hand-deliver or post this to the regional or central office.

Remote Device

A small portable computer is required for CAPI enumeration. Portable computers currently take the form of laptops, notebooks PCs, Handheld PCs (HPCs) and Palmtop PCs. The choice of which of these is used depends on their price, size (most notably weight), processing power, ability to support the software used, ease of data entry, level of technical expertise for support of the computers and the field conditions.

Survey Software

The choice of a software product for building CAPI instruments depends on a number of factors, including the availability of products in the country, the type of hardware being used and the complexity of the particular surveys. The type of hardware may be the more important consideration as this will govern what software can be used efficiently and effectively.

Essential criteria for CAPI software products are that the product must:

- Provide the necessary functionality for a CAPI system according to stated requirements (e.g. skip logic, editing functionality, speed of operation)
- Be easy to use by survey developers, interviewers in the field and programmers. These are issues for all of these, but in terms of importance ease of use by survey developers and interviewers takes priority, as programmers will always find a way to make the system work.
- Have adequate support and training in the system provided by vendor.
- Be recommended by other statistical organizations (in that they have not had significant problems with the system or vendor)

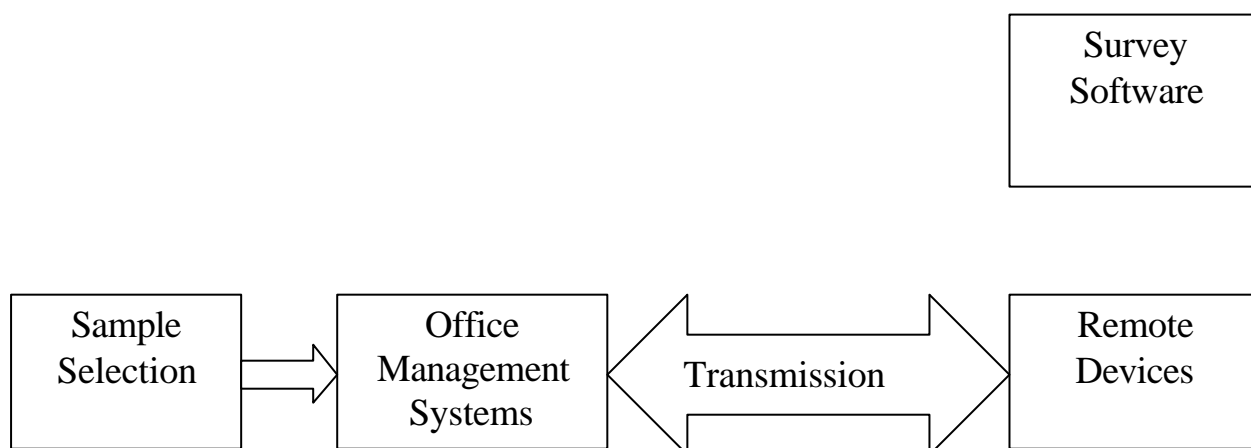


Figure 3.1 CAPI Components

Hardware and Software need

Going from paper-and-pen to computer assisted interviewing, CAPI requires investment in the form of computer hardware, software, and maintenance for each interviewer.

Computer Requirement Criteria:

Since interviewers using CAPI travels, then even though small portable computers can be used on a table or in one's lap, interviewers prefers handheld computers when conducting interviews on a doorstep, because small portable computers would be needed for this environment. Some criteria required: Screens visibility and readability, computer speed and memory, portable and light weight, etc.

Screens Visibility and Readability.

Many types of screens available: cathode-ray tube (CRT), liquid crystal display (LCD), backlit supertwist LCD, gas

Computer Speed and memory.

plasma, DC plasma, and electroluminescent display. Factors that should be also considered are screen contrast, resolution, blur, scroll, size, adjustability, and power consumption. The screen should be readable in the field environment such as lighting condition.

Portable and light weight computer

Speed and memory requirement is driven mostly by software needs. Software techniques may be employed to improve performance such as to use RAM disk files that are frequently accessed, delay I/O operations, and use machine language routines for CPU intensive operations

Battery charge

The carrying case should be utilized to protect the computer it is dropped or banged

Modem

The computer should not be lasted in the middle of the interview. The battery charge on a portable computer may last up to four hours, but some computer allows inserting battery packs as needed. Respondents might allow the use of AC outlets, and a car battery charger is useful on the road. Low battery indicator is helpful to interviewer.

Modem is required for transmitting interview results that performed by an interviewer to the central office. Transmission is by the modem based file transfer mechanism, though care needs to be taken in when handling multiple workloads and/or surveys.

3.3 Advantages and Disadvantages of CAPI***Advantages (compared to paper-and-pen interview)***

- a. No routing errors.
If the system is correctly programmed, routing errors, that is, errors in the sequence of questions, skipping and branching, will not occur. Therefore, missing data because of routing and skipping errors will never happened. Also, questions that do not apply to a specific respondent are automatically skipped. Hence, automatic routing reduces the number of data errors.
- b. Automatic validity checks
Data can be checked immediately because the program will perform some internal validity checks. The checks that can be performed, for example, are range check and consistency check. When errors are detected, there is an opportunity to correct range and consistency; therefore CAPI should lead to fewer data entry errors and missing data.

c. Questions formulation

The computer offers new possibilities to formulate questions. One example is giving each respondent with a unique sequence of question by randomizing the order of questions in a scale. This will eliminate systematic question sequence effects. Response categories can also be randomized, which avoids question format effects. The computer can also assist in the interactive field coding of open questions using elaborate coding schemes, which would be unmanageable without a computer. CAPI can also be used to employ question formats such as drawing line lengths as in psychological scaling.

d. Automatic clean data

There is no separate data entry phase. It means that the results will be available soon after the data collection phase is ready. In this case, CAPI has a real advantage when the results must be available as fast as possible.

e. Automatic interviewer control

The knowledge that the system accurately records information about the interview itself (e.g. time and duration of the interview, the internal relation between interviews and the order in which they are carried out) prohibits interviewers to cheat. CAPI provides a research organization with greater interviewer control and offers a protection against unwanted interviewer behavior.

f. More privacy

Using a computer could also lead to the expectancy of greater privacy by respondents since responses are typed directly into the computer and they cannot be read by anyone who happens to see the questionnaire.

g. More confidence and professional interviewer

Trained interviewers may feel more self-confident using a computer and behave more professionally. This could lead to more confidence of the respondent in the interviewing procedure. Social exchange theory, as applied to survey process, predicts that this should be to more willingness to comply with the interview's requests

a. The construction and programming of a questionnaire of CAPI takes considerable time.

Disadvantages

- b. When a respondent is totally unfamiliar with computers, then the respondent could refuse the interview or may provide undesirable answers especially for sensitive questions.
- c. Inexperienced interviewers may direct much of their attention to keeping the computer running and getting the answers correctly. If interviewers cannot type blindly, typing in long answer may lead to less eye contact between interviewers and respondents, causing the interviewers to miss nonverbal reactions of the respondents. Especially, when the computer is located between the interviewer and the respondent. The methodological survey-literature stresses the importance of good (nonverbal) communication and harmony between interviewers and respondents. If using the computer weakens the relation between interviewer and respondent, the interview will not be conducted optimally, and in consequence the data quality may suffer.

The role of Interviewer

The role of interviewer is a major factor in a CAPI interview

3.4 Operation and Management of CAPI system

Workflow

The following diagram (adopted from ABS) shows the core processes used during CAPI survey operation.

Survey Instrument Design

The operation and management of CAPI system are dependent on the availability of a good authoring system. If an authoring system is available, the typical survey instrument designer with little or no computer experience can prepare the CAPI survey instrument. Computer programming assistance will be needed to write the case management and output portions of the software. These portions of the survey, usually, vary with each survey or survey instrument; consequently, they must be custom programmed.

However, if an authoring system is not available, the entire CAPI instrument must be custom programmed with either a general-purpose language or a special purpose CAPI language. In either case, computer-programming expertise is required. The level of expertise dependent on the language selected. In addition, the survey instrument preparation will require the services of survey instrument designer who will need to work very closely with computer programmers.

Training Needs

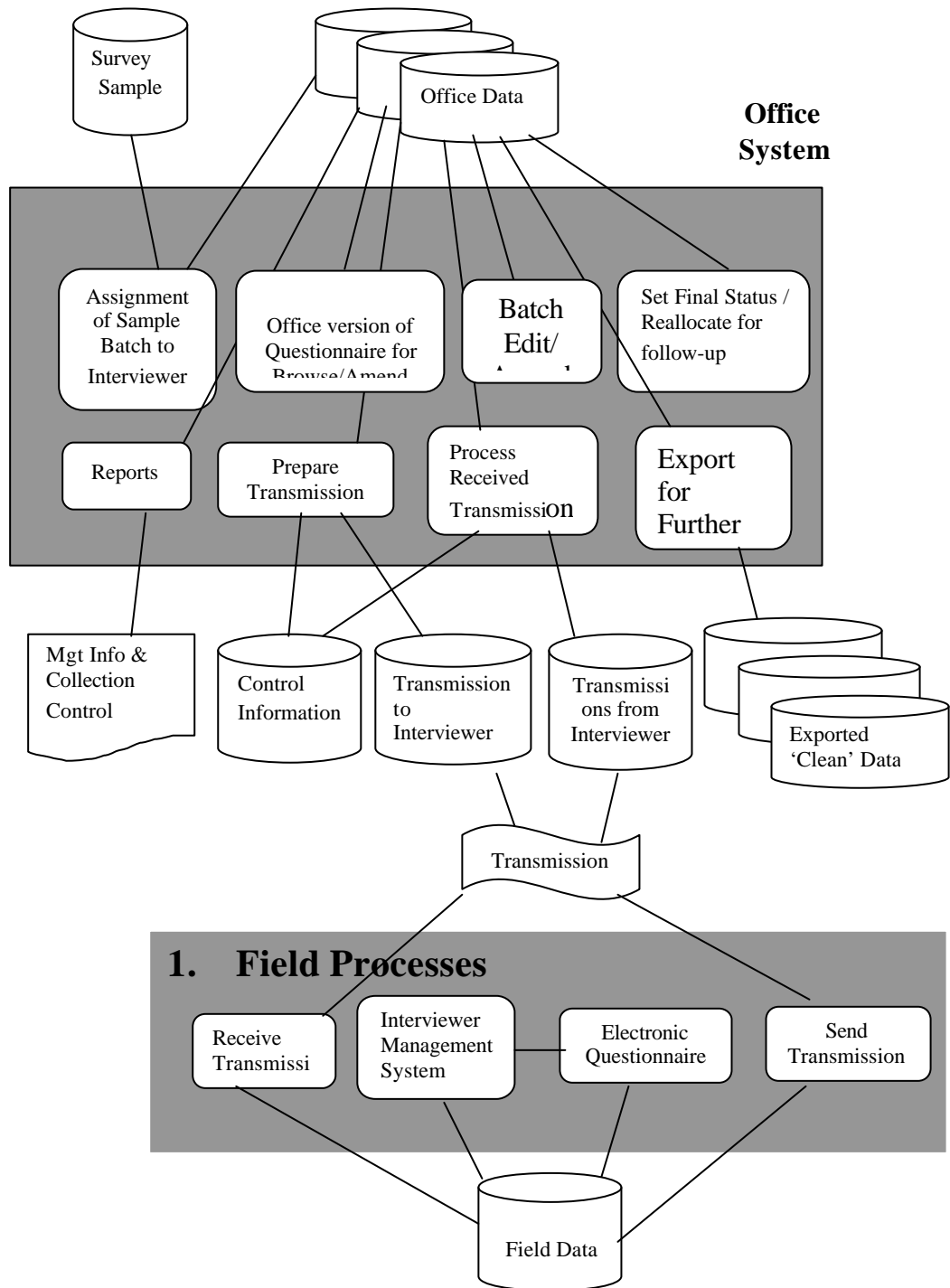
- In addition to conventional training, basic interviewer training in CAPI needs to include training in handling a computer and using the interview software. In many cases the training can be less costly because many topics (skipping, branching, selection rules) need not be taught since the interview software can handle this matter. Also, executives, research directors and field managers have to learn and appreciate computer assisted interviewing.
- The effect of CAPI on the interview process depends strongly on the amount of training and/or experience the interviewers have with this method of data collection.
- In CAPI survey we need interviewers that are well trained and experienced in computer assisted data collection techniques. Given that we have well trained and experienced interviewers, the altered interview situation is likely to have more advantages than disadvantages, especially with sensitive questions.
- Once trained, most interviewers prefer CAPI to paper-and-pen interviewing. With good training, even older interviewers without any previous computer experience can also enjoy the computer and conduct good interviews.

Cost Efficiency

- After the initial investments are made, CAPI may be cheaper than traditional data collection, but it all depends on the study, its complexity, its size, and its questionnaire. To evaluate the cost efficiency, a distinction should be made between front-end processing and back-end processing. In general, a well-designed computer assisted data collection requires investing more time, effort, and money in the beginning of the research (front-end processing), and time that is saved at the end stage (back-end processing). Especially the design and implementation of range- and consistency-checks (front-end) reduces the time needed to prepare the data for the analysis (back-end), and no questionnaire have to be printed and coded.
- Even computers are used frequently in the fieldwork; it will take about a year before the investment starts to pay back. Besides frequency of use, sample size is also a key factor for cost efficiency. Only with large sample sizes are the cost savings in printing, dispatch, data entry and editing (back-end costs) greater than the extra costs of questionnaire design and implementation (front-end costs).
- The overall CAPI is more expensive, mostly because of

added costs in training and supervising interviewers. The cost reduction in entering and cleaning data could not be large enough to offset the higher training and supervision costs. However, speed and accuracy compensate the cost.

TYPICAL CAPI PROCESSING



3.5 The Impact of Using CAPI

Data Quality

- The acceptance of computer assisted face-to-face interviewing is high for both respondents and interviewers, and there are no indications that using a computer disturbs the interviewing situation. In addition, a well-implemented CAPI system prevents many interviewer mistakes. As a result, we may expect that compared to traditional paper-and-pen methods, computer assisted interviewing has a positive effect on data quality.
- The percentage missing data is clearly lower in CAPI, mostly because interviewers cannot make routing errors.
- Presumably CAPI is less affected by social desirability bias, such as a sensitive question about alcohol consumption. In CAPI, the respondent answers sensitive question by typing their own answer into the computer, unseen by interviewers, like a self-administered questionnaire.

Respondents

- Although the first users of CAPI were afraid of a negative on the response rate, however studies report very low percentages of spontaneous negative reactions by respondents. Most reactions are neutral or positive.
- When respondents are explicitly asked for a reaction to using the computer, they generally react positively and are found to prefer reports that most respondents find CAPI interesting and amusing, and attribute a greater degree of professionalism to CAPI. The social interaction with interviewer is generally described as comfortable and relaxed.
- In CAPI, respondents were more positive about data-privacy, and judged answering sensitive question as less unpleasant.

Interviewer

- One important complaint by interviewers is about the difficulty of grasping the overall structure of the questionnaire, because CAPI still have more constraints than paper-and -pen methods.
- At first, interviewers may experience problems with open-ended questions. When they are not keyboarding literate and lack typing skills, entering a detailed answer to an open-ended question can be slow and laborious. However, when interviewers have keyboard experience they become fast enough typists to correctly records answer. Interviewers are well able to use the CAPI functions on which they are trained. By careful

analysis of the mistakes made, they come with several suggestions for further improving interviewer performance. Among these are providing laptops with key templates for prevention of function key errors and a better ergonomic design for laptop computer keyboards.

Case Study 2

Use of Handheld Computers in Philippine Data Collection

To increase the use of information technology in gathering relevant information from the sample households, The Philippine National Statistics Office (NSO) has utilized handheld/palmtop computer (HPCs) in data collection. Two hundred and fifty units of PSION Series 5 HPCs funded by The United Nations Development Program and 550 units of HP 360LX palmtop computers supported by the World Bank and the European Union. These HPs that run under Windows CE Services environment utilized the Visual Basic Software along with the Toolkit Accessory Software.

The HP 360LX palmtop computers were proportionally distributed to the provinces except in National Capital Region and Zamboanga City where hired enumerators conducted the interview of the sample households and in Batanes where there were communication and transportation problems due to weather conditions.

The HPCs were used as encoding devices for entering the data from the Annual Poverty Indicators Survey (APIS) questionnaires. The experiences of the five provinces from the pilot APIS gave them confidence and advantage over the 68 provinces during the conduct of the nationwide APIS. They could handle the computer carefully and operate them easily. Though the computers are different from the ones that they used in the pilot survey, they could run the program and follow the instructions of the manual correctly. For the rest of the provinces this has been a new and exciting experience.

However, problems not only in the computer program but also in the hardware. The problems on the hardware were to negligence in the use of the main batteries and backup lithium battery, and the limited size of the monitor made the program and the file display burdensome for the eyes of some enumerators. Other problems encountered were related to the kind of software necessary for writing the computer program

Though there were problems encountered, the main benefits that can be derived from using HPCs is the fast delivery of outputs or results of the survey. The used of HPCs reduced the waiting time of the data users by more than one-half, say, if the preliminary results of survey are to be released six months after the survey and the final results by more than six months. However, with the HPCs, the data can be released right away three months after the survey operation.

NSO is looking forward to maximize the use of HPCs in other surveys like the Labor force survey without the use of questionnaires. Likewise, in collecting the prices of commodities for the CPI, and HPC may figure prominently in the data capture and the data entry in the forthcoming census of population and census of agriculture and fisheries.

For further information: Contact Tomas P. Africa and Josie B. Perez, National Statistics Office.

Case Study 3 CAPI in ABS

Over the last three years the Australian Bureau of Statistics (ABS) has utilized CAPI for a number of their surveys. The surveys are including two waves of the longitudinal Survey of Employment and Unemployment Pattern, the Survey of Mental Health and Well-being and the Survey of Disability, Aging and Cares. CAPI will also be used in the forthcoming Housing Survey.

The interviewers for gathering the information use IBM color notebooks. The machines included disk encryption and access control software and the modem plugged in 'at home', extra battery and 'door stop' interviewing stands.

Software on the notebook is essentially based on Blaise and DOS. Blaise has proved to be useful and powerful environment for CAPI application (as well as some straight data entry application).

The interviewer workforce adapted rapidly to CAPI collection, and acceptance of it was high. Interviewer training for new CAPI instruments slightly longer than non-CAPI instruments. The workload for each interviewer may need to be merged with already captured information for some interviews for example for incomplete response or multiple interview instruments. Problems usually occurred when the workloads were not quite standard –e.g. when the household system was 'adapted' for the longitudinal person-based Survey of Employment and Unemployment.

Respondent reactions were a little more mixed, many were favorably disposed but some were less happy.

Instrument distribution used two techniques, either pre-loading instrument to disk before dispatching to interviewers, or transmitting the new files via the modem based file transfer system.

Field capture was essentially by direct entry at time of interview, but sequencing was always fully automated and this provided the ability to field instrument with much more complex sequencing paths (though causes some instrument validation problems). Verification was at-time-of interview when possible; improving data quality and reducing downstream problems, but edit override and later editing were also enabled on some fields. Remarks could be inserted against any fields, and post-interview editing could make use of these comments.

For further information: Contact Dr Rob Edmondson, Director, Technology Services Division, Client Relations Manager for Population Statistics Group and Methodology Division, Australian Bureau of Statistics (ABS); E-mail: rob.edmondson@abs.gov.au