# Direct Data Entry Using Microcomputers—A Travel Survey Pilot Project

GERALD N. STEUART, BRUCE MORI, AND PETER J. NOEHAMMER

**One aspect tested in the pilot project for a Toronto region origin-destination travel survey was direct data entry using microcomputers. Four interviewers utilized a computer program that performed (a) file management of the sample, and (b) data prompting for the interviewer. Twelve interviewers used conventional pencil and paper forms. A comparison of the two methods proved the automated process just as expedient as the manual method during actual telephone interviewing, but far superior during sample processing. Large potential benefits from direct data entry are possible after interviewing is complete as paper forms must further be entered onto a computer by keypunch. The costs associated with providing each interviewer with a microcomputer and program balance the costs of keypunching paper forms. No extra cost or risk over reliable manual methods is evident for direct data entry to automate the survey process. Furthermore, productivity gains during interviewing provide the potential for direct data entry to surpass conventional manual methods in terms of cost-effectiveness.**

Transportation planning in virtually every community in North America has been formulated on the basis of one, and usually only one, large-scale and comprehensive inventory of urban travel. Most of these studies were carried out in the 1960s using home interviews as the data collection technique. Information regarding all trips made by each member of a household during a specified 24-hr period was recorded. The resulting travel data have become the foundation on which much of transportation planning is based.

Toronto is typical of this historical development. The first and only comprehensive inventory of urban travel in the Toronto region was the Metropolitan Toronto and Region Transportation Study (MTARTS) Home Interview Survey, which was carried out in 1964. This information has served as the basis for most of the large-scale transportation planning efforts since that time. After approximately 20 years, the information is of questionable value as a basis for estimating current urban travel. The data will always be useful for looking at time trends; however, to serve this purpose new origin-destination data must be collected.

All transportation planning agencies in the Toronto region have realized the inadequacies of 20-year-old information and the limitations of incomplete information. They have formed a committee to promote a coordinated effort to improve the situation. The Toronto Area Transportation Data Collection Steering Committee has the objective of determining the best collective course of action to improve the quality of travel data in the region. The committee has recommended a comprehensive origin-destination (O-D) travel survey for a universe of approximately 4 million people encompassing six regional municipalities in the Toronto area.

The Ontario Ministry of Transportation and Communications financed a pilot project *(1)* in March 1986 aimed at testing several proposed features of a full-scale study. Primary features tested in the pilot project were a proposed set of questions to obtain pertinent household, person, and travel information; a procedure for informing the sample of an impending interview; a test of the sampling procedure; and most important from the standpoint of this paper a test was carried out in which some of the interviewers used microcomputers to enter travel data directly to computer files. Interviews were conducted March 4–9 and 18–20, 1986.

## DIRECT DATA ENTRY

The objective of the microcomputer experiment was to verify that direct data entry could be used in a travel survey. In this paper, the findings of an experiment in which microcomputers were used by an interviewer to enter travel information directly to a computer file while a home interview was taking place over the phone, that is, direct data entry, are reported. Observations and conclusions are aimed specifically at the feasibility of using direct data entry for a full-scale O-D study.

The data collection process in a large-scale O-D survey, whether performed by direct data entry or by more conventional manual methods using paper forms, contains three distinct phases:

1. Sample processing. Assemble a sample of the universe of households and continuously update the status of each record in the sample;
2. Interviewing. Select a record from the sample, prompt an interviewer with questions to be asked, and record the data;
3. Data processing. Check the answers received for validity, assign a numerical code for spatial locations, and record all information in computer files.

These phases are closely related, as the output of one serves as input to the next. O-D surveys in the past have done all record keeping and data recording on paper forms. This method requires an extensive filing system for managing the paper forms, both completed and to be completed.

Conceptually, the entire process could be automated on a computer system with a network of terminals on either a

Department of Civil Engineering, University of Toronto, Toronto, Ontario, M5S 1A4, Canada.

mainframe or minicomputer. However, the process must operate without interruption. Unfortunately, the development of a large-scale computer system with multiple on-line users, including networking of terminals, is usually subject to a breaking-in period during which the system is unreliable. It is required that the system development costs be rationalized for one short application. This process normally implies the use of existing software packages.

A better strategy is to use the computing device to aid or improve already proven techniques without taking the risks associated with developing a complex computer network in a short period of time. The most common and most widely used software packages for data management are available for microcomputers. An added incentive for developing microcomputer procedures is the potential for use in other cities. Other O-D studies have a better potential for access to compatible microcomputers than identical minicomputers or mainframe computers with compatible operating systems.

The rate at which interviews are conducted can be adversely affected by computer system failures and operator errors. Thus, contingency plans and error recovery strategies are important components of this project. In addition, the computer hardware or software should never dictate the pace of the interview. The computer program should not jeopardize or hinder the rate at which an interviewer can process the sample. The computer screen should merely provide the electronic counterpart of a paper form, and the program should handle sample updating and database management tasks normally done manually by filing forms.

During the pilot project, all aspects of the actual interviewing were the responsibility of a market research firm. This included the hiring, training, and supervising of all interviewers. Approximately 20 interviewers were trained for the project and 5 were selected for direct data entry. Four of these were regular interviewers, each one using the same computer and sample every session, while the fifth became the supervisor. Instruction on the use of the computer was carried out by the authors of this paper. In addition, project members were available in person early in the project to detect problems. On the basis of these observations and input from the interviewers, several minor modifications were made to the computer program during the course of the pilot project.

The computer program was written in a popular database language available for microcomputers operating under MS–DOS. The program was then compiled into an executable program for distribution. This step allowed the necessary features of the database language to be used without having to purchase a program license for each computer, and did not infringe on copyright regulations. The program was configured for an IBM-compatible personal computer with at least 256 kilobytes of random access memory, a monochrome monitor, and two 360-megabyte floppy disk drives. The interviewer was given two disks every evening. One disk contained the program and the interviewer's sample file, with a record of all attempts to contact each household. A record was ignored after five unsuccessful contact attempts. The second disk recorded all the data collected during the course of interviewing and was replaced every evening. An update of the global database could be made every day.

## SAMPLE PROCESSING

The first phase in an O-D survey is to assemble a sample of the universe of households in the study area, assuming the study area has been defined and a sample size selected. The sample should be obtained in the form of a data tape and loaded onto the appropriate file management system in random order with sequential sample numbers. The first task of file management is to provide input to the printing of preinterview letters. It is possible to automatically print for mailing out an appropriate number of preinterview letters containing a current date once the sample is set up in a computer file.

The pilot project was conducted on a sample of households from eight postal districts consisting of approximately 5,000 records. This sample was selected sequentially from the current telephone directory by Tele-Direct, a subsidiary of Bell Telephone, and prepared on magnetic tape. Every fourth record was copied to a separate computer file, which became the sample for the direct data entry experiment. The remaining records were to be processed using conventional paper forms.

The sample was divided into two sets. Preprinted letters from one set were sent out on Friday and Saturday, March 1 and 2, 1986. These letters notified households of interviews to be conducted March 4–9, inclusive, for travel on March 3–7, inclusive. Interviews were not conducted during the following week, which coincided with the public school's March break. The second set of letters was to have been mailed out on March 11; however, a sufficient sample was available from the first mailing. As a result, the second mailing for the remaining 3 days of interviewing was never carried out. Some adjustments to the sample allocation were necessary because direct data entry processed records faster, thereby leaving insufficient sample for three remaining evenings of interviewing. One in every four records remaining in the paper form sample was transferred to the direct data entry sample. The expanded sample proved advantageous because the additional 220 transferred records could be distributed according to the needs of each interviewer. Although the interviewers were initially assigned the same number of records, each processed at a different rate.

There is no reason to believe that direct data entry will decrease the length of time an interviewer spends actually conducting interviews on the telephone. Productivity gains, if there are to be any in direct data entry, are likely to come from efficiently handling the records in the sample and automatically sequencing the calls for the interviewer. For this reason, considerable effort was spent on file management by the program developers.

Two aspects of automatic file management of the sample records improved interviewer productivity and improved the likelihood of making contact with the respondent. Strategies for attempting repeated contacts were automated for "no answer" and "busy" situations. "Call backs" were used to accommodate respondents at a more convenient time, or to obtain complete information from a member of the household who was not at home. For a "call back" situation, the record and an optional message were automatically redisplayed for the interviewer at a prespecified combination of time and day when interviewing was taking place. "No answers" would automatically reappear on the screen, either the same evening or the

next evening, after a specified period of elapsed time. Noninterview time consumed 66 percent of the interviewers' time for the conventional paper entry mode of Operation 1. File management by computer has the potential to reduce this time, and reduce the average time per completed interview by as much as 2 min. Furthermore, the interviewer would spend more time interviewing and less time handling forms.

Management of the sample files began with the preparation of the database file containing the sample that included a series of fields from the information provided by Tele-Direct:

- Respondent's name
- Street address (and apartment number, if applicable)
- Municipality
- Postal code
- Telephone number

Each record was then completed by adding the following fields:

- Sample number
- Five repetitions of three fields (initially blank or zero)

  Data of response
  Time of response
  Response code

- Interview completion time

An evening's interviewing begins with the program's searching the sample file for the first record that has not been completed. It subsequently displays the information on the screen in the form shown by Figure 1. The left-hand side of the form contains a display of the information for all the fields of that record. As such, the form displays a complete history of any previous action. The upper right-hand side of the screen displays the current date and time taken from the computer's internal clock. The box on the lower right contains a reminder of all the valid actions the interviewer can take on the sample by depressing the appropriate key. These actions, or responses, were taken only after the interviewer had dialed the telephone number displayed.

One additional precautionary step was taken in the pilot project. The sample number, respondent's name and telephone number, and the response were written down on paper for every telephone call made by the interviewer. This procedure was done primarily for supervisory staff records, but also served as a precautionary measure in the event of a computer malfunction.

Response 1 was assumed to be the most common response and given the first position. This response was designated if the respondent confirmed that the name and address were correct and that they were willing to be interviewed. The record was then internally marked in a manner to prevent it from reappearing on the screen.

Response 2 was designated if the telephone was not answered after an appropriate number of rings. The final version of the program, used in the last few days, attempted to patch a strategy onto the existing program structure that would redisplay "no answer" records 2 hr later.

Response 3 was designated if the respondent, for whatever reason, refused to be interviewed. The record was then internally marked in a manner to prevent it from appearing on the screen again.

Response 4 turned out to be a useful device enabling the sample record to reappear on the screen at a designated time. A comment field, which recorded information to be presented on the screen to remind the interviewer of details regarding the call back, was also incorporated at the request of the interviewers. The procedure operated as follows when Response 4 was selected. A new screen appeared asking the interviewer for a date and time of the call back. A default value of the current date was displayed and could be overwritten. The interviewer was then given the option of entering three short lines of comments. The screen then displayed the next sample record to be interviewed. The record registering a call back came back on the screen at the appropriate time, as the program searched the call back file for dates and times.

Response 5 was designated if the operator came on the line to advise that the number was not in service, the number represented a business rather than residential phone, or the respondent had moved out of the study area but kept the same telephone number. The record was then internally marked in a manner to prevent it from reappearing on the screen.

Response 6 allowed the interviewer to change the informa-

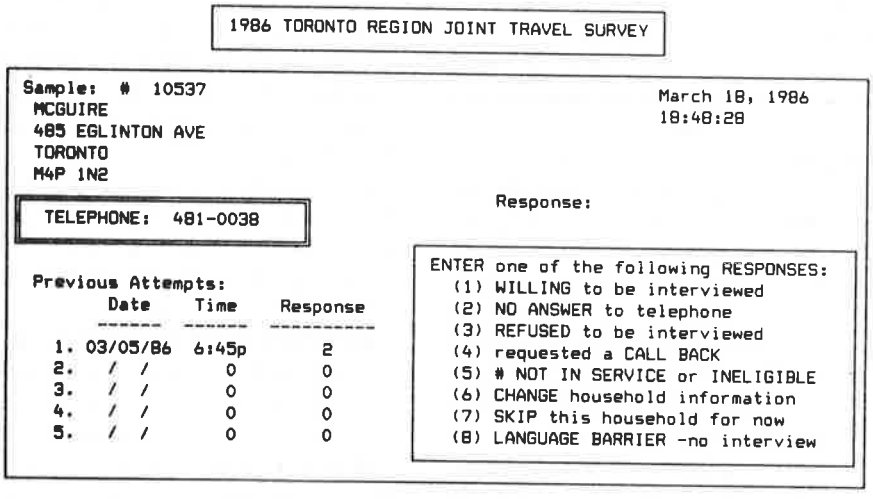


**FIGURE 1 Direct data entry—Screen 1 format.**

tion in the fields describing the name and address of the respondent. This action was important because the home address was stored in memory and became part of the trip record. When the editing was complete, the original screen returned and the interviewer could select the appropriate response to proceed.

Response 7 allowed the interviewer to skip the current sample record. The record would not appear until the computer was turned off and the start-up procedure carried out.

Response 8 was designated if the interviewer encountered a language that was indeterminable, or if the pool of interviewers did not include anyone with sufficient knowledge of the language to conduct an interview. The record was then internally marked in a manner to prevent it from appearing on the screen again.

Every action an interviewer took on a record in the sample was timed to the nearest minute by the computer's internal clock and recorded in the sample file. The recorded data showed file management to be important. Of the 142 hr that interviewers spent processing sample records, just less than 67 hr were spent entering trip records on the basis of information gathered over the telephone. The remaining 75 hr were spent attempting calls, encouraging interviews, and doing other file management tasks.

## INTERVIEWING

Computer specialists immediately visualize a serial process whereby the interviewer, when telephone contact is established, would be presented on a visual display terminal with the name of the potential respondent, then with a series of precisely worded questions each followed by a prompt for data entry. Each entry would be checked for validity, then either rejected and reprocessed or stored in a database file. Unfortunately, travel survey interviews seldom progress sequentially as answers are provided in a conversational form as opposed to a rigidly structured form. The interviewer is used to, and benefits from, seeing the complete household's travel behavior rather than simply a portion of a single trip. Travel survey interviewers have always been accustomed to seeing the evolution of a household's complete trip-making behavior on a single form. This procedure provides information about duplicate trips by different members of the household and allows the interviewer to recognize previous errors and to move directly to correct them. Interviewers are also better able to encourage information and help the respondent along because they are fully aware of which trips have been made. Should a series of conventional, sequential forms be used, many errors would unfortunately go unnoticed, and much would be left to the memory of the interviewer to go back to the intended form.

This approach, however, requires that interviewers be conversant with the fundamentals of microcomputer operation and be able to type about as quickly as they can write. Experienced interviewers in market research have traditionally used a pencil and paper. Fortunately, there are two compensating forces at work. First, interviews are conducted in the evening, allowing for a larger selection of experienced interviewers, some of whom will be able to type. Second, there is a growing interest in the field of market research for direct entry of information at a computer terminal. This fact implies that more experienced interviewers will be able to use direct entry and that many others are interested in developing the skill.

Another implication is the need for an interviewer to be able to move freely and quickly from field to field on the interview form. The interview form will be large and complex by database management standards and is not of fixed length. The length of the form will depend on the number of persons and the number of trips made by each person. Most database management software allows the operator to move sequentially from field to field; however, moving on the screen to another section of a large and complex form of unspecified length may be time-consuming. Training an interviewer to move the cursor about on the screen with the use of key strokes and to bring up a new screen when required must be accomplished in training sessions.

The screens designed for this project proceed in the following manner. When an interviewer selects Response 1, indicating a respondent is willing to be interviewed, a new screen appears on the monitor as shown in Figure 2. The screen represents two database files. The box on the left contains all information to be collected about the household (the number of persons and number of automobiles). The screen prompts for this information and when entered stores the information in the household database with the sample number as an identifier.

The screen then displays, one line at a time in the right-hand box, a request for information about each person in the household. The computer automatically numbers every person and requests a name or other identifier for the interviewer to refer to that person (free format, eight characters maximum), approximate age (two-character numeric with 99 indicating refusal), sex (M or F), employment status (P part time; F full time; S student; H home; U unemployed), and driver's license (Y or N). A 9 entered in any field denotes a nonresponse or unknown information. The interviewer is able to edit any of the information about any person while in this box; however, the interviewer cannot change any information in the previous box. Each field is checked for valid information according to the indicated allowable codes. When the interviewer enters information for the last person correctly, all information remains on the screen and a request for trip information is presented as shown in Figure 3.

The first line of trip entry begins with a numeric identifier for the trip, which is automatically assigned by the computer, followed by a numeric field for person number. The origin of the first trip taken by the person in question is recorded in a

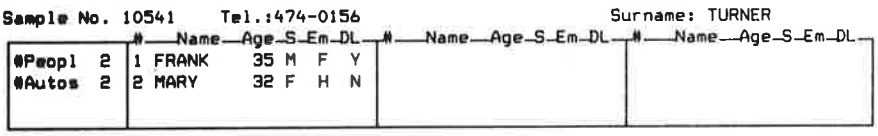


**FIGURE 2 Direct data entry—Screen 2 format.**

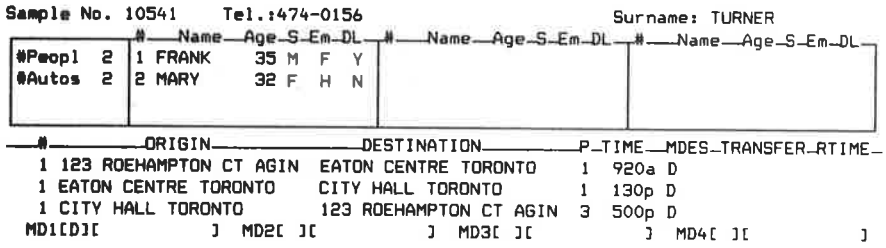


**FIGURE 3 Direct data entry—Screen 2 format, including trip records.**

free format field of 22 characters maximum. The address information of the home location was previously stored in computer memory and can be recalled anytime by the interviewer by simply selecting Function Key F9 on a standard keyboard. The screen then prompts for the destination of the trip, which is also entered into a free format field of 22 characters maximum. The program automatically stores this field in computer memory and defines Function Key F10. This action facilitates the use of F10 on the next trip to recall the previous destination, as it was often the origin of the next trip. The screen then prompts for trip purpose, which is a numeric code selected from a table presented on a card and usually placed on top of the monitor. The next field to be entered is the starting time of the trip using a 12-hr clock with a.m. and p.m. designations (the field is four numeric characters followed by one alphabetic character, limited to a or p). A 12-hr clock was chosen over the conventional 24-hr clock used in most travel surveys because it was more convenient for the interviewers and would also eliminate any errors made while converting from a 12- to a 24-hr clock. The request for entry then moves to the second line to record modal information on every segment of the person's trip. Trips are allowed a maximum of four segments made up of any combination of the available modes. Information is entered beginning with a single character field designating the mode

B Bus
S Subway
G GO train (regional commuter system)
V Via rail
D Automobile driver
P Automobile passenger
T Taxi
C Bicycle
O Other
9 Unknown or not available

followed by a free-format field of 10 characters maximum for a description of a public transit mode, such as a bus name or number. The transfer field is an eight-character free-format field used to record the location of transfers between public and private modes of transit. This field was included at the request of the local transit authority. The last field is used to record the return time, when appropriate, in the same format as the previous time field. A return time is entered if the following trip is identical to the current trip but in the reverse direction. The fields can be selected out of order with the use of the keyboard's arrow keys. The last field can always be reached in one step by depressing the END key on a standard keyboard.

The efficient use of screen space is demonstrated by the manner in which trip information is first recorded and then displayed as subsequent trips are recorded. When the interviewer moves on to recording the next trip, which is automatically done by depressing the RETURN key from the last field on a line, the next trip entry form overwrites the second line of the previous trip entry. However, at the same time, the four single-character designators of mode for each segment of the previous trip are displayed in a space that was left vacant beside the departure time field. In this manner the interviewer has all information about every trip displayed on one line, with the exception of the descriptions of the modes.

Trip information can be edited at any point in the process by selecting Function Key F1. A prompt asks, "Which trip?" The operator then selects the trip number to be edited and enters the new information. Function Key F1 is also used to move to the next household sample record when required and to exit the program when an evening's interviewing is complete.

TABLE 1 DIRECT DATA ENTRY—SUMMARY OF INTERVIEW TIMES

AUTOMATIC DATA ENTRY
Number of Interviews and Average Interview Times

| Interview Day | Station 1 No. | Station 1 Avg. | Station 2 No. | Station 2 Avg. | Station 3 No. | Station 3 Avg. | Station 4 No. | Station 4 Avg. | All Stations No. | All Stations Avg. |
|---|---|---|---|---|---|---|---|---|---|---|
| 04 | 10 | 8.50 | 9 | 10.33 | 6 | 18.00 | 16 | 10.00 | 41 | 10.88 |
| 05 | 16 | 7.94 | 9 | 13.11 | 8 | 12.75 | 8 | 9.88 | 41 | 10.39 |
| 06 | 18 | 5.22 | 13 | 7.54 | 9 | 14.11 | 17 | 5.76 | 57 | 7.32 |
| 07 | 6 | 9.67 | 6 | 9.67 | 4 | 16.00 | 6 | 4.67 | 22 | 9.46 |
| 08 | 19 | 7.74 | 16 | 7.75 | 11 | 11.55 | 20 | 7.60 | 66 | 8.34 |
| 09 | 19 | 6.89 | 13 | 7.23 | 8 | 9.25 | 12 | 9.17 | 52 | 7.86 |
| 18 | 25 | 4.88 | 15 | 7.67 | 7 | 18.00 | 16 | 6.88 | 63 | 7.51 |
| 19 | 23 | 5.39 | 18 | 8.61 | 12 | 12.08 | 18 | 7.44 | 71 | 7.86 |
| 20 | 18 | 4.89 | 15 | 6.00 | 11 | 11.64 | 12 | 10.29 | 56 | 7.67 |
| All days | 154 | 6.49 | 114 | 8.42 | 76 | 13.04 | 125 | 8.20 | 469 | 8.48 |

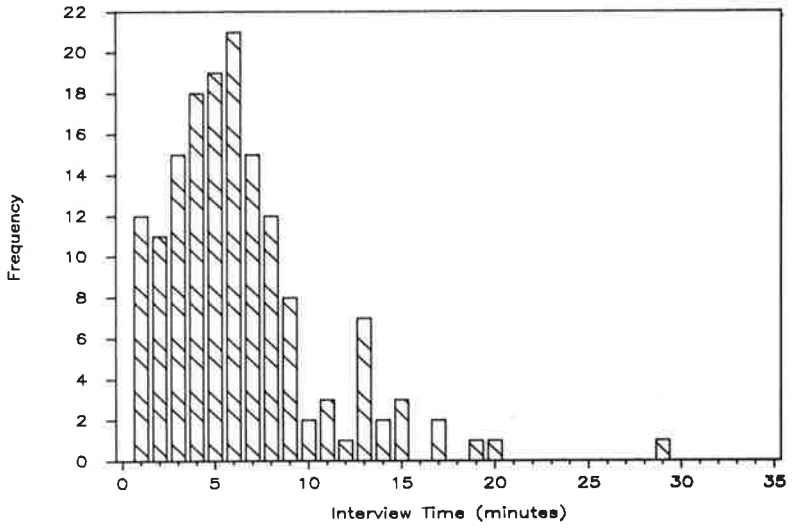


**FIGURE 4 Direct data entry—Histogram of interview times, Station 1.**

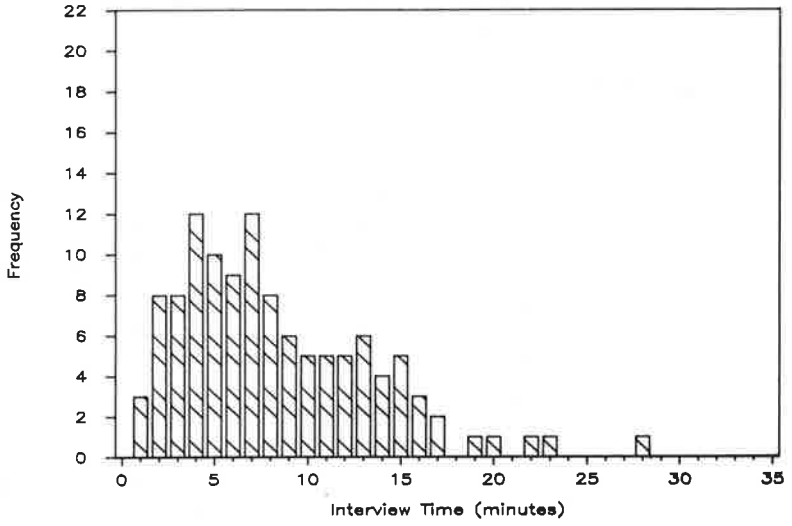


**FIGURE 5 Direct data entry—Histogram of interview times, Station 2.**

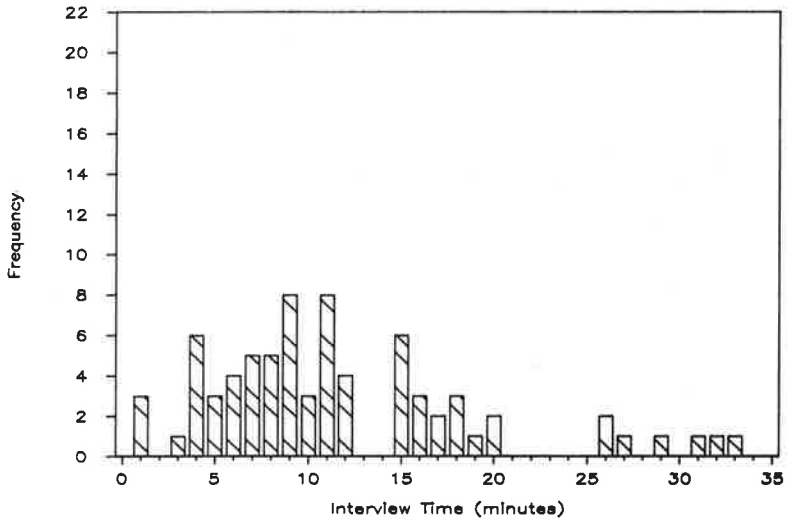


**FIGURE 6 Direct data entry—Histogram of interview times, Station 3.**

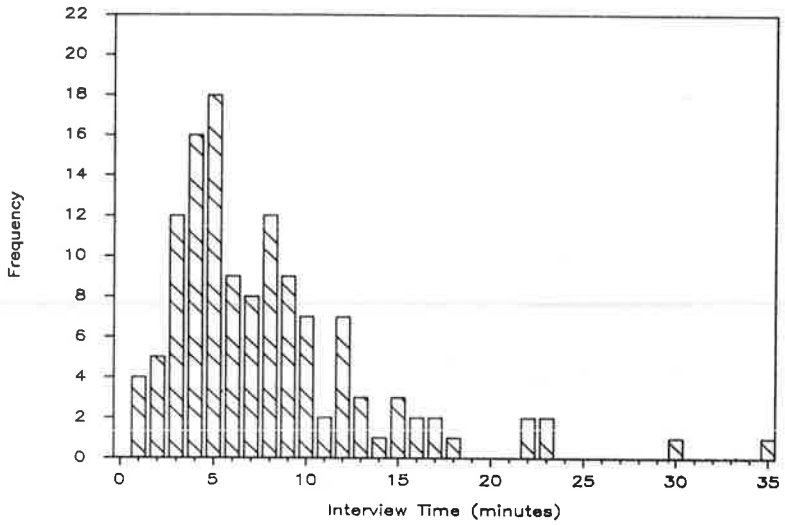


**FIGURE 7 Direct data entry—Histogram of interview times, Station 4.**

TABLE 2 DIRECT DATA ENTRY—SUMMARY OF ACTIVITIES FOR ALL STATIONS

TORONTO TRAVEL SURVEY
PILOT PROJECT
Summary of Automatic Data Entry Activities
'Day' designates the interview day

ALL STATIONS

| Response Code | Day 03/04 | Day 03/05 | Day 03/06 | Day 03/07 | Day 03/08 | Day 03/09 | Day 03/18 | Day 03/19 | Day 03/20 | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 41 | 41 | 57 | 22 | 66 | 52 | 63 | 71 | 56 | 469 |
| 2 | 52 | 29 | 33 | 25 | 46 | 42 | 44 | 63 | 67 | 401 |
| 3 | 16 | 5 | 10 | 4 | 9 | 10 | 9 | 4 | 8 | 75 |
| 4 | 32 | 35 | 49 | 38 | 67 | 44 | 46 | 53 | 36 | 401 |
| 5 | 4 | 3 | 4 | 4 | 3 | 4 | 8 | 7 | 10 | 47 |
| 8 | 0 | 1 | 0 | 0 | 1 | 3 | 1 | 3 | 4 | 13 |
| Total | 145 | 114 | 153 | 93 | 192 | 155 | 171 | 201 | 182 | 1406 |
| [1]Total eligible contacts | 57 | 47 | 67 | 26 | 76 | 65 | 73 | 78 | 68 | 557 |
| [2]Refusal rate | 0.28 | 0.11 | 0.15 | 0.15 | 0.12 | 0.15 | 0.12 | 0.05 | 0.12 | 0.13 |
| [3]Response rate | 0.72 | 0.87 | 0.85 | 0.85 | 0.87 | 0.80 | 0.86 | 0.91 | 0.82 | 0.84 |
| Completions per interviewer hour | 2.31 | 3.50 | 3.06 | 2.08 | 3.50 | 3.29 | 4.21 | 4.24 | 3.45 | 3.32 |
| Minutes per completed interview | 26.0 | 17.1 | 19.6 | 28.9 | 17.1 | 18.2 | 14.3 | 14.2 | 17.4 | 18.1 |

Total number of records processed in 9 days of interviewing: 801

[1]Completed + refusals + language barriers
[2]Refusals/total eligible contacts
[3]Completed/total eligible contacts

Interviewers worked for several evenings before they were comfortable with the compact design of the screens. The initial reaction was for additional prompts and descriptions of each field or step in the process. By the end of the first week, however, they were able to read the complete history of a household's travel information at a glance. When asked in the second week, all interviewers agreed that the screen and computer entry was easier to use than paper forms for this unstructured type of interview. Elaborate screen prompts are not desirable for travel surveys for which the same interviewers are using the same program for up to 10 weeks (i.e., the expected duration of the Toronto O–D Survey).

The rate at which interviewers can process records is made up of two components, interview times and record-processing times. The automatic recording of begin and end times for every interview allows a detailed study of interview times. On the other hand, record-processing times can only be examined in combination with interview times in the form of total records processed in the total number of hours that interviewers spent working. Average interview times for every station and every period of interviewing are presented in Table 1. Improvement is apparent as the interviewers became more experienced. The average interview time settled down to a value of approximately 8 min per interview.

A more complete presentation of interview times is presented in Figures 4–7. The area under the histograms represents the total number of interviews complete during the pilot project. Stations 1, 2, and 4 appeared to interview at similar rates, much more efficiently than Station 3. The interviewer at Station 3 never became completely comfortable with direct data entry using a microcomputer; this lack emphasized the need for interviewers to have at least some typing ability.

Table 2 presents data on the number of completed interviews processed per hour. The best test of the process occurred on the evenings of March 18 and 19 because the interviewers had become more experienced with direct data entry and software problems were solved by this time. The results of March 20 were not as favorable because they represented the final night of interviewing. Consequently, most of this particular evening was spent trying to contact households that had not been reached on previous attempts. The average processing time settled down to a value close to the average of 14.1 min per completed interview experienced for 25,000 interviews in Vancouver in 1985. However, considering the problems Station 3 was having with the process and the precautions taken to keep a record of each sample on paper, an average processing time of 12 min should be achievable using direct data entry when interviewers are screened and become familiar with the process.

## DATA PROCESSING

Even though significant productivity gains are possible with direct data entry due to efficient file management, major cost savings for the automated process are realized during the data processing phase. These savings are possible because the data are handled only once in the automatic process. The same data are handled by at least two other persons in the manual process in which the data are entered from paper form at a keyboard using double entry for verification. Apart from the potential manpower savings, the automatic process should also produce higher quality data by eliminating errors associated with double or triple handling.

Based on an O-D travel survey conducted in Vancouver in 1985, estimated costs for data entry were compared to

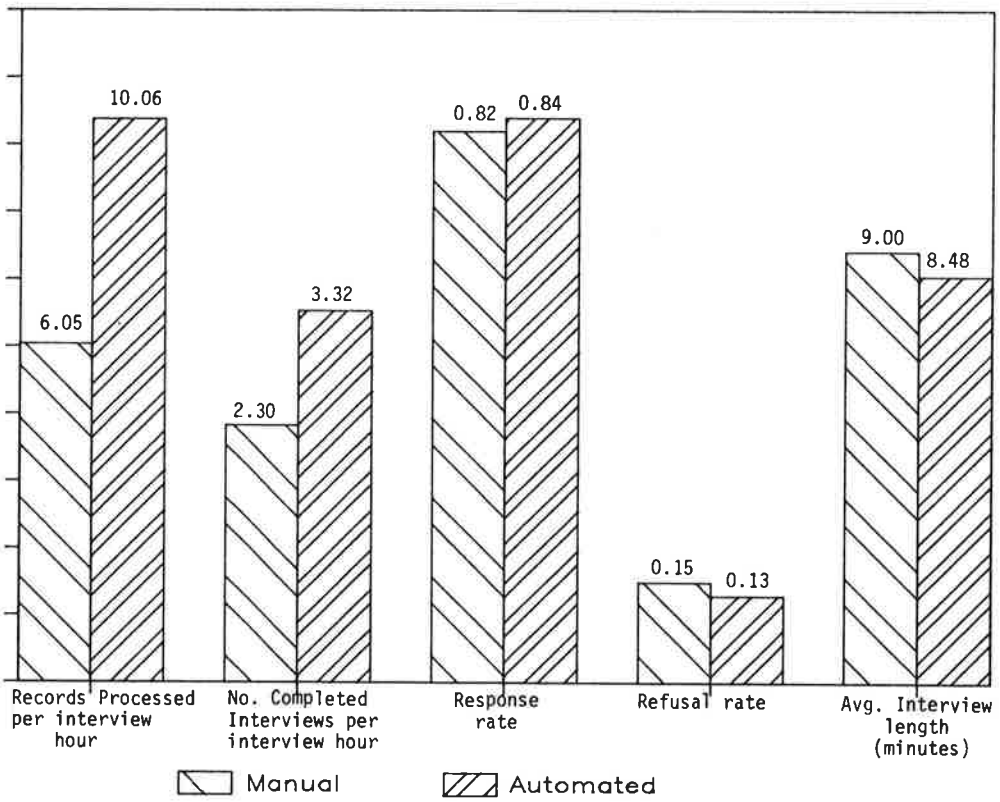


**FIGURE 8 Comparison of survey methods—Manual versus automatic.**

estimated microcomputer rental and program development costs associated with a full-scale Toronto region O-D travel survey *(2)*. It was calculated that the cost of direct data entry is virtually identical to the expense involved in keypunching paper forms. These cost estimates are based on the same number of interviewers working on both methods. One should be careful to observe, however, that improved productivity of interviewers using direct data entry truly makes this method more cost-effective (i.e., direct data entry produces more completed interviews for the same cost as manual methods).

## RESULTS

Figure 8 shows the benefits involved with direct data entry and automatic file management on microcomputers. The comparison is drawn with the paper form method in order to show marked differences in processing times, but similar response, refusal, and interview rates. During the experiment, there were 67 percent more sample records processed per interviewer-hour and approximately 50 percent more completed interviews per interviewer-hour using the direct data entry method.

## CONCLUSIONS

The direct data entry experiment, as part of the pilot project for a Toronto region O-D travel survey, provides new optimism for improving survey efficiency. The experiment proved ideal for comparing reliable manual methods (i.e., paper interview forms) to the automatic mode of data collection and entry. The two groups of interviewers allowed a fair comparison to be drawn between conventional and new methods. Productivity gains using direct data entry amounted to about 50 percent more interviews per hour. Sample processing is greatly improved by the use of an automated method (67 percent more records processed), whereas actual interviewing time, response rates, and refusal rates are comparable for both methods.

The largest benefit from direct data entry surfaces after interviewing is complete. Information collected in the automated process needs only to be coded to a zone system or geocoded; however, the manual method requires paper forms to be entered by keypunch onto a computer storage device. This step not only introduces potential error but also tremendously increases cost. This additional cost is virtually identical to the cost of providing each interviewer with a microcomputer and program if the same number of interviewers are used for both methods. Direct data entry using microcomputers can be incorporated in O-D surveys at no extra cost or risk over reliable manual methods. Furthermore, productivity gains using an automated method provide the potential to greatly surpass conventional methods in terms of cost-effectiveness.

## ACKNOWLEDGMENT

The research described in this paper was funded by a research contract with the Ontario Ministry of Transportation and Communications. The views expressed in this paper are those of the authors and not necessarily those of the Ministry.

## REFERENCES

1. *MTC Origin Destination Survey Pilot Study.* Environics Research Group Ltd., Toronto, 1986.
2. G. N. Steuart. *Survey Coding and Processing Techniques: Applications to a Travel Survey for the Regions of Durham, Halton, Hamilton, Peel, Toronto and York.* Ontario Ministry of Transportation and Communications, Toronto, 1986.