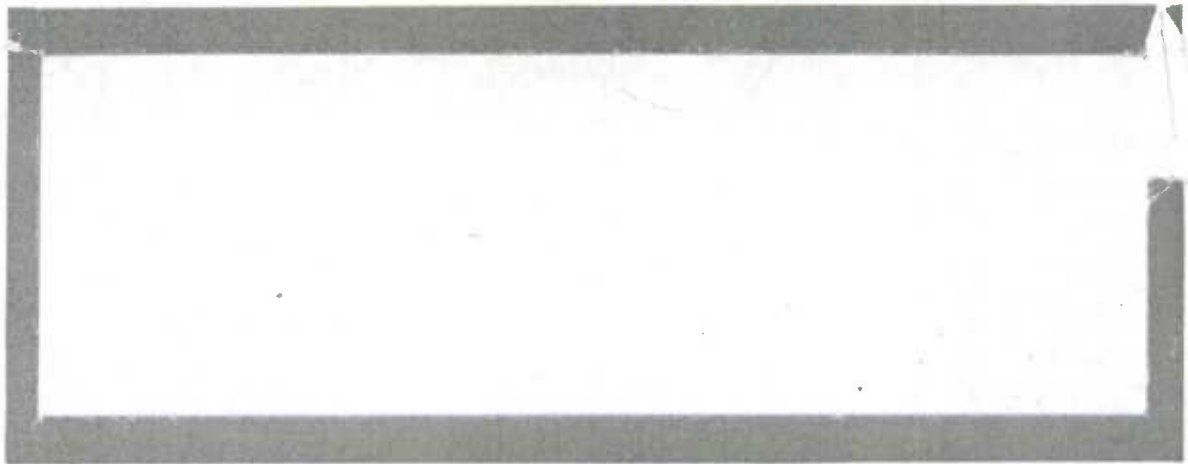


11-619E  
no. 95-09  
c. 2

 Statistics Canada  
Statistique Canada



Methodology Branch

Household Survey  
Methods Division

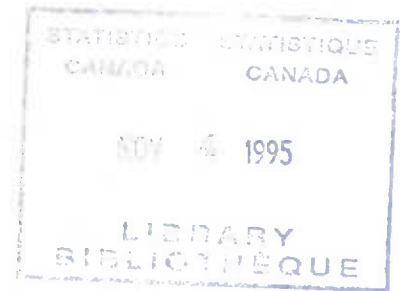
Direction de la méthodologie

Division des méthodes  
d'enquêtes-ménages

Canada

WORKING PAPER

METHODOLOGY BRANCH



**CONVERTING THE LABOUR FORCE SURVEY  
TO COMPUTER-ASSISTED INTERVIEWING**

HSMD - 95 - 009E

J. Dufour, R. Kaushal, C. Clark and J. Bench

Household Survey Methods Division  
Household Survey Division  
Survey Operations Division

Statistics Canada

October 1995

---

The work presented here is the responsibility of the author and does not necessarily represent the views or policies of Statistics Canada

# CONVERTING THE LABOUR FORCE SURVEY TO COMPUTER-ASSISTED INTERVIEWING

J. Dufour, R. Kaushal, C. Clark and J. Bench

Household Survey Methods Division  
Household Survey Division  
Survey Operations Division

## ABSTRACT

As the first phase of a multi-year redesign project, Statistics Canada recently converted the mode of data collection for the Labour Force Survey (LFS) from paper and pencil interviewing (PAPI) to Computer-Assisted interviewing (CAI). Conversion took place from November 1993 to March 1994. This paper gives an overview of how such a fundamental change to the survey's methodology was achieved. More precisely, it discusses: the collection before and after adopting CAI, the tests that provided evidence and experience, the preparation of the electronic infrastructure and the operational structure to support it, the strategy of conversion and the impact of conversion on data quality. The paper concludes with a brief evaluation of the conversion period and outlines the future plans regarding the LFS and CAI.

## LA CONVERSION DE L'ENQUÊTE SUR LA POPULATION ACTIVE AU MODE D'INTERVIEW ASSISTÉE PAR ORDINATEUR

J. Dufour, R. Kaushal, C. Clark et J. Bench

Division des Méthodes d'enquêtes-ménages  
Division des enquêtes des ménages  
Division des opérations des enquêtes

### RÉSUMÉ

Comme toute première phase d'un projet de remaniement étalé sur plusieurs années, Statistique Canada a entrepris récemment la conversion du mode de collecte des données de l'Enquête sur la population active (EPA), du mode d'interviews papier et crayon (IPC) au mode d'interview assistée par ordinateur (IAO). La conversion s'est effectuée du mois de novembre 1993 au mois de mars 1994. Ce document de travail explique comment un tel changement à la méthodologie de l'enquête a pu être réalisé. Cet article décrit plus précisément: les procédures de collecte avant et après l'adoption du mode IAO, les essais qui ont apporté l'évidence et l'expérience pour cette conversion, la préparation de l'infrastructure électronique et de la structure des opérations pour supporter le mode IAO, la stratégie de conversion et l'impact d'un tel changement sur les indicateurs de qualité. Ce document conclut avec une brève évaluation de la période de conversion et expose les plans futurs en ce qui a trait à l'EPA et au mode IAO.

## TABLE OF CONTENTS

<b>1. INTRODUCTION</b> .....	1
<b>2. THE LABOUR FORCE SURVEY</b> .....	3
2.1 Sample Design .....	3
2.2 PAPI Collection .....	3
2.3 CAI Collection .....	7
<b>3. TESTING</b> .....	8
3.1 Early Testing .....	9
3.2 Data Quality Test .....	9
<b>4. ELECTRONIC INFRASTRUCTURE</b> .....	11
4.1 CAI Equipment .....	12
4.2 Notebook Software .....	12
4.3 Case Management System .....	13
<b>5. OPERATIONAL INFRASTRUCTURE</b> .....	15
5.1 Communication Strategy .....	15
5.2 Training Field Operations Staff .....	16
<b>6. PHASING IN CAI</b> .....	18
6.1 Phase-in strategy .....	18
6.2 Data Quality .....	19
<b>7. EVALUATION AND STABILIZATION PERIOD</b> .....	25
<b>8. FUTURE PLANS</b> .....	27
<b>GLOSSARY</b> .....	28
<b>REFERENCES</b> .....	30

## 1. INTRODUCTION

Computer-Assisted Interviewing (CAI) has been promising survey takers data that is both more timely and of better quality for many years now. The earliest CAI systems were developed for centralized telephone interviewing. The development of systems for personal interviewing followed later as technology for notebook computers evolved. However, it is only recently that technology has reached the point where it is cost efficient for large government organizations to adopt computer-assisted personal interviewing as a standard for data collection.

In March 1994, Labour Force Survey (LFS) data collection was converted to CAI. It is the first of four major initiatives to redesign the LFS. A redesigned sample was introduced in October 1994. The LFS questionnaire is also being redesigned to take advantage of features of computer-assisted interviewing such as complicated branching and customized questions. Redesigned processing systems to support the questionnaire changes and new content will be introduced in 1997 along with the new questionnaire. This paper will focus on the motivation to and process of computerization of data collection in the LFS.

At STC, the following benefits motivate converting the LFS to CAI:

- i) **Timeliness** - It is anticipated that collection time will be reduced due to electronic transmission and the integration of data capture with interviewing. Timeliness is critical for a monthly survey like the LFS, where estimates are released within 2 weeks of collection. The quality of the final product, the released information, can be improved if the time saved in collection can be used in analysis and in otherwise improving products.
- ii) **Cost** - CAI is expected to improve the efficiency of the LFS by saving on data capture costs. After incorporating hardware replacement costs, net savings are anticipated.
- iii) **Data Quality** - An increase in the quality of data is expected for several reasons. First, data are edited at the time of collection thus allowing invalid data to be corrected at source. Also, some sources of human error, such as incorrect skip and branching patterns in questionnaires, are eliminated because of automation.
- iv) **Generalized Collection Tool For Household Surveys** - Much of the collection infrastructure in place for the LFS is used by other household surveys. Thus, converting the LFS simplifies the path to CAI for these surveys.
- v) **Powerful Questionnaire** - Notebook hardware and software can be exploited in redesigning the questionnaire by employing, for example, features such as complicated branching and customized questions. A new questionnaire utilizing CAI to its full potential is currently being developed.

Adopting CAI for a large regular survey in a government statistical agency is an extensive and expensive undertaking. At STC, data collection for the LFS was recently converted to CAI from Paper And Pencil Interviewing (PAPI). Conversion to CAI took place over five months from November 1993 to March 1994. Five major steps were involved in making the transition to CAI:

---

i) **Testing** - The first step towards conversion was testing to investigate feasibility, to determine the potential impact on quality, timeliness, and cost, and to determine new requirements for field operations. Three large tests have taken place since 1987. The last, from July 1992 to January 1993, was an opportunity to test notebook hardware and software in "near production" conditions. The test provided survey estimates from a CAI sample and from a PAPI sample. A statistically significant difference between the estimates could not be detected.

ii) **Electronic Infrastructure** - The second step involved building the electronic infrastructure to support the collection cycle. This included purchasing hardware and customizing software for interviewers. A large initiative was the in-house development of case management software for all of the required platforms at the interviewer, Regional Office (RO), and the Head Office (HO) level.

iii) **Operational Infrastructure** - Organizing and preparing the operational infrastructure for such an extensive change is the next step. This includes not only training interviewers and their supervisors and managers but determining how to meet new requirements such as the demand for hardware and software support in the field. The change to CAI has shifted the type of resources required in the ROs from clerical and data entry to computer operations and technical support.

iv) **Phase-in** - Phasing in of CAI interviews took place over five months from November 1993 to March 1994. In the first month, one-third of the interviewers in each RO completed their assignment using a notebook computer. In December, another one-third were converted while the final third switched to CAI in March 1994. The strategy for conversion was constrained by training and technical support resources and availability of back-up options. The phase-in was designed to make estimates comparable.

v) **Stabilization** - The benefits of CAI related to timeliness, cost, and data quality were not immediately realized during the phase-in period. Rather, the period provided a learning experience that resulted in some disruptions in timeliness and data quality. Some of these situations were traced to shortcomings of the new technology which resulted in lost and corrupt data. This final phase of converting to CAI is a period of stabilization where collection operations are tuned to the stable, efficient, and timely state that had been enjoyed under PAPI. Beyond this point, it is anticipated that the promised benefits of CAI will be more evident.

This paper gives an overview of how such a fundamental change to this survey's methodology was achieved. The organization of this paper is based on the above mentioned five steps involved in the conversion process. In the next section, a background describing the LFS collection process before and after adopting CAI is provided. In section 3, the testing that provided evidence and experience for CAI are described. Sections 4 and 5, have descriptions of the steps in the development of the electronic infrastructure and the operational structure to support it, respectively. Section 6 gives the phase-in strategy for converting collection operations, and the impact of converting on data quality. The stabilization period is discussed in section 7. Future directions with CAI are outlined in the last section. A glossary of terms is given at the end of the paper.

---

## **2. THE LABOUR FORCE SURVEY**

The LFS provides monthly estimates on the size, composition, and characteristics of Canada's labour force. These estimates are important economic indicators for Canada and are used extensively by all levels of government, the private sector, international agencies, and the general public. Monthly movements in employment and unemployment rates are reported widely by the news media.

For the purposes of this report, changes in the LFS since the conversion to CAI are disregarded. Descriptions of sample design and collection are those that were current during the conversion. At the time, the LFS estimates were based on a sample of 59,000 households. About 1000 local interviewers and 100 senior interviewers collected the data each month over one week by a combination of personal visits and telephone calls. Data collection operations were managed by eight ROs across Canada and a group at HO in Ottawa. In the next three sub-sections, the sample design and collection process of the LFS under PAPI and CAI are described.

### **2.1 Sample Design**

The LFS sample is selected using a rotating multi-stage design (Singh et al. 1990). In the LFS, within provinces, the population is stratified into: self-representing areas (SR), which are larger urban centres; non-self-representing areas (NSR), which are mostly rural areas; and special areas, which are primarily remote areas and institutions. Within SR and NSR areas primary sampling units (PSUs) are city blocks, called clusters, and groups of enumeration areas, respectively. The NSR has an additional stage of sampling, where groups of EAs are sub-sampled to get clusters in small urban areas and groups (which are EAs or collapsed EAs) in rural areas. Special areas follow a sample design similar to NSR but tailored to the characteristics of the specific area. All but the last stage use probability proportional to size sampling schemes. In the final stage of selection, dwellings are sampled systematically from a cluster/group. Interviewer assignments cover dwellings in several clusters/groups. A dwelling remains in the sample for six consecutive months before rotating out. Survey information is collected from the current inhabitants of the dwelling, called the household. If a dwelling is vacant, a business, incomplete, uninhabitable, etc., it is considered out of scope.

### **2.2 PAPI Collection**

The LFS collects demographic and labour force information on Canadian residents. Two questionnaires are used: Form 03 collects data on the composition of the household including basic demographic information; and, Form 05 records labour force activity during the reference week<sup>1</sup> for all persons in the household aged 15 and over. The LFS also serves as a general vehicle for household surveys; information for other Statistics Canada surveys is often collected at the same time.

---

<sup>1</sup> The reference week is usually the week containing the 15th of the month.

---



The interviewing and collection take place the week following the reference week. Interviewing is generally done in person in the first month a dwelling is in the sample and by telephone in subsequent months. The interview is carried out at the doorstep if a member<sup>2</sup> of the household over 15 years old is available. The interviewer fills the questionnaire out in pencil and edits the questionnaire at home before shipping it to the RO. During the interview, the interviewer may use abbreviations for the industry and occupation information. As part of the editing at home, the abbreviations are expanded in full. If the interviewer is unable to reach respondents the first time, an attempt is made to determine the last name of a member of the household to be able to contact the household by phone. In such a case, the follow-up and the birth interview are done by phone. If this is not possible, several personal follow-ups are done by phone. If this not possible, several follow-ups are done.

Figure 1 summarizes the PAPI survey cycle. The monthly sequence begins when the sample is selected at Head Office in Ottawa. Specification of the dwellings that are in the sample is sent to the ROs so that a control list of all dwellings in selected clusters, referred to as listings, can be prepared for each assignment. Hardware encryptors in place between HO and the ROs insure the confidentiality of the survey data.

The ROs use a Honeywell computer to receive and store the sample, print the questionnaires with pre-filled information. Control information is pre-filled for dwellings entering the survey for the first time. For the remaining households that were in the survey during the previous month, demographic and some job description information from the previous month is pre-filled. In the ROs, assignments are prepared and shipped to the interviewers. The assignment packages contain introductory letters, blank forms and pre-filled forms.

Interviewers receive their assignment and plan their work one week before interviewing. This includes scheduling when interviews will be conducted during the next week and planning a route to follow for conducting personal interviews. Data for completed interviews are returned to the RO daily. With PAPI, data are manually edited by both interviewers before returning completed questionnaires to their ROs and by RO clerical staff. At the ROs, the data are captured using the mini-computer, which is subject to quality control procedures. Finally, the captured data are transmitted back to the HO.

Survey data are received, edited and corrected, coded and tabulated in Head Office. Data are also automatically edited at HO with some manual resolution. The last two weeks are used to process, weight and analyze the data. The conclusion of the monthly cycle comes with preparation and translation of the text for the release and dissemination of estimates.

---

<sup>2</sup> For both personal and telephone interviews, proxy interviews are accepted from persons aged 15 or more.

---

# LFS PAPI SURVEY CYCLE

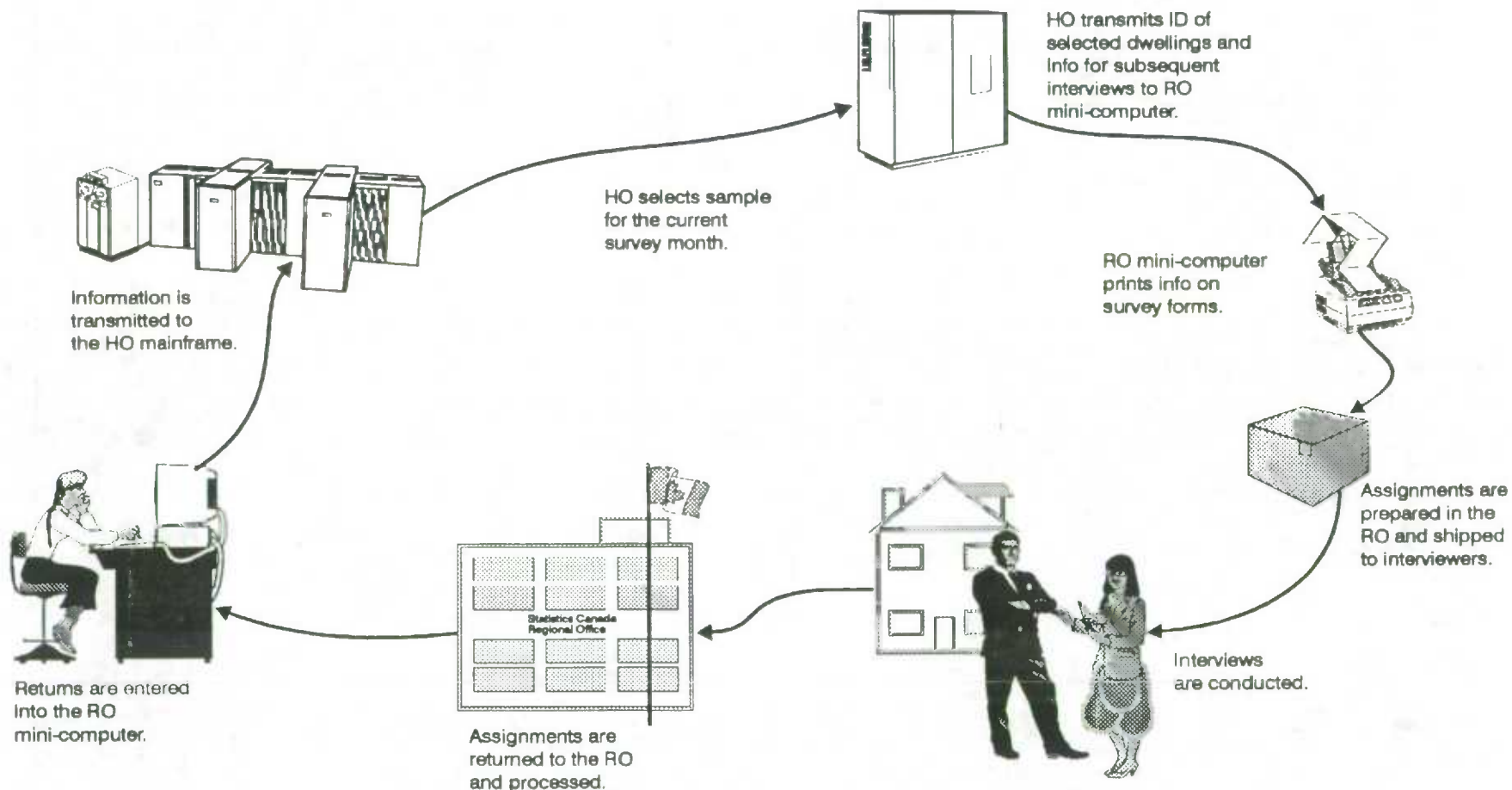


Figure 1: PAPI Collection Cycle

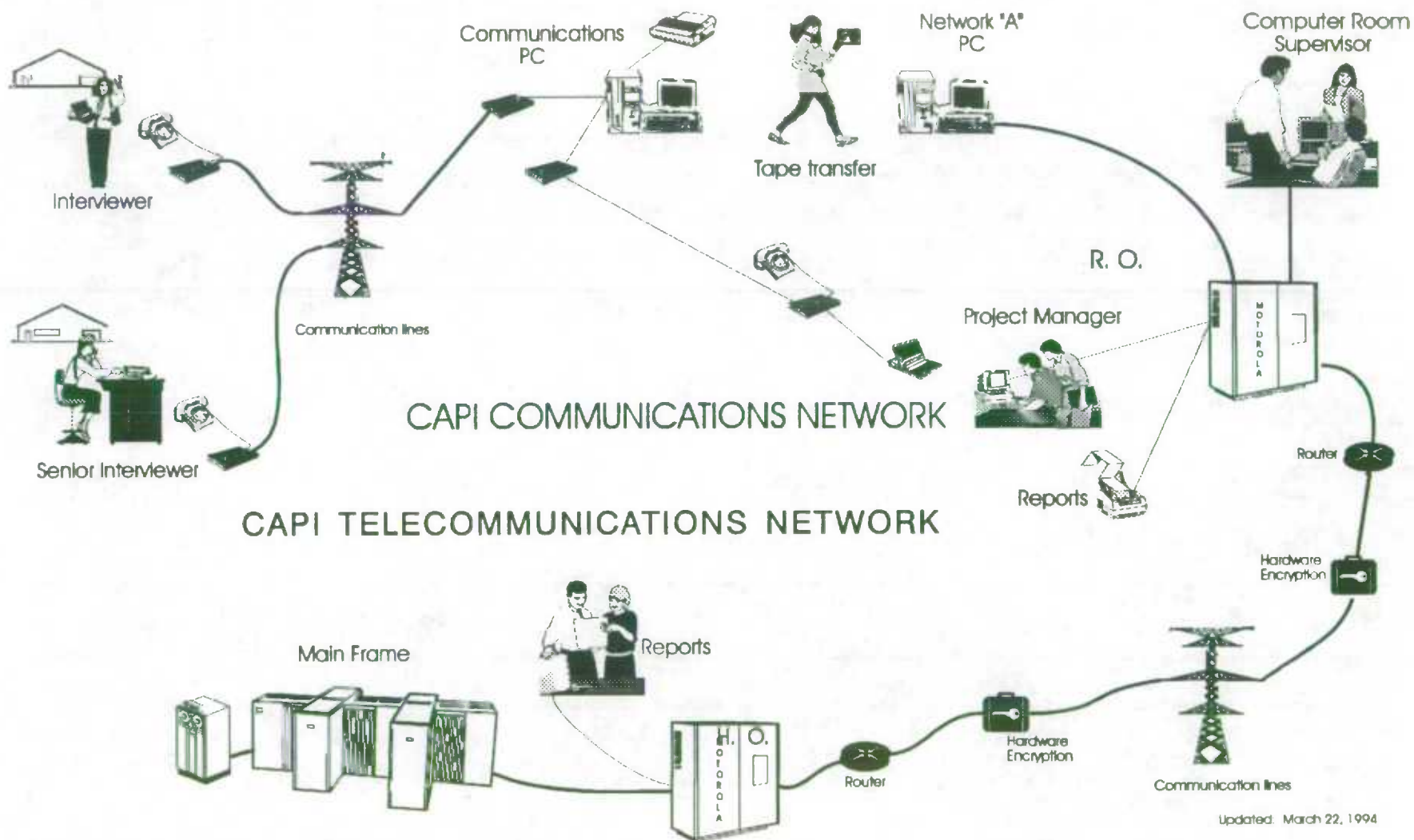


Figure 2: CAI Electronic Infrastructure

### 2.3 CAI Collection

No changes were made to the questionnaire or concepts, when Statistics Canada converted data collection for the LFS to Computer-Assisted Interviewing (CAI). Paper questionnaires were translated directly into electronic forms. In the process of integrating collection with capture, some of the quality control measures that were in place at the RO are now eliminated. These have been replaced with validity and cross edits on the collection instrument. While the information collected was not changed, the entire collection process was computerized. Figure 2 presents a summary of the electronic infrastructure supporting collection.

No substantial changes have been made in the HO-RO communications, but several interfaces were introduced at both ends to accommodate the changes in data format. The data flow for getting sample out to ROs continues to be from the Head Office (HO) main computer to the Survey Operations Division's UNIX system. This system's role is to split the sample up and relay it to the Honeywell system of the appropriate RO. In the RO, the manual processes related to paper questionnaires no longer exist. The RO staff prepare interviewer assignments using Case Management software before routing it to interviewers.

The process of sample selection and transmission between the HO and RO has changed considerably with the introduction of CAI. In every RO, the respective assignments are downloaded from the UNIX system to a micro-computer that belongs to the confidential Statistics Canada network (network A) where they are encrypted and downloaded to tape. Backups are made and hard copy audit reports are produced. The tape is physically transferred to the Communications PC that belongs to the Statistics Canada's external network (network B). The tape is copied to the Communications PC and backups are made. The Communications PC is responsible for dispatching the assignments by modem to each interviewers' notebook. In the CAI environment, the questionnaires take on an electronic form which resides on a notebook computer. Interviewers receive and return their assignments by hooking up their notebooks to a modem and dialling in to the RO.

Each interviewer is now equipped with a notebook computer that contains the assignment, electronic questionnaires, and case management software which moves data from one process to another and provides tools to organize and manage work. The interviewer enters a respondent's data directly into the computer at the time of the interview. A separate data capture step is eliminated since interviewers enter the data directly on to their notebook computers. Field editing is automated because some of the edits are programmed into the notebook. A message is displayed if an entry is invalid or inconsistent. Interviewers can choose to change data at that point or indicate that the data is correct by not making any changes.<sup>3</sup>

Getting assignments to and from interviewers is done through telephone lines using modems. Interviewers receive their assignment by hooking up their notebook computers to a modem and then instructing their notebook to "Transmit and Receive Cases" by selecting a menu item. The software takes over from there to connect and transmit. Work planning is done using tools available on the notebook such as the ability

---

<sup>3</sup> There are no plans to reduce the amount of editing done at the HO as a result of converting to CAI since the risk of corrupt data from the multi-platform distributed CAI systems is deemed, at this time, to be too high.

to sequence an assignment by user-defined codes, or the ability to view the demographic composition of households that were in the sample last month. During interviews, responses are entered directly into the electronic questionnaires. Entries are edited "on-line" as described above.

The process works in reverse for returning data except that it takes place every day for up to one and a half weeks during the collection period. As the interviewer completes a case, a back-up to diskette is done and the case is compressed for transmission to the Communications PC. The communications software is usually configured to transmit immediately or overnight. Interviewers must hook up their modem and notebook prior to transmission time. After survey data are transmitted to the Communications PC, the seniors interviewers and the project managers' notebooks dial up and sign on to the Communications PC and receive data to dynamically produce screen reports of the transmission activity. Any transferred cases may be picked up at this time. When Seniors or Project Supervisors transfer cases, the Interviewer's notebook pick up the new cases when connecting to the RO. These transfers require some manual intervention from RO staff. The Communications PC has error and transmission logs, which are reviewed to ensure problems are detected early. Data received by the Communications PC is downloaded to tape daily, and a backup of the files is made. The tape is transferred to the confidential Statistics Canada network PC. On this network PC, the tape is mounted and the data files are uploaded. At this stage, the files are software decrypted and decompressed, backups are made and hard copy audit reports are produced. Next, completed cases are transmitted to the Regional UNIX system. On the UNIX system, the Case Management system generates a series of survey status reports. Completed cases are forwarded to the Head Office UNIX system. The machine produces an error log which must be checked to ensure there were no transmission problems. One more time, the files are backed up; and finally on the Head Office UNIX system, status reports are generated, and the completed cases are transmitted to the mainframe computer for processing.

Maintaining the confidentiality of data, as required by the Statistics Act, was a critical requirement of the notebook computer and its supporting systems. This was one of the STC's first experiences in transmitting confidential data from private homes. There were concerns over both data being stored in notebooks in the possession of interviewers and data transmitted on unsecured lines. Several procedures were established to preserve confidentiality. First, the step of physically transferring data from one PC (confidential network) to another network (external network) in the ROs ensures that no internal network is accessible to the public by modem. Confidentiality is also ensured by encrypting data before the cases leave the interviewer's notebook for transmission over telephone lines and before they are transmitted from the ROs to Head Office in Ottawa. Last, an access control system (PCDACS) is built into the notebook computer so that only the interviewer has access to the information stored in his/her computer.

### 3. TESTING

The decision to convert to the CAI mode was the result of several years of research. Since the late 1980s, many tests have been performed on the LFS, with the specific objective of evaluating the potential of computerized data collection. Early testing included a controlled study comparing CAI in a centralized environment with PAPI and a feasibility study of using new technologies in the LFS. At the beginning

---

of the 1990s, the LFS underwent a major data quality test that led to the decision to convert the LFS PAPI mode of data collection to the CAI mode.

### **3.1 Early Testing**

The first test of CAI at STC was in the form of a centralized CATI (computer-assisted telephone interviewing) test, which took place from February to September 1987. Centralized CATI is telephone interviewing done from centralized locations, like ROs, rather than from decentralized locations, like interviewers homes. The test was a joint Statistics Canada/U.S Bureau of the Census research project that used the USBC Census CATI System and the LFS as the collection vehicle (see Catlin and Ingram, 1988). This large scale test of centralized CATI used random digit dialling to contact a monthly sample of 2,000 households. The sample was randomly assigned to either centralized CATI or PAPI. The objective of the test was to evaluate the effect of CATI on survey costs and data quality. The results showed that the item based error rates were reduced with CATI. Response rates were lower for CATI with no difference in the refusal rate but the non-contact rate was higher due largely to the longer CATI interview. It was concluded that although the longer interviews would contribute to higher costs, savings could be realized in reduced time between interviews, in reduced supervision, and in combining collection and data capture. Due to the combined effect of longer CATI interviews and lower contact rates, centralized CATI was not recommended.

The second attempt was a small scale testing of decentralized CAI (see Kaushal and Laniel (1995)) in 1991. The objective was to study the feasibility of using current notebook computer technology for the LFS. There were concerns about portability, durability, suitability for winter use, durability of batteries, and the availability of software. This test evaluated Datelite, a portable micro-computer which used touch sensitive screens to input data directly onto an internal hard disk. Nine interviewers and one senior interviewer collected data using Datelite for seven months from about 500 dwellings a month. Nonresponse rates were higher in the first 3 to 4 months due to an initial learning period. Software problems contributed to higher error rates. However, interviewers and respondents were found to be very receptive to changing the collection mode. The life of the micro computer's NICAD batteries, at 1 to 1.5 hours, and the 7.5 kg weight of the laptop, were critical failings of the Datelite hardware. The conclusion of this test was that notebook computer technology was adequate to meet the demands of Statistics Canada for the LFS. Technology for touch sensitive screens was promising but not yet sufficiently advanced for LFS purposes.

### **3.2 Data Quality Test**

This section describes the Data Quality Test, conducted from July 1992 to January 1993 (refer to Kaushal and Laniel, 1995). This test also served as a precursory exercise in CAI for the LFS. The objective of this test was to predict the impact of the implementation of CAI. More specifically, the impact on the labour force series, quality and interviewing cost were estimated. Operational development of CAI and evaluation of CAI electronic infrastructure were other objectives.

---

The reliability requirement for the test was to detect a difference of one percentage point in the unemployment rate due to change in mode. A control group of PAPI interviews and a test group of CAI interviews were conducted under similar circumstances to allow a comparison of the two modes. The required sample size to meet the reliability criterion was 3,150 households per month for the control group of interviews. The test group was of the same size. A total of 6,300 households were interviewed per month with 2,100 households being interviewed in each of the 3 non-LFS weeks. Forty-two trained LFS interviewers, in four different regional offices, were part of the test. The test was conducted for 6 months in non-LFS weeks to avoid contamination of the LFS.

Test and control interviewers were paired and assigned to cover the same geographic area. In SR areas, sample interpenetration was done at the blockface or cluster level. In NSR areas, however, the interpenetration was at the PSU level due to operational reasons. While it would have been preferable to nest the treatments within interviewers to avoid an interviewer effect and other confounding effects, this option was not undertaken for operational reasons. To isolate the CAI effect, the test design was structured so that the confounding effects could be estimated. In the fourth month, the test group (CAI trained) used PAPI before returning to CAI for the remaining months. The fourth month, when both groups used PAPI, was used to estimate systematic non-mode differences that were confounded with the mode effect.

The non-mode confounding factors were deemed important enough to be incorporated into estimation. As can be seen from table 2.2A, statistically significant differences were observed in November, when both sets of interviewers were using the same instrument. The confounding factors in the unemployment rate were statistically significant at the 5% confidence level whereas the employment to population ratio effect was significant at the 10% level. Although the confounding factors were not statistically significant for the participation rate, all the rates were adjusted for the confounding factors. It can be seen from Table 2.2B, the unemployment rate, employed to population ratio and the participation rate were not affected significantly by the change in mode.

**Table 2.2A**  
**Confounding Effects in the Labour Force Rates:**

Rates	Statistics	Estimates
Employment to Population Ratio	Estimate	1.51
	Std Dev	0.89
	P-Value	0.09
Unemployment Rate	Estimate	-1.20
	Std Dev	0.55
	P-Value	0.03
Participation Rate	Estimate	0.89
	Std Dev	0.87
	P-value	0.31

**Table 2.2B**  
**CAI Effect on the Labour Force Rates**

Rates	Statistics	Estimates
Employment to Population Ratio	Estimate	-0.10
	Std Dev	0.55
	P-Value	0.86
Unemployment Rate	Estimate	0.08
	Std Dev	0.38
	P-Value	0.83
Participation Rate	Estimate	-0.16
	Std Dev	0.53
	P-value	0.76

The overall mode effect on the *nonresponse rate* was estimated to be statistically non-significant at 0.20 percentage points with a standard deviation of 0.36. The mode effect for the *vacancy rate*, was estimated to be 0.93 percentage points with a standard deviation of 0.44. Therefore, based on a 95% confidence level, there is a significant mode effect on the vacancy rates. A possible explanation for the mode effect is that interviewers did not have access to a questionnaire after completion and were unable to make corrections if new information became available. *Edit failure rates* were somewhat lower for edits that were programmed into CAI but not for others. There appeared to be no apparent modal pattern for *industry and occupation automatic coding rates*. When excluding data capture costs, *cost projections* based on administrative data did not appear to be significantly different for the two modes.

The operational evaluation of the CAI environment highlighted two problem areas. First, it was determined that the hardware and software combination was not fast enough to carry on an interview that was as smooth as one with a paper questionnaire. Second, the batteries used (Ni-Cad) were not suitable because they developed a memory for charge and needed to be drained completely before being recharged.

The conclusion from the test was that when implemented, CAI would not have a discernible impact on the series, overall quality or cost. Also, the observed operational problems could be remedied through hardware and software changes. With these recommendations and experience, CAI was implemented in the LFS in the fall of 1993.

#### 4. ELECTRONIC INFRASTRUCTURE

The first activity in the process of converting field operations to CAI was the selection of the collection instrument. The criteria for selection were low cost, low weight, option of operating on batteries, possibility of recharging batteries without exiting the application and powering down the notebook, a



power management chip (called an SL chip) and a specific keyboard configuration. The next step was to develop the software for the notebook to generate electronic questionnaires and to plan, monitor and control the activities of data collection. The latter is called a Case Management System. These three main components of the electronic infrastructure, namely the CAI equipment, the notebook software and the Case Management System are discussed in this section.

#### **4.1 CAI Equipment**

The hardware chosen to gather the survey data is an AST PowerExec 3/25SL notebook computer. This portable computer was selected for its rapidity, lightness and reliability. At the time of purchase in 1993, the cost of each unit was \$1,800. The size of the notebook is 21.6cm X 24.2cm (8.5" X 11.5") and its weight is 4 kg with the battery. The screen displays text and graphics using an illuminated Liquid Crystal Display (LCD). These machines have a 60 MB removable hard disk drive and 4 MB of RAM. There is also one floppy disk drive which can read, write to, or format high-density 1.44 MB or double-density 720 KB 3.5 inch diskettes. Operating speed is variable (3.125 MHz, 6.25 MHz or 25 MHz).

Nickel metal hydride batteries with a four-hour lifetime are an alternative power source to AC. They can be recharged using the computer or the charger provided. These batteries were an improvement over the Ni-Cad ones used for previous testing. A critical failing of the Ni-Cad was a "memory effect" whereby the efficiency of the batteries reduced over time if they were not fully depleted before charging. Each interviewer receives three batteries and an AC adaptor.

ZyXEL modems with 14,000 baud are used by interviewers for transmitting and receiving data. This is a considerable improvement over modems with 2,400 baud used for previous testing.

The AST PowerExec 3/25SL is equipped with a removable hard disk drive which is a must on a CAI instrument. This was not a part of the specifications for choosing the technology, but it has proven to be practical since it is easier to solve problems by replacing hard-drives instead of transporting the notebooks.

#### **4.2 Notebook Software**

The development of the software for the notebook began largely with preparation for the Data Quality Test in 1992. The version implemented for production varied only slightly from what had been used in this test. The notebook software was a combination of commercially available software and software developed internally. There are two components to the software; namely the electronic questionnaires and the Case Management system. The questionnaire software includes skip patterns and edits. There are two questionnaires, DEMOG and LFS, that correspond to the household roster (FO3) and the labour force questionnaire (FO5). The Case Management system, which transmits data to and from the notebook, is part of a larger system described in the next subsection.

---

Involving interviewers in designing the software gave the developers important insight into client perspective. Once the software was developed, it was sent to a few interviewers for their input. Interviewers provided valuable comments on the design of the screens and the ease with which they could navigate through the software. The result was an interviewer-friendly and operationally efficient tool. Inclusion of interviewers in this process not only aided development, but also increased interviewer knowledge of CAI and support for the change. Similarly managers were involved in developing the management information reports.

When designing the software, it was important to be consistent in the way information was displayed on the screen from one survey application to the next. As well, keeping any function keys consistent throughout the applications was desirable. For example, if F7 means go back to the previous question, it should have that same function in all applications. The Labour Force Survey interviewers conduct several other household surveys during the course of the year.

Different types of software were considered for programming the questionnaires. Software from the Centre for Human Resources Research (CHRR), Ohio State University was selected for programming both the demographic (FO3) and the Labour Force (FO5) modules. However, portions of the FO3 had to be programmed in CLIPPER and Turbo Pascal because the FO3 had several questions to be addressed on one screen. In CHRR, only one question can be addressed at a time.

Maintaining the confidentiality of data, as required by the Statistics Act, was a critical requirement of the notebook computer and its supporting system. Several procedures were established to preserve confidentiality. Software, called MAILSFAE, encrypts data before leaving the interviewer's notebook for transmission over telephone lines and before data is transmitted from the Regional Offices to Head Office in Ottawa. An access control system built into the notebook computer is PCDACS (Personal Computer Data Access and Control System) which was approved by Electronic Data Processing (EDP) Security at Statistics Canada. Only the interviewer has access to the information stored in the computer; a user ID and a password are required.

### **4.3 Case Management System**

The Case Management system is a sophisticated system that was developed in-house by the Survey Infrastructure Applications (SAI) section of Statistics Canada to handle all surveys using CAI. UNIX scripts, Oracle SQL, C language programs, batch files, DOS scripts, and Clipper were all used to create the applications. One characteristic of this complex computer system is its flexibility. This system manages cases which can be defined differently depending on the survey. For example, in the LFS, a case refers to a sampled dwelling whereas in other surveys it may refer to a person. The collection period is another flexible component of the system since it is not the same for all surveys. The system manages the survey activities according to the beginning and the end of the survey cycle. The survey managers also have the possibility of defining their own outcome codes.

The Case Management system performs three main functions, namely routing, reporting and assistance. The routing component is responsible for all movement of cases during the survey period. For the LFS,

---

this component basically dispatches the cases to the appropriate interviewer's notebook and returns the completed cases to the Head Office. In other words, the system handles the normal flow of survey data through Head Office, from Head Office to Regional Offices, and from Regional Offices to individual interviewers, then back again. This component is also able to handle special needs such as transferring cases from an interviewer to a senior interviewer or a project manager and back to an interviewer, if necessary. For the routing of cases, the system is equipped with an important module called the import/export function.

Since the electronic infrastructure is made up of many platforms, the routing function has to transfer cases across multiple computer platforms. To achieve this, every time a case reaches a new piece of equipment, a control process has to be performed to accept the data. The same process can also be applied whenever a case leaves a machine. Some programs, for example, require parameters such as the date and the RO identification before being executable. In addition, the system itself involves some control and verification processes. Before accepting an FO5 form, for example, the system must have received the corresponding FO3 form.

The second function of Case Management consists of reporting on the data collection activities. The reports generated by the system can be used to describe the status of the survey, to monitor the survey or to evaluate both staff performance and progress. This reporting component controls information related to interview completion, interview duration, number of contacts, edit failures, status of cases, type of nonresponse, etc. A wide variety of survey indicators can be generated at different levels, such as at the case level, supervisor level, operations manager level, survey manager level or higher management levels. For example, a project manager can easily monitor the survey by obtaining the progress of the previous day of interviewing. The reporting component also has the facility of storing and retrieving information regarding progress, time, cost, and resource utilization at different levels.

The third function of Case Management is assistance to the interviewers during the progress of the survey. This function helps interviewers to perform their work efficiently. The system, for example, offers not only the choice of language to respondents, but has been programmed to provide the interviewers with the same option. The interviewers can also use this function to plan their work by sorting their assignment by current status, sample identification or appointment. They also have the possibility of sorting their cases by address, for instance, to help in planning their route for face-to-face interviews.

The interviewers use the Case Management system to select a case for interviewing, to create new FO5s for new household members, to generate FO5s for birth households and to close a case, to name just a few examples. The assistance function also has the capacity for maintaining respondent information and for keeping in its memory pertinent notes related to a case. In addition, help keys that assist the interviewers in their work are available.

Finally, the assistance component serves as an information source. After data transmission, for example, this function shows the results of the transmission: how many cases were exported that are transmitted and how many cases imported that are received. If a problem occurs during the transmission, the interviewers are informed so that an action can be taken.

---

## **5. OPERATIONAL INFRASTRUCTURE**

The operational infrastructure to support CAI differs considerably from the one needed for PAPI. In the ROs, the need for clerical and data entry resources is all but eliminated. In a CAI environment, there is an increased need for computer operations staff to run and support the large computer systems in the RO. There is also a need to provide expert technical support for the notebooks including trouble shooting and maintenance of notebook software and hardware. Effective and comprehensive training of all staff affected by the introduction of CAI was crucial to successful implementation. To keep all staff abreast of changes due to CAI, a critical component of converting field operations was the development of a communication strategy.

### **5.1 Communication Strategy**

A communication strategy was developed and implemented to ensure that all staff affected by the move to CAI were informed about their role in the conversion, how they would contribute to the success of the project, changes that were being made and what training they would receive. The information any one group received had to be consistent and clear since the information needs differed. Existing communication vehicles such as management meetings, telephone conference calls, monthly letter to ROs and electronic mail were used. As a part of the communication strategy, a question and answer flyer was distributed to all ROs and field staff to address concerns raised by staff at all levels. Another goal of this flyer was to respond to potentially negative questions by pointing out and emphasizing the positive aspects of CAI. To complete the communication strategy, two new tools were created: i) a video that introduced the CAI application to field staff and ii) a bi-monthly newsletter that were directed at interviewers.

Communication between field staff unfamiliar with computers and computer-literate staff unfamiliar with interviewing was a particular challenge. Keeping everyone abreast of software updates is critical. Involved parties need to be kept informed of plans and progress in a vocabulary that is familiar to them. A newsletter for interviewers and senior interviewers was created for this purpose. Involving interviewers in the development of software also kept them informed and enthusiastic about the upcoming changes.

Given the size and complexity of the survey cycle, the number of transmission processes, the technology used and the number of people involved, technical problems were anticipated. To deal with hardware and software problems involving the notebook, a hierarchical approach along reporting lines is used to provide technical support. Interviewers contact their senior to report situations which they are unable to resolve themselves. If the senior cannot solve the problem, then it is referred to the project supervisor. Problems for which the supervisor could not provide a solution were sent to RO technical support staff. Finally, if RO technical support staff were unable to provide a solution, the problem would be referred to HO technical support. At the HO, the strategy adopted was that technical support staff would deal with hardware problems, while the software developers would provide software support.

To deal with all problems related to the application, laptop and UNIX system, a Help Desk was set up in Head Office. During the phase-in, the help desk was fully operational during the reference, survey

---

and post survey weeks. As well, key technical and operational staff were available by cellular telephone after working hours and on weekends.

## **5.2 Training Field Operations Staff**

A CAI collection process requires a considerably different pool of skills from that needed for PAPI. In the ROs, the need for clerical and data entry resources reduced considerably while there was a requirement to operate the new RO systems and to support the notebooks. These needs were filled by some retraining and by some hiring.

Effective and comprehensive training of all staff affected by the introduction of CAI was crucial to a successful implementation. In developing the training strategy and programs for the conversion to CAI, the experience gained from LFS CAI field test was used, as well as the experiences of colleagues in Britain and in the United States.

A large part of training was formal training sessions. Sessions were held for five staff groups: i) Regional Managers and Project Supervisors, ii) Technical Support staff, iii) Computer and Operations staff, iv) Senior Interviewers and v) Interviewers. Among RO staff, regional managers and supervisors in the ROs each received 8.5 days while computer operations and technical support staff received 5 days of training. Among field staff, senior interviewers received 8.5 days of training while interviewers received 7.5 days of training. The total number of training hours for 1,050 interviewers and senior interviewers was about 8,000 hours or 36 person years. This figure does not take into account of on-the-job training, which comprised a considerable portion of training.

The options for a training strategy were limited by cost. Given that most of the training was provided to interviewers, cost was minimized by having senior interviewers train their interviewers after receiving training from Head Office. The objective of the training for interviewers was to prepare them to effectively complete LFS interviews using the notebook computer. This included allowing enough time in training sessions to practice new skills. In addition, for those who would be training others, training had to ensure they would be confident in their ability to deliver training. A training package has also been created for new hires who had to learn survey concepts, interviewing skills, and the notebook computer in the same number of training hours as would have been used to train someone for PAPI interviewing. This not only saved training costs but gave senior interviewers an opportunity to thoroughly learn the new tool. The last part of the training given to senior interviewers was one-day session on training skills.

Training for seniors and interviewers consisted of four modules; namely Home Study, Data Collection Classroom, Practice Cases and Practice Assignment. Trainees were able to familiarize themselves with the basic care. They were also able to setup their notebook and modem before classroom sessions began. The hardware was shipped to their home for the three-hour Home Study module which also contained a typing tutorial. Trainees then spent three days in classroom training for the Data Collection module which consisted largely of working through scripted interviews. During the course of the training, trainees were able to observe in group mock interviews, to practice scripted interviews with a partner and

---

to practice their own interview scenarios both with a partner and on their own. Each notebook computer was loaded with a number of practice cases that could be worked through independently of classroom training. Practice cases can be reset thus providing an ongoing training tool on the notebook. The last part of training interviewers was to provide them with a small practice assignment of 15 dwellings. These assignments were created using off-line LFS sample used for the CAI operations field test. Each assignment contained both birth and subsequent dwellings which allowed interviewers to conduct both types of interviews. This also allowed them to practice their new skills in a real interview situation before using the notebooks for the survey.

Two impacts of CAI on the job of the senior interviewer, beyond the change in collection mode, should be noted. First, senior interviewers will complete a small assignment of about 15 dwellings every month in order to practice CAI collection and to keep up to date on software changes. Second, the timeliness and accuracy of information on the progress of each interviewers's collection work has improved. Senior interviewers receive daily reports on their notebooks that give progress as of the previous day of interviewing.

Training for technical support and computer operations staff was designed in two modules: one specific to each group (Training Course for Technical Support and Computer Operations Training) and one common (Technical Support Follow-up Training). The training for the technical support staff gave an overview of the CAI process from the interviewer's notebook computer to HO mainframe computer as shown in figure 2. The overview also covered the life cycle of a case, roles and responsibilities, system maintenance contracts and system updates on the notebooks. The rest of the training was spent covering configuration and installation, system security, operating and recovery procedures, and maintenance procedures. The training for computer operation staff included information on file manipulation, trouble shooting on the PC and the RO Unix. Approximately two months after all interviewer assignments had been converted to CAI, a technical support follow-up training was given to technical support and computer operations staff in each RO. The objective of this common session was to give an overview of the entire process, including the roles and responsibilities of each area.

Regional managers and project supervisors received a training consisting of five modules. Four of them (Technical Support Follow-up training, Home Study, Data Collection Classroom and Train the trainer) were described previously. The fifth module was a Case Management System Overview. This module focused on the RO components that administer and control the CAI field activities. This session was also devoted to hands-on training in the functions that would be used most often by regional staff: assigning sample clusters, running reports, and maintaining the interviewer/senior - interviewer/project manager reporting relationships.

In addition to formal training, the initial surveys done by CAI provided a rich opportunity for on-the-job training, especially for RO staff. Some of the clerical and data entry personnel had opportunity to acquire the more technical skills necessary for production support activities and this became part of the RO technical support group.

---

## **6. PHASING IN CAI**

On the basis of the findings of the Data Quality test, it was decided that the LFS data collection method would be converted to CAI as of fall 1993. During June, July and August of 1993, an operational test was performed. The first objective of the operational test was to ensure that each RO completed a full CAI operational cycle from transmission of the sample to the production, at RO, of edited data. The second objective was further testing of the CAI instrument although time frames did not allow much progress on this objective. The operational test took place in all ROs and consisted of about 125 dwellings a month for 2 or 3 interviewers per RO.

In November 1993, the LFS was ready for this major transformation that would be phased-in during a six-month phase-in. Data collection was converted from PAPI to CAI over a period of six months from November 1993 to March 1994. The factors considered in selecting the phase-in strategy included minimizing the lapsed time between training and using the new mode for the survey. The potential impact of any learning effect on survey estimates was a concern. This concern was dealt with by converting gradually over time, by ensuring sufficient training and practice, by randomly selecting interviewers, and by converting across the country at the same time. Operational complexity and commitments to other surveys were also taken into account. The phase-in strategy also allowed a comparison of data quality indicators obtained with CAI and PAPI, for each month of the implementation period.

It is important to note that some of the benefits of CAI are not realized until some time after the initial conversion. The phase-in period was a time of disruption for LFS field operations. This is evidenced by the data quality statistics presented in the next section. There is a significant period of stabilization required before field operations return to the same level of high efficiency as was enjoyed with paper interviewing. It is only then that the anticipated benefits of CAI related to timeliness and data quality are consistently observed. The stabilization period began in March 1994 and will likely continue for about a year.

### **6.1 Phase-in strategy**

For the phase-in design, interviewers were randomly divided into three equal groups. Interviewers across the country were randomly allocated to these three groups to minimize risks, in particular, the risk of introducing a bias into the estimates. In November, the first group of the interviewers used CAI to complete their assignment. In December, the second group switched to CAI. No further conversions were done in January and February due to unrelated operational requirements. However, the two-month break provided an opportunity to address problems encountered in November and December, and to learn more about CAI. The final third of interviewers moved to CAI in March. Since assignments are more or less the same size of about 75 dwellings, the proportion of dwellings by collection mode during the phase-in period was the same as the proportion of assignments by collection mode. All senior interviewers were converted in the first two months. When an interviewer converted to CAI, his or her whole assignment had to be completed with CAI and that interviewer continued with the computer-assisted method during the succeeding months. Table 6A shows the proportion of the sample that was interviewed using PAPI or CAI in each month survey for Canada.

---

During the phase-in period we instituted a contingency plan, should the laptops breakdown during survey week. For the first several months, all interviewer assignments were printed on the paper forms, sorted into assignments packages and ready for immediate dispatch to the interviewer(s) should they be required. Toward the end of the phase-in period, individual paper assignments were printed on an as needed basis.

**Table 6A**  
**Proportion of the sample interviewed with CAI and PAPI in Canada**

Data Collection Mode	Survey month					
	Oct 93	Nov 93	Dec 93	Jan 94	Feb 94	Mar 93
PAPI	1	2/3	1/3	1/3	1/3	0
CAI	0	1/3	2/3	2/3	2/3	1

## 6.2 Data Quality

During the CAI implementation period, many data quality indicators were monitored and examined monthly by the LFS Data Quality Committee which consists of quality experts, analysts and operations managers. This section presents five quality indicators that were closely analyzed during the implementation: nonresponse rates, vacancy rates, interviewer turnover rates, item quality and conversion rates from CAI to PAPI. These indicators are presented at the national level for November 1993 to February 1994 for the CAI and the PAPI samples. More information about data quality indicators monitored during the phase-in period can be found in Simard and Dufour (1995).

### Nonresponse rates

Figure 6.2A presents the graph of the national nonresponse rates broken down by nonresponse component for CAI versus PAPI for November 1993 to February 1994. Five components are presented: Temporarily absent, No one at home, Refusal, Technical problems (i.e. no form received due to transmission problems) and Other category (which represents households where the interview could not be conducted for other reasons). The overall nonresponse rate is obtained from the combination of these components.

The CAI overall nonresponse rates were higher by 1.0 to 3.4 percentage points. PAPI nonresponse rates from November to February varied from 4.0% to 5.0%. Most of the difference between CAI and PAPI is explained by an increase in nonresponse that is due to technical problems (such as hard disk failures, insufficient memory allocation, transmission problems, etc.) rather than to any increase in the number of persons who were not at home, temporarily absent, or who refused to participate. If the technical problems are eliminated from the overall nonresponse rate, the average rate for CAI is 4.9%, which is close to the rate for PAPI (4.7%). Furthermore, average nonresponse rates for the LFS during the



corresponding months in 1991 and 1992 were similar at about 4.8%. Consequently, it seems that technical problems influence the overall nonresponse rate. However, when all technical problems are addressed, it is anticipated that the nonresponse rates under CAI will be similar to those under PAPI.

The impact of CAI on the refusals was a concern due to the presence of a notebook computer at the respondent household. Many felt that respondents would be more reluctant to respond because of privacy concerns, even during subsequent telephone interviews. This was found not to be the case. The refusal rates were stable at 1.2%, regardless of the method of collection. Furthermore, in January and February 1994, the refusal rate was slightly higher for PAPI than for CAI (1.3% vs. 1.1% and 1.5% vs. 1.4% respectively). The LFS refusal rate is usually around 1.5%.

### NATIONAL NONRESPONSE RATES BY SURVEY MONTH – CAI VS PAPI

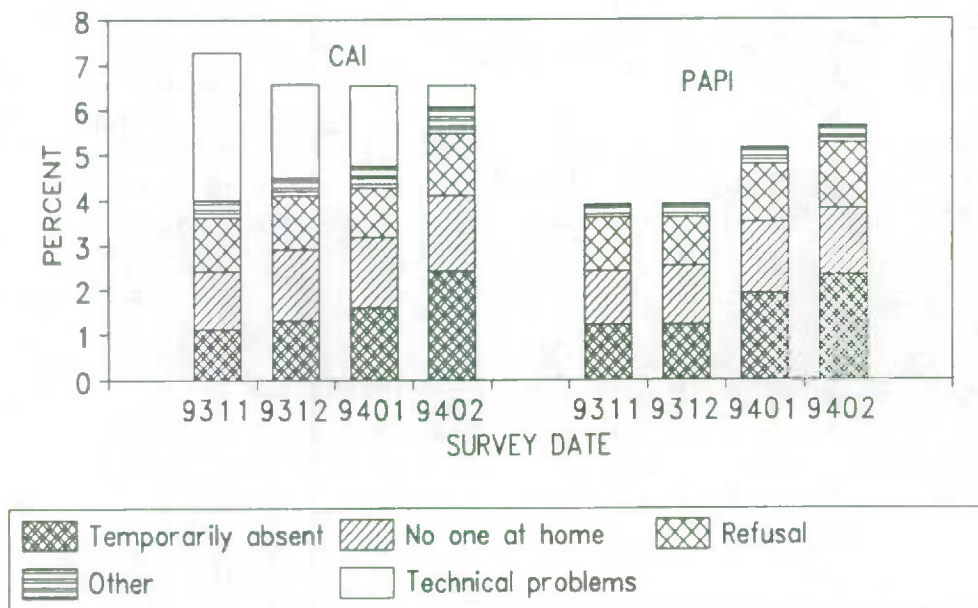
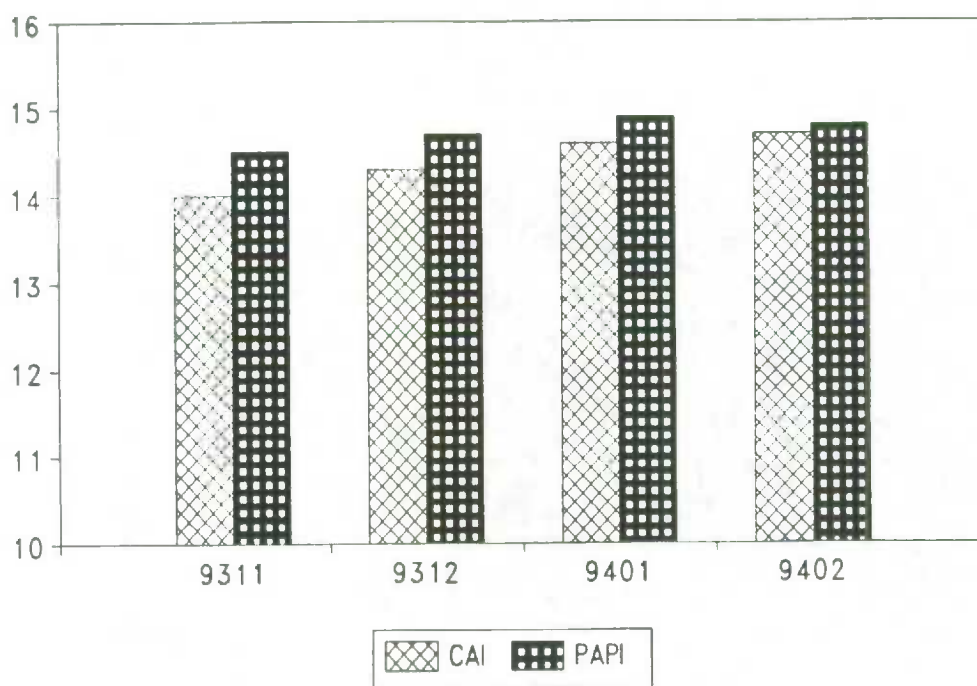


Figure 6.2A

The proportion of "Technical problems" was at 3.3% during the first month of the implementation but ended the implementation period at 0.5%. Such a trend was expected because of the learning curve of the interviewers, Regional Offices and the Head Office. Many tests were performed, but problems still occurred during the actual survey. It was found that the import/export function of the Case Management system, which transmits cases from the ROs to the interviewers and the reverse, is mainly responsible for the technical problems. This function will be rewritten.

### NATIONAL VACANCY RATES BY SURVEY MONTH - CAI VS PAPI



**Figure 6.2B**

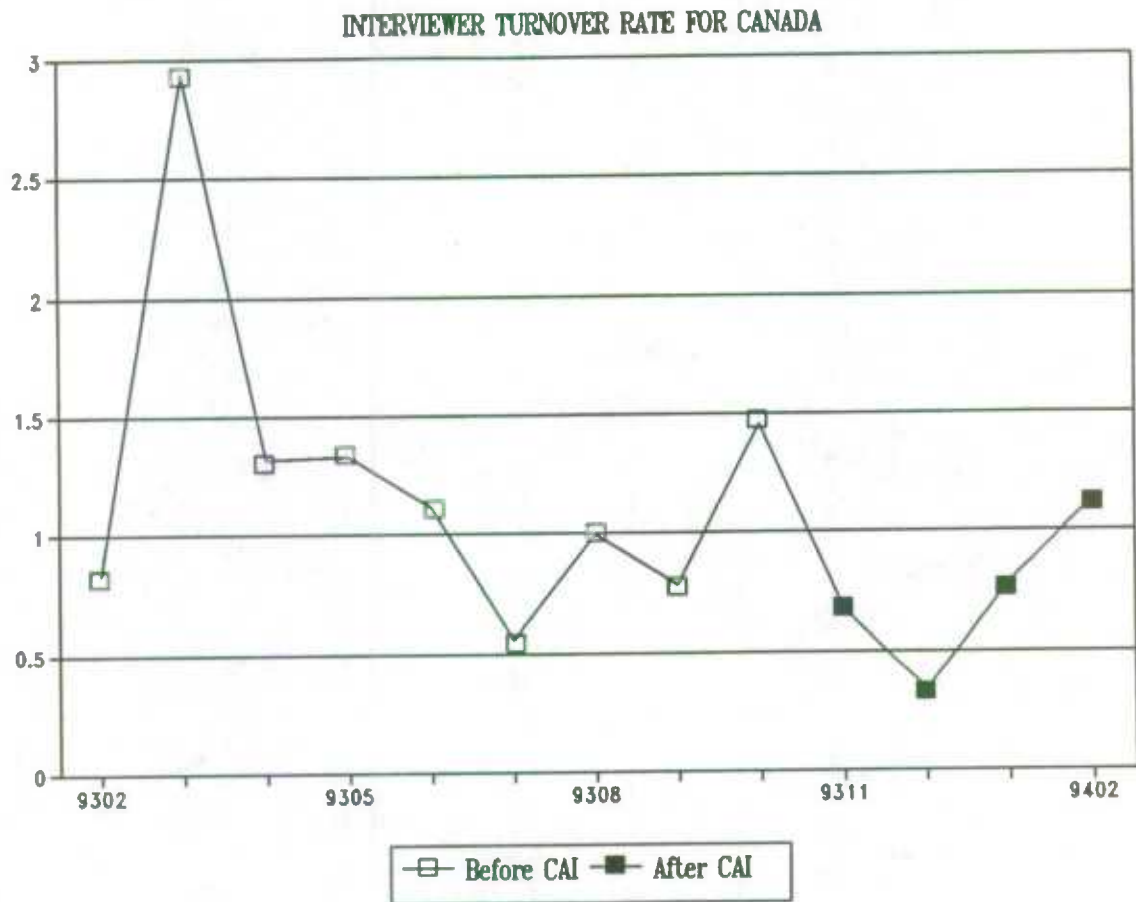
#### Vacancy rates

Figure 6.2B shows the vacancy rates for CAI and PAPI for November 1993 to February 1994. For both samples, the vacancy rates were stable over the four-month period, ranging between 14.0% and 14.9%. For the CAI sample, a slight upward trend was observed during the phase-in. A similar pattern was seen

for the PAPI sample for the first three months. However, while the vacancy rates for the CAI sample were systematically lower than the PAPI ones, the small differences between the two decreased over the course of the implementation period. During the quality test, the opposite was observed. No specific cause for the difference has been identified.

**Interviewer turnover rates**

The interviewer turnover rates are produced to monitor survey data quality. During the CAI implementation, this measure was monitored to evaluate the interviewer acceptance of CAI. Interviewers had to deal with a major change in their work methods. They had to become familiar with new tools such as the keyboard, the portable computer, and the various computerized procedures (saving data, charging batteries, transmitting by modem). They also had to adapt their interviewing style to the requirements of CAI.



**Figure 6.2C**

Figure 6.2C presents the interviewer turnover rates twelve months ending February 1994. The graph shows 12 months as opposed to just 4 for two reasons: i) it was not possible to make the distinction between PAPI and CAI interviewers during the implementation so data were combined for a comparison at an overall level, and ii) it was important to show the seasonality, so the range of the rates does not lead to an erroneous conclusion. Consequently, it is shown that for the last three months, the interviewer turnover rate was increasing from 0.3% in December 1993 to 1.1% in February 1994. However, this trend is not alarming, since 0.3% has been the lowest rate since the last LFS redesign and 1.1% is in the usual turnover range. The peak observed in the last year (March 1993) coincides with a sample size decrease. Hence, the implementation of CAI has not resulted in any major change in turnover rates so far.

### Item quality

One of the main promises of computer-assisted interviewing is an improvement in item quality. There are several reasons for such improvements. First, it is the CAI application that determines the next question to be determined rather than the interviewer; thereby eliminating human error. The second source of item quality improvements is on-line editing which allows interviewer and respondent errors to be corrected at source. On-line edits for consistency were built into the CAI application on both the Demographic (FO3) and Labour Force Activity (FO5) questionnaires. The edits built into the Demographic questionnaire included most of the consistency edits performed at Head Office on PAPI questionnaires while, due to operational time frames and respondent burden, only a subset of the FO5 consistency edits were programmed into the FO5. On-line editing does not mean that data is necessarily clean when it returns from the field since interviewers can choose not to change data items when an edit failure occurs. It should be noted that during HO processing all of the edits that were performed on the PAPI data were also performed on the CAI data. This was done to minimize redevelopment costs but also protected against any short-term human or technical problems from adapting to CAI.

Improvements to *discrepancy rates* during the phase-in period are marginal. Item quality for LFS data under PAPI has traditionally been quite high. The average discrepancy rate from April 1991 to March 1993 was 0.1% for the Demographics questionnaire (F03) and 0.9% for the Labour Force Activity questionnaire (F05). Table 6.2C gives the percent of all items that failed editing at Head Office. From December to February, discrepancy rates were lower for CAI sample than for PAPI sample. Rates for CAI sample ranged from less than 0.1% to 0.1 for the demographic questionnaire (Form 03) and from 0.7% to 1.1% for the labour force activity questionnaire (Form 05). Rates for PAPI sample ranged from 0.1% to 0.2% for the demographic questionnaire (form 03) and from 0.9% to 1.2% for the Labour force activity questionnaire (Form 05).

---

**Table 6.2C**  
**Discrepancy Rates at the Canada level**  
**CAI vs. PAPI - Phase-in months**

Phase-In Month	Type of questionnaire			
	Demographics (Form 03)		Labour Force Activity (Form 05)	
	PAPI	CAI	PAPI	CAI
Nov 93	0.1	0.1	0.9	1.1
Dec 93	0.1	0.0	1.0	0.9
Jan 94	0.2	0.1	1.2	0.7
Feb 94	0.2	0.1	1.2	0.7
Mar 94	NA	0.1	NA	0.7

**Conversion rates**

The use of CAI as the data collection method introduced new problems that the LFS had not previously experienced. Notwithstanding the power of the new technology, every month of the phase-in period ushered in a new set of technical problems. Whether the causes lay with the notebook computers, the hard disks, the modems or the telephone lines, the result was the same: some data were not available for processing on time. To minimize this problem, PAPI was used as a contingency plan. After March 1994, it was still possible to go back to PAPI but only in some particular circumstances. The PAPI mode is being kept as a contingency plan until all problems are reasonably resolved.

The conversion rates were designed to monitor the performance of the CAI mode. This rate represents the number of dwellings assigned to be interviewed by CAI that were converted to PAPI. During the phase-in, complete or partial interviewer assignments were converted to PAPI if technical difficulties were encountered.

Table 6.2D gives the national conversion rate from CAI to PAPI by survey month during the phase-in and for March 1994, the first month of full implementation. The conversion rate varied during the implementation from 2.9% to 5.7%. The decrease recorded in January was expected since no new sample was converted to CAI in January. Hence, the increase observed in February was surprising because the CAI interviews in January and February were conducted by the same interviewers. Consequently, it seems that the volume of conversions is not only related to the lack of interviewer

experience but to a multitude of other factors. The lowest conversion rate was reached in March when the entire sample was done in CAI mode (a rate of 1.0% represents 668 dwellings) which indicates the collection mode is working well.

**Table 6.2D**  
**Percentage of Dwellings Converted From CAI to PAPI**

	SURVEY MONTH				
	Nov 93	Dec 93	Jan 94	Feb 94	Mar 94
Conversion rates	3.0%	3.4%	2.8%	5.7%	1.0%

**7. EVALUATION AND STABILIZATION PERIOD**

With the completion of the phase-in came the evaluation of the conversion process and the new collection vehicle. Conversion to CAI was a period of major disruptions for collection operations. The main sources of these disruptions were technical problems. Hence, the benefits of CAI related to timeliness, cost, and data quality were not immediately realized during the phase-in. To put CAI experience in perspective, 204 PAPI surveys have been done using the same collection methodology and questionnaire, while with CAI only a few surveys have been completed since conversion. In a few months, after a period of stabilization, the benefits of CAI are expected to be realized. In this section, an evaluation of CAI and resulting enhancements are outlined.

As described in section 5.1, good communications was emphasized and a communication strategy was developed. In general, the information conveyed was consistent and timely. There were minor unanticipated problems, which were rectified by improving and/or streamlining communication. One challenge was that the technical support staff and field interviewers did not share the same technical vocabulary. Initially, once a problem was reported, support staff sometimes took a day to get back to the interviewer with a solution to the problem. With a six day window to complete their survey assignment, interviewers required a quicker response mechanism for technical problems. Over time both of these situations have been resolved. Having overcome the learning curve, interviewers are now both more familiar with the hardware/software of the notebook and are able to communicate their problems. With a streamlined communication process, problems are logged and responded to within a few hours of the problem being reported.

Technical problems during transmission of data had a serious impact on data quality as can be seen by an increase in nonresponse. During the conversion period, an increase in the number of cases still outstanding at the end of the collection period was responsible for an increase of 5% to 25% of overall nonresponse. This problem is largely believed to be the result of hardware and software failings in moving data from one platform to another. Otherwise, no differences in nonresponse patterns have been

observed. Specifically, refusal patterns have not changed. Although the overall nonresponse rate was affected adversely, the occurrence of missing or invalid data has decreased. This represents a marginal gain in data quality since item quality rates were already high under PAPI.

No gain in timeliness has been realised yet as a consequence of transmission problems. There may be some gains in moving data between the interviewer and the RO, but it has taken longer to reach the point of having sufficient data to produce survey estimates.

On the average, it takes longer to do an interview with CAI than it did using PAPI. Although the interviewer's learning curve and the requirement to follow paths precisely play a role, the main cause for the increased time is the time that it takes the notebook to move between modules. Field staff experienced frustrating time lags in moving between components of the questionnaire. This has increased respondent burden and challenged the interviewer's skills.

It is expected that the costs of recreating the same quality of field work with CAI will outweigh any cost savings from eliminating data capture during the stabilization period. Also, the cost of retraining can be anticipated to be more than the cost of formal training since a considerable amount of learning is done on-the-job.

Software has been put in place to address problems with transmission and time lags between components of the application. This has reduced the length of the interview, cost, and respondent burden, as well as improve timeliness and quality. In addition, some lessons learnt from conversion have been put into practice with the software rewrites. After several problems were reported, it was noticed that interviewers do not necessarily read everything that is on the notebook screen. In the rewrites, instructions and error messages are simple, user friendly and easy to understand.

New software to reduce the length of the interview was implemented for the September 1994 survey. Substantial improvements have been realized. A significant period of stabilization following the period of conversion is to be expected before the benefits of CAI related to timeliness, reduced cost, and improved data quality are fully realized, and before field operations return to the same level of high efficiency enjoyed with PAPI.

---

## **8. FUTURE PLANS**

The software rewrite to address transmission problems related to lost cases is scheduled to be implemented in the summer of 1995. This should improve nonresponse. There is now a staff of trained and experienced CAI interviewers to undertake other upcoming CAI household surveys.

The next major development for LFS collection will be the introduction of the new questionnaire early in 1997. The questionnaire will take advantage of CAI as well as contain improvements and new content.

---



## GLOSSARY

### Non-interview:

During survey week every month, approximately 1,000 interviewers and 100 senior interviewers enumerate all dwelling in the LFS sample. Each interviewer is assigned a set of dwellings and makes every reasonable effort to interview all eligible persons who are members of households belonging to these dwellings.

For various reasons, interviewers are not able to obtain an interview at every dwelling. These non-interviews occur for the following reasons:

Household nonresponse: including reasons such as refusal, no one at home, household temporarily absent, no interview due to weather conditions, no interviewer available, no interview due to circumstances within the household or technical problem due to the CAI instrument.

Vacant dwellings: including unoccupied dwellings, vacant seasonal dwellings, dwellings under construction and dwellings occupied by persons not eligible to be interviewed.

Non existent dwellings: including dwellings which were demolished, converted into business premises, moved, abandoned or listed in error.

### Nonresponse rate:

The nonresponse rate represents the number of nonrespondent dwellings divided by the number of dwellings in the sample expressed as a percent.

### Vacancy rate:

The vacancy rate represents the number of vacant dwellings in the sample divided by the number of existent dwellings in the sample.

### Edit failure rate:

The edit failure rate represents the total number of Forms O3 (or DEMOG) or Forms O5 (LFS) with at least one discrepancy as a percentage of the total number of FO3s, or FO5s, completed.

### Edit discrepancy rate:

The edit discrepancy rate for a specific form is obtained by dividing the number of discrepancies for the form by the total number of entries for the form and then multiplying by 100. A discrepancy between original and edited item response occurs when one or more edits change the original value.

---

**Conversion rate:**

The conversion rate represents the proportion of interviews that should have been conducted with CAI but, for one reason or another, were conducted with PAPI.

**Interviewer turnover rate:**

The interviewer turnover rate is the percentage of interviewers who terminate their employment and who will be replaced.

**Automated coding rate:**

The automated coding rate represents the proportion of records that are coded automatically by the automated system for the following three variables: industry, occupation and class of worker.

---



## REFERENCES

- Bench, J., Clark, C., Dufour, J. and Kaushal R.** (1994). *Implementing Computer-Assisted Interviewing for the Canadian Labour Force Survey*. Proceeding of the 1994 ASA Conference.
- Catlin, G. and Ingram, S.** "The Effects of CATI on Cost and Data Quality". *Telephone Survey Methodology*, edited by R.M. Groves et al., New York: John Wiley and Sons.
- Dowd, D.** (1992) Labour Force Survey CAPI Implementation - Training Strategy. Internal Report, Survey Operations Division, Statistics Canada.
- Kaushal, R. and Laniel, N.** *Computer Assisted Interviewing Data Quality Test*. Proceedings of the Bureau of the Census 1993 Annual Research Conference. pp 513-524.
- Kaushal, R. and Laniel, N.** *Overview of Computer Assisted Interviewing Data Quality Test*. Branch Working Paper (1995) .
- Simard, M. and Dufour, J.** (1995) *Impact of the Introduction of Computer-Assisted Interviewing as the New Labour Force Survey Data Collection Method*. Quality Report, Household Survey Methods Division, Statistics Canada.
- Singh, M.P. et al.** *Methodology of the Canadian Labour Force Survey 1984-1990*. Statistics Canada Publication. Catalogue number 71-526.
-