TRANSPORTATION TOMORROW SURVEY

2006

DESIGN AND CONDUCT OF THE SURVEY

TRANSPORTATION TOMORROW SURVEY

2006

A Telephone Interview Survey on
Household Travel Behaviour in the
Greater Toronto and Hamilton Area and the Surrounding Regions
Conducted in the Fall of 2005 and the Fall of 2006
with Extensions into the Winter of 2006 and the Spring of 2007

DESIGN AND CONDUCT OF THE SURVEY

Prepared for the
Transportation Information
Steering Committee

by the

Data Management Group
University of Toronto
Joint Program in Transportation

January 2010

Participating Agencies:

Ministry of Transportation, Ontario • City of Barrie • City of Brantford • City of Guelph
City of Hamilton • City of Kawartha Lakes • City of Peterborough
City of Toronto • County of Dufferin • County of Peterborough
County of Simcoe • County of Wellington • GO Transit • Regional Municipality of Durham
Regional Municipality of Halton • Regional Municipality of Niagara
Regional Municipality of Peel • Regional Municipality of Waterloo
Regional Municipality of York • Toronto Transit Commission • Town of Orangeville

Acknowledgements

Twenty-one agencies were represented on the TTS Technical Committee that planned and directed the 2006 survey. The people who served on the technical committee were:

Teresa Marando, Chair	Ministry of Transportation
Arthur Tai, Secretary	Ministry of Transportation
Jeff Parent	City of Barrie
Richard Forward	City of Barrie
Russ Loukes	City of Brantford
Rajan Phillips	City of Guelph
Mohan Philip	City of Hamilton
Ken Becking	City of Kawartha Lakes
Todd Nancekivell	City of Peterborough
Vince Alfano	City of Toronto
Trevor Lewis	County of Dufferin
Byran Weir	County of Peterborough
Nathan Westendorp	County of Simcoe
Gordon Ough	County of Wellington
Dan Francey	GO Transit
Chris Leitch	Regional Municipality of Durham
Melissa Green-Battiston	Regional Municipality of Halton
Phil Bergen	Regional Municipality of Niagara
Murray McLeod	Regional Municipality of Peel
Eric Chan	Regional Municipality of Peel
Paula Sawicki	Regional Municipality of Waterloo
Loy Cheah	Regional Municipality of York
Mark Hanlon	Regional Municipality of York
Richard Hui	Regional Municipality of York
Bernard Farrol	Toronto Transit Commission
Allen Reid	Town of Orangeville

The survey was managed by the Data Management Group at the University of Toronto. The management team consisted of:

Prof. Gerry Steuart	Project Director
Peter Dalton	Project Advisor
Reuben Briggs	Coding Manager
Susanna Choy	Project Coordinator
Sharon Kashino	Interview and Site Manager
Michael O'Cleirigh	Computer System Manager

The hiring and training of interview staff was supervised by Ian Fisher, assisted by Lucy Balaisis.

Renno Fo Sing and Jackey Ho provided computer support services.

Lorine Jung provided administrative assistance.

The interview team leaders for the 2006 component of the survey were:

Wilma Cromwell	David Piller
Marilyn Nugent	Joy Brennan assisted by Lance Hafner

The daytime supervisor was Stewart Kemp.

Trevor Pitman, of the Toronto Transit Commission, provided daily assistance in reviewing the logic and consistency of all the transit route information collected.

The Java based rewrite of the TTS Survey Software Suite (Direct Data Entry, Sample Control Software and Geocoding Software) was developed and supported by Michael O'Cleirigh of the Data Management Group. Harold Connamacher managed the prioritized task list and contributed to the development between September 2004 and August 2005. Jingda Zhang and then Yousif Atique assisted with the development process between October 2004 and January 2005.

Special thanks to the more than 400 interview and coding staff, who got the job done, and to all those who contributed so much to the success of the previous surveys.

This report was prepared for the Transportation Information Steering Committee (TISC) by the Data Management Group (DMG) at the Department of Civil Engineering, University of Toronto. The Steering Committee, formerly known as the Toronto Area Transportation Planning Data Collection Steering Committee (TATPDCSC), which also conducted the 1986, 1991, 1996 and 2001 TTS, is represented by the Ontario Ministry of Transportation, Cities of Toronto and Hamilton, Regional Municipalities of Durham, Halton, Peel and York, GO Transit and the Toronto Transit Commission. The contributions of the above supporting agencies to the production of this report and to the ongoing work of the DMG are gratefully acknowledged.

Table of Contents

Table of Contents ____ *iii*

Table of Figures ____ *vii*

Table of Tables ____ *viii*

Section 1 Introduction ____ *1*

Section 2 Planning and Organisation ____ *3*

2.1 Organisation ____ **3**

2.2 Survey Design ____ **4**

2.3 Survey Content ____ **6**

2.4 Fall 2005 Survey (Areas External to the GTHA) ____ **7**

2.5 Fall 2006 Survey (GTHA) ____ **8**

2.6 May 2007 Survey (Wilmot) ____ **10**

2.7 Sample Design ____ **11**

2.8 Sample Selection ____ **11**

2.8.1 Area A – External to the GTHA ____ 13
2.8.2 Area B – GTHA ____ 14
2.8.3 Area C1 ____ 18
2.8.4 Area C2 ____ 18
2.8.5 Area C3 ____ 18
2.8.6 Area D – New Hamburg & Baden ____ 18

2.9 Mailing Plan ____ **18**

2.10 Sample Management ____ **20**

2.11 Publicity ____ **21**

2.11.1 Letter to Local Officials ____ 21
2.11.2 Press Release ____ 22
2.11.3 Advance Letter ____ 22
2.11.4 MTO Info ____ 22

Section 3 Software Development ____ *23*

3.1 System Design ____ **23**

3.1 Sample Management System (SMS) ____ **23**

3.1.1 Sample Check-out Processing ____ 24
3.1.2 Sample Check-in Processing ____ 25
3.1.3 Nightly Transition Process ____ 25

3.2 Direct Data Entry (DDE) 27
3.2.1 SMS Provided Interviewer Features 28
3.2.2 Management Control Features for Interviewers 28
3.3 Geocoding Console (GC) 29
3.3.1 Coding Reference Database 29
3.3.2 SMS Provided Geocoder Features 31
3.3.3 Management Control Features for Geocoding 32
3.4 Monitoring Console (MC) 32
3.5 Administration Console (AC) 32
3.5.1 Sample Management Operations 33
3.6 External Reporting 33
3.6.1 Real Time Script Examples 33
3.6.2 Daily Report Examples 34
3.7 Operating System 35
3.8 Open Source Components 36
Section 4 Equipment 38
4.1 Computer Network 38
4.1.1 Servers 40
4.1.2 Clients 40
4.1.3 Backup 42
4.1.4 Resale 42
4.2 Telephones 42
Section 5 Conduct of the Survey 44
5.1 Historical Overview of Survey Statistics 44
5.2 Interview Staffing 45
5.3 Training 47
5.4 Rates of Pay 47
5.5 Hours of Work 48
5.6 Incentive Bonuses 49
5.6.1 Other Work Environment Incentives 49
5.7 Quality Control 50
5.7.1 Logic Checks 50
5.7.2 Monitoring 51
5.7.3 Performance Statistics 51
5.7.4 Visual Review 51
5.7.5 Callbacks 53
5.7.6 Feedback from the Coding Process 53

5.7.7 Rotation of Sample Between Interviewers 53
5.7.8 Random quality control audits 54
5.7.9 Paper Management 54

5.8 Answering Machines (Voice mail) 55
5.8.1 Call-in From Voice Mail 56

5.9 Survey Interruptions 56

5.10 Non-English Callbacks 57

Section 6 Completion Statistics 58

Section 7 Coding 65

7.1 Staffing and Training 65

7.2 Coding Activity 65
7.2.1 Coding in 2005 65
7.2.2 Coding in 2006 66

7.3 Post-Processing 67

7.4 Statistics 68

Section 8 Survey Budget and Costs 70

8.1 University Overhead and Taxes 70

8.2 Cost Summary and Comparison with Previous Surveys 71
8.2.1 Software Development 71
8.2.2 Interview Staff and Training 71
8.2.3 Coding Staff 71
8.2.4 Equipment 72
8.2.5 Other Direct Expenses 73
8.2.6 Management 73

8.3 Unit Cost Comparison with Previous Surveys 73

Section 9 Conclusions and Recommendations 75

9.1 Data Quality 75

9.2 Software 76

9.3 Hardware 76

9.4 Supervisory Staff 76

9.5 Interview Site 77

9.6 Advance Letter 77

9.7 Productivity 78

9.8 Student Population 79

9.9 Sample Selection and Management 79

9.10 Geocoding 80

9.11 Coding Reference Databases 80

Section 10 Recommendations for 2011 82

10.1 Background 82

10.2 A Feasible Approach 82

10.2.1 Survey Method 1 82

10.2.2 Survey Method 2 82

10.2.3 Survey Method 3 83

10.3 Issues Requiring Early Attention 83

10.3.1 Sample Selection 83

10.3.2 Browser Based Interview 83

10.3.3 Development of Cost Estimates 83

Appendix A Letter to Local Officials 84

Appendix B Advance Letter GTHA 88

Appendix C Advance Letter Areas External to the GTHA 91

Appendix D Advance Letter In French 94

Table of Figures

Figure 2.1 Layout 2005 ____ 7
Figure 2.2 Layout 2006 ____ 10
Figure 2.3 Sample Lifecycle ____ 20
Figure 3.1 Nightly Transition Process ____ 27
Figure 4.1 Main Network Set-up ____ 39
Figure 4.2 Local Area Network Set-up ____ 39
Figure 5.1 2005 Interview Staff ____ 45
Figure 5.2 2006 Interview Staff ____ 46
Figure 6.1 Completion Rates for Toronto Postal Areas ____ 61
Figure 6.2 Completed Interviews by Day ____ 63
Figure 6.3 Completed Interviews per Logged Hour ____ 64
Figure 7.1 Post-Processing DDE Screen ____ 68

Table of Tables

Table 2.1 Schedule of Key Events 5
Table 2.2 Purchase of Sample Lists 13
Table 2.3 Mailing Plan 19
Table 5.1 Historical Overview of Statistics 44
Table 5.2 Average Rates of Pay 48
Table 5.3 Typical Performance Printout 52
Table 6.1 Completed Interviews by Agency 58
Table 6.2 Completion Statistics 60
Table 6.3 Disposition of Phone Calls 62
Table 6.4 Completed Interviews by Trip Day 62
Table 7.1 Location Types verses Address Types 69
Table 8.1 Actual Expenditures for TTS's in 1996, 2001 and 2006 72
Table 8.2 Unit Cost Comparisons for TTS's in 1986, 1991, 1996, 2001 and 2006 74
Table 9.1 Productivity and Quality Measures 78

Section 1 Introduction

The 2006 Transportation Tomorrow Survey (TTS) is the largest and most comprehensive travel survey ever conducted in Ontario and perhaps anywhere in North America. The survey was conducted on behalf of 21 local, regional, provincial and transit operating agencies in the greater Toronto and Hamilton Area and surrounding regions. The TTS data contains detailed demographic information on all members of the surveyed household and a ledger of travel information for an entire weekday.

The TTS is a joint undertaking by the agencies represented on the Transportation Information Steering Committee (TISC), formerly known as the Toronto Area Transportation Planning Data Collection Steering Committee (TATPDCSC). The Committee was established in 1977 for the purposes of setting common transportation data collection standards and for coordinating data collection and dissemination between the member agencies. Membership of the committee includes the Cities of Toronto and Hamilton, the Regional Municipalities of Durham, York, Peel, Halton, the Toronto Transit Commission, GO Transit and the Ontario Ministry of Transportation.

The 2006 survey is the fifth in a series of surveys conducted every five years. The first TTS, conducted in 1986, obtained completed interviews for a 4.2% random sample of all households in the GTA. After completion of the 1986 survey, the Data Management Group was formed at the University of Toronto with one of its prime objectives being the management and distribution of the 1986 TTS data. The Data Management Group was also requested to manage the second TTS undertaken in 1991. The 1991 survey was a smaller update of the 1986 survey focusing primarily on those geographic areas that had experienced high growth since 1986. The survey area was expanded slightly to include a band approximately one municipality deep surrounding the outer boundary of the GTA for the purpose of obtaining more complete travel information in the fringe areas of the GTA.

The 1996 TTS was a new survey, not an update. The survey area was expanded to include the Regional Municipalities of Niagara and Waterloo, the counties of Wellington, Simcoe and Victoria and Peterborough, the Cities of Guelph, Barrie and Peterborough and the Town of Orangeville. Approximately 115,200 interviews were completed representing a 5% random selection of households throughout the survey area. Based on Census information, the survey area covered 60% of Ontario's total population. A technical sub-committee of the TATPDCSC was established that included representation from all the participating agencies. The Data Management Group was responsible for all aspects of the management of the survey.

The 2001 TTS was essentially a repeat of the 1996 survey with approximately 137,000 completed interviews. The survey area was the same as in 1996 except for the exclusion of the Regional Municipality of Waterloo and inclusion of City of Orillia and all of the County of Simcoe. The organizational structure and the role of the Data Management Group were also the same as for the 1996 survey.

The 2006 TTS covered all of the area involved in the 2001 survey plus the Regional Municipality of Waterloo, which had been surveyed in 1996 but not in 2001, and the City of Brantford and

County of Dufferin which had not been surveyed in previous versions of the TTS. The survey involved cooperation from seven cities, ten regional and county governments, one town, two transit operators and one provincial ministry. In order to provide contiguous coverage in the area surveyed, Brant County was also surveyed during the training of interview staff. Altogether approximately 149,000 households were successfully interviewed.

The 1996, 2001 and 2006 surveys are three of the largest travel surveys ever undertaken anywhere. The 1986, 1991 and 1996 surveys each involved a major element of technology development. The use of automated geocoding was a key development in the 1986 survey. On-line Direct Data Entry (DDE) was introduced in the 1991 survey and networked computers in the 1996 survey. The survey methods were essentially unchanged in 2001 with only minor revisions to some of the computer software.

The survey methodology and questionnaire in the 2006 survey was the same as the previous surveys. However, the sample control, interviewing and geocoding software were re-written to take advantage of the experience and knowledge gained in the conduct of such surveys in order to provide better performance and quality control. A telephone interview with on-line Direct Data Entry (DDE) and automated geocoding of all geographic information collected was adopted as the proven most cost effective and reliable means of collecting large quantities of travel data.

The interviews for the 2006 TTS were conducted in three stages. Areas external to the GTHA were interviewed in the fall of 2005, the GTHA in the fall of 2006. A small number of additional interviews were conducted in May 2007 to correct for problems identified in the original sample selection.

Section 2 Planning and Organisation

The selection of the Data Management Group to manage the 2006 survey ensured continuity from the initial planning and design of the survey through to the dissemination of the final database and subsequent analysis of results. The selection also took advantage of the experience gained from the 1986, 1991, 1996 and 2001 surveys, ensuring consistency in survey methods and results.

2.1 Organisation

A Transportation Tomorrow Survey in the year 2006 was initiated by a long standing Transportation Information Steering Committee (TISC) in the Greater Toronto and Hamilton Area (GTHA). TISC asked the Data Management Group (DMG) to manage the survey and approved an initial budget based on the DMG's initial work plan and schedule. A collection of agencies external to the GTHA that had participated in past surveys was invited to participate in the 2006 Survey. Two new agencies (County of Dufferin and the City of Brantford) asked to be included. A TTS Technical Steering Committee was assembled consisting of a representative from each of the participating agencies. It met once every three to six months to receive progress reports from the Project Director and to make, or confirm, decisions on critical items.

The Management structure was established based on the need to draw on the experiences gained in the conduct of the previous surveys at the same time as broadening the base of experience that might be used in the conduct of future surveys. A Management Team was assembled in 2005 and met on an informal, as required, basis to discuss all aspects of the design and conduct of the survey. The composition of the Management Team was as follows:

Gerald N. Steuart, Project Director
Gerald has been involved in every TTS starting with 1986. He is the Manager of the DMG and served as Project Director for the 1996, 2001 and 2006 TTS.

Peter M. Dalton, Project Advisor
Peter is currently a private consultant and has been involved in a senior management role in every TTS (1986, 1991, 1996, 2001 and 2006).

Susanna T.T. Choy, Project Coordinator
Susanna was Coding Manager in the 2001 survey and was involved in the conduct of the 1991 survey and post survey processing of the 1996 survey data. A long time employee of the Data Management Group her responsibilities have included the ongoing maintenance and distribution of the TTS data.

Reuben Briggs, Coding Manager
Reuben operated as a support person on the 2001 TTS and played a significant role in the development of improvements to the coding process. He is a long time employee of the Data Management Group with responsibilities that include the ongoing maintenance and distribution of the TTS data.

Sharon Kashino, Interview and Site Manager
Sharon is currently a freelance consultant. She began her TTS experience providing software support in addition to being an Interviewing Team Leader in 1996. She assumed responsibility for telephone interviewers in 2001 and continued in that role in 2006. She was extensively involved in the post processing stages of both the 1996 and 2001 TTS.

Ian Fisher, Manager of Interviewer Training
Ian is a freelance consultant with experience on every TTS (1986, 1991, 1996, 2001 and 2006) He personally interviewed more than 350 potential telephone interviewers and gave each their introduction to the interviewing procedures used in the 2006 TTS.

Michael O'Cleirigh, Computer System Manager
Michael is a full-time employee of the Data Management Group. He began his experience as the lead software developer of the TTS software re-write undertaken by the DMG. The task began in 2004 and his responsibilities increased as the production phases of the project began to take place.

Pentti Soukas from the Ontario Ministry of Transportation acted as liaison with the Ministry and as the secretary of the TTS Technical Committee. Louise Hominuk was the MTO Info contact.

Trevor Pitman of the Toronto Transit Commission was seconded to the project to review and edit all transit routes in all jurisdictions recorded by the interviewers. Mr. Pitman was also an active member in the conduct of the 1996 and 2001 TTS.

In 2006, after a few safety concerns were raised by staff, the TTS site was toured by personnel from the Ontario Labour Board and the University of Toronto Environmental Health and Safety Office. A Health and Safety committee was established at the TTS site to deal with any future concerns. The Health and Safety committee consisted of 6 persons: one representative from each of the management team, the four interviewing teams and the geocoding team. This committee met regularly and inspected the site on two occasions undertaking to bring any health and safety concerns to the attention of management who would take any necessary action.

2.2 Survey Design

The success and cost effectiveness of the 1986, 1991, 1996 and 2001 surveys, together with the need for a consistent time series, resulted in the same survey methods being adopted for the 2006 survey. The basic survey methods consisted of an advance letter mailed to each of the selected households followed, about a week later, by a telephone interview to collect demographic data and travel information for the previous weekday for each member of the household. A universal co-ordinate system was used to record geographic information to allow assignment to any zone system.

Although the survey methods and procedures remained the same, significant changes to the computer system supporting these methods and procedures were necessary. The underlying

software used in 1996 and 2001 was proving to be unsuitable for a survey of this magnitude. The software re-write began in 2004.

Experience gained in the 1996 and 2001 surveys reinforced the conviction that management and supervision costs per interview increased when a call centre was larger than 4 teams of approximately 25 interviewers per team. This meant that the survey needed to be conducted in two phases, one in the Fall of 2005 and the second in the Fall of 2006. To be certain that school was in session during the interviews, the intent was for each session to start in September and finish as early as possible in December. An added benefit was that the estimate of the universe of households in the survey area from Statistics Canada in May 2006 would fall conveniently between these two phases. An adjustment to household data from Statistics Canada would, therefore, not be necessary.

Based on anticipated interviewer productivity, the objective of the first phase was to complete 35,000 interviews in the areas outside the GTHA. The objective of the second phase was to complete 115,000 interviews for households in the GTHA. Productivity problems occurred in both phases which meant that completion targets could only be met with interviewing extended into January and February of 2006 and into January of 2007.

The 2001 TTS demonstrated a clear advantage for the interviewing site to be located close to a subway station in the central area of Toronto. Fortunately it was possible to conduct both phases of the 2006 survey from the same location at 500 University Avenue in central Toronto. A significant number of interviewers returned from the 2001 TTS. In addition, having the same site location for the second stage of the survey proved to be beneficial in terms of being able to re-hire many of the same interviewers.

Table 2.1 Schedule of Key Events

Fall 1986	Conduct of the 1986 TTS (61,708 households interviewed)
August 1988	Release of the 1986 TTS database (Version 2.0)
December 1989	Data Management Group appointed to manage the 1991 TTS
Fall 1991	Conduct of the 1991 TTS (24,507 households interviewed)
June 1992	Release of the 1991 TTS database (Version 2.1)
January 1995	Data Management Group appointed to manage the 1996 TTS
Oct./Nov. 1995	Conduct of the Waterloo component of the 1996 TTS (7,556 interviews completed)
Sep-Dec 1996	Conduct of the main portion of the 1996 TTS (108,850 households interviewed)
August 1997	Release of the 1996 TTS database (Version 2.1)
May 1999	Data Management Group appointed to manage the 2001 TTS
Sep-Nov 2000	Conduct of external portion of the 2001 TTS (22,000 household interviews)
Sep-Dec 2001	Conduct of the main portion of the 2001 TTS (101,000 households interviewed)

May 2002	14,000 additional interviews conducted
December 2002	Release of final 2001 TTS database (Version 1.0)
December 2004	First meeting of the 2006 TTS Technical Committee
June-July 2005	500 University Ave. selected as survey site for stage 1 of the 2006 TTS
July 2005	Installation and testing of phones, computer systems and software at 500 University Ave for stage 1 of the 2006 TTS
August 2005	Initial recruitment and training of interview staff for stage 1
Sep.2005-Feb 2006	Conduct of external portion of the 2006 TTS (37,000 household interviews)
May 2006	National census (Statistics Canada)
July 2005	Selection of site, installation and testing of phones, computer systems and software at 500 University Ave for stage 2 of the 2006 TTS
August 2006	Initial recruitment and training of interview staff for stage 2
Sep 2006-Jan 2007	Conduct of the main portion of the 2006 TTS (115,000 households interviewed)
May 2007	2,000 additional interviews conducted
December 2008	Release of final 2001 TTS database (Version 1.0)

2.3 Survey Content

No changes were made in survey content relative to the 2001 survey.

The survey consists of the following questions:

Household Data

- Home Location
- Type of dwelling unit
- Number of persons
- Number of vehicles available for personal use

Person Data

- Gender
- Age
- Possession of a driver's licence
- Possession of a transit pass
- Employment status
- Occupation
- Usual work location
- Availability of free parking at place of work
- Status as a student
- Usual school location (Name of school)
- Origin of first trip

Trip Data (Only collected for persons 11 and older)

- Location of destination
- Trip purpose
- Start time
- Method of travel

For Trips made by Public Transit

- Method of access
- Sequence of transit routes and/or boarding & alighting stations (maximum of 6)*
- Method of egress

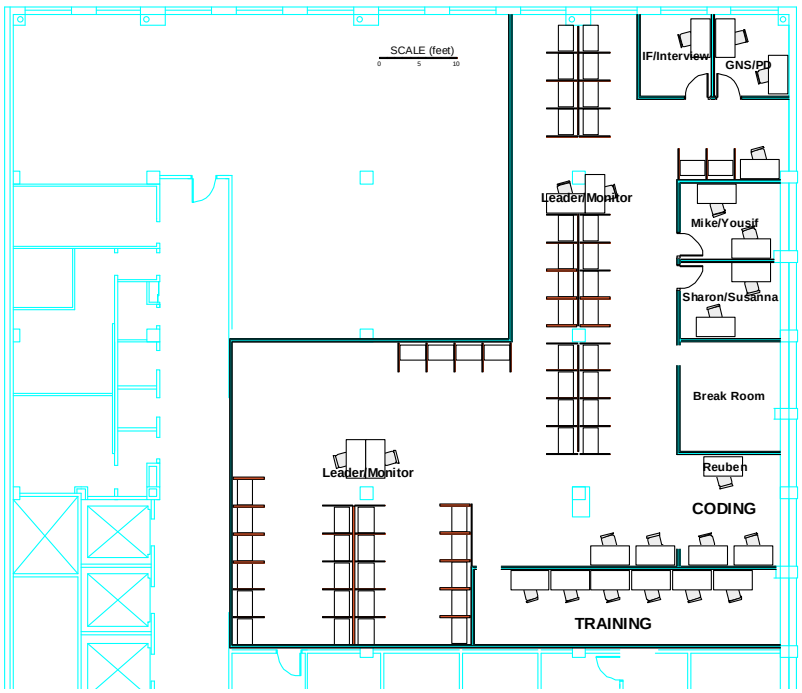
* The transit route is recorded for each segment of a transit trip made by bus or streetcar. The access mode, egress mode, each transit route used (maximum 6) as well as boarding and alighting stations (where subway, GO Rail or RT are used) are recorded as parts of a single trip.

Details of all the response categories and definitions are contained in both the Interview Manual (2006 Transportation Tomorrow Survey Working Paper Series: Interview Manual, August 2006) and the Data Guide (2006 Transportation Tomorrow Survey: Data Guide Version 1.0, October 2008).

2.4 Fall 2005 Survey (Areas External to the GTHA)

The search for an appropriate interview site commenced in May 2005. Basic requirements were identified as approximately 500 square meters of open floor space in downtown Toronto with good access to the subway. Appropriate space available from August 1st to the end of December was found at 500 University Avenue, which is the same location (not the same space) as 2001. A layout of the survey site for the first phase in the Fall of 2005 is shown in Figure 2.1

Figure 2.1 Layout 2005



The site was equipped with approximately 65 used Dell Optiplex GX260 computers (Pentium 4, 1.8 Ghz) that were obtained at the end of their lease from the Engineering Computing Facilities at the University of Toronto. Two interviewing teams were composed of approximately 30 stations, each with two monitoring stations; the remaining computers were utilized by the management team, for geocoding, and training purposes. All stations were setup using Debian Linux Stable and further customized to create specific and limited profiles for each of the training, interviewing, reviewing and geocoding roles. Each of the two monitoring stations was able to mirror the screen of any of the 30 workstations, while at the same time listening to the interview in progress on a silent telephone monitoring system. None of the non-management computers were allowed to connect to the internet which was provided by the Faculty of Medicine. Geocoders were allowed web access through a transparent proxy server that allowed management oversight and the capability to restrict access to non-geocoding related websites. Two Dell Power Edge 1800 servers were used; one ran the sample control software while the other provided a development platform and network file server. All of the computer equipment and telephone equipment was retained for use in the main part of the survey in 2006.

The 5% sample requirement translated into a target of 37,000 completed interviews. A randomly distributed sample of residential phone listings was purchased from InfoCanada, a private company specialising in the maintenance and distribution of phone and mailing lists. An initial list of 34,689 residential phone listings (name, address and phone number) was obtained in early August for use in training and the initial start up of the survey. A second list of 59,407 was obtained in mid October. The purchase of the 2nd list was delayed until October in order that students moving into University and College residences in September would be included.

The survey commenced on Wednesday, September 7, 2005 and ended on Wednesday, December 21, 2005. Interviewing resumed on Tuesday, January 10, 2006 and ended on Thursday, February 9, 2006. A total of 201 interviewers and 5 geocoders were recruited and trained. 2 staff members originally recruited as interviewers and team leaders subsequently became geocoders increasing the total coding staff complement to a maximum of 7. 37,442 interviews were completed successfully. A small number of records were subsequently discarded as being incomplete or outside the survey area.

2.5 Fall 2006 Survey (GTHA)

The only real difference between the 2005 and 2006 components of the survey was in the scale of operation. The minimum space requirement was identified as approximately 800 square meters of open floor space. An opportunity arose that allowed for the space occupied in the 2005 phase to be expanded to 750 square meters. The space was renovated and occupied for 5 months from August 1st to the December 31st. Free access was granted in the last week of July for installing wiring, computers, telephones and furniture. Because the space was too small for all activities, supplementary space was found in the same building on another floor. This space was occupied temporarily during interviewer training from August to October, 2006.

To accommodate the larger staffing needs for the 2006 portion of the survey an additional 65 Dell OptiPlex GX520 (Pentium 4, 2.8Ghz) computers were purchased new from Dell. Network setup was simplified by having exactly the same configuration. Only two distinct Debian Linux images

were required. Four teams of approximately 30 interviewing with two monitoring stations each (8 in total) were established. Three computers were separated and dedicated to processing respondent call-ins. By supporting all user profiles on all survey workstations it was possible to have geocoders and reviewers situate themselves at any location within the call centre and allow full management control over where such activities would take place. Unlike the 2001 TTS other language interviews could be conducted on any interviewing stations so long as the interviewer was configured appropriately with the sample control software. Most of the computers were re-sold on completion of the survey.

The majority of the interview stations were separated from each other by 5-foot high screens for the purpose of sound attenuation. The exception to this occurred in the area used by the geocoders during the day and interviewers at night, where only three of the 23 stations had screens. The monitoring/supervisor stations were located in open areas with an optimum view of the interview stations they were set up to monitor. Each of the semiautonomous teams was set up in a similar manner to the fall 2005 site with approximately 30 interview stations and two visual and telephone monitoring stations. Initially two separate rooms, one with a large boardroom table and one with computer stations were available for nights 1, 2 and 3 training of new interviewers prior to their going live on the telephone. Midway through the project these rooms became unavailable and training was moved to a smaller area on the interview floor. The site facilities included a meeting and break room equipped with a fridge and coffee maker allowing interviewers to take their breaks without leaving the premises. Additional space as well as a microwave was available on the main floor of the building. Access into the building and use of the elevators was limited by the use of a pass card, of which a limited number were available. Ease of access should be a consideration in the selection of future TTS sites. A layout of the survey site is shown in Figure 2.2

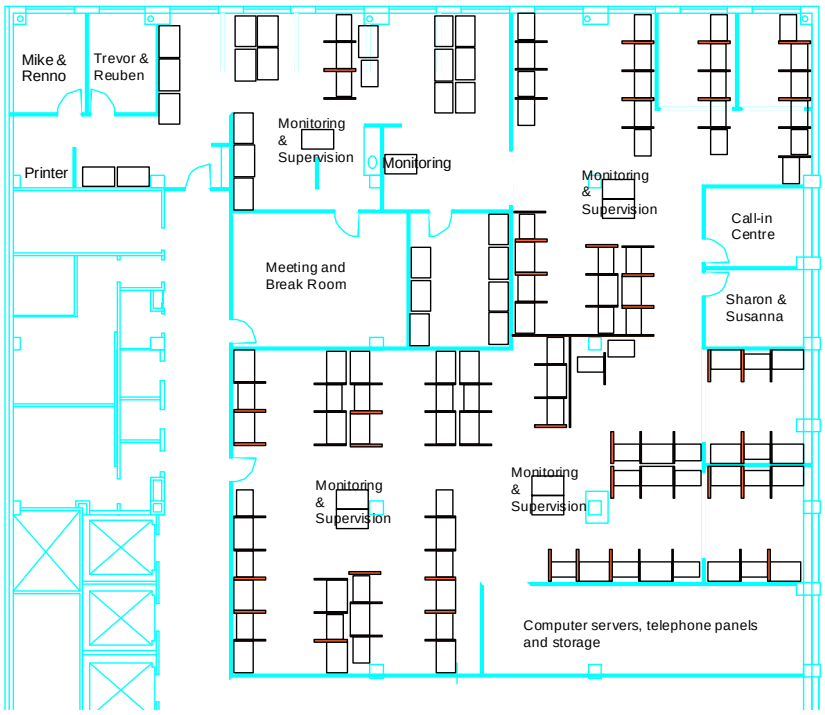
Figure 2.2 Layout 2006

The 5% sample translated into a target of 115,000 completed interviews. As in the fall of 2005, the sample was purchased in multiple stages. Midway through July the 1st list was obtained containing 90,787 phone listings. A 2nd list of 175,992 was obtained in late September after completion of the September updates to the residential phone lists. A final list was received in mid December with 32,402 listings. An additional list of 486 households was purchased in May, 2007 in a new postal code area that had unknowingly been excluded previously. Additional interviews were conducted in May 2007 to correct for this missing area.

Live interviewing commenced on Wednesday, September 6, 2006 and finished on Wednesday, December 20, 2006 with a continuation of callbacks for the remainder of that last week. Interviewing resumed Tuesday, January 9, 2007 until Thursday, January 25, 2007. Interviews were also conducted for one week in May, 2007. A total of 370 interview staff and 16 geocoders were recruited. 3 staff members originally recruited as interviewers and team leaders subsequently became geocoders increasing the total coding staff complement to a maximum of 19. A total of 113,003 interviews were successfully completed, including 260 done during training (in August) for Brant County.

2.6 May 2007 Survey (Wilmot)

The Region of Waterloo was included in the 1st phase of TTS conducted in the fall of 2005. As the final results were being validated it was determined that part of the Township of Wilmot was missing in the sample. In order to have a true representation of the region, an additional 200 interviews were conducted in this township over a one week period in May, 2007.



2.7 Sample Design

The survey target was to achieve completed interviews for a 5% random selection of households throughout the survey area. The listing of households included in the survey was obtained from InfoCanada, a private company specializing in the maintenance of phone and mailing lists for the market research and tele-marketing industries. InfoCanada obtains the white page phone listings from Bell Canada with regular monthly updates. The information supplied by InfoCanada for each household in the sample list consisted of:

- Name
- Street Address
- Municipality
- Postal code
- Phone number

CRTC regulations, introduced in 1991, do not allow Bell Canada to release information that is not contained in the telephone directory. Apartment numbers are generally not included in directory listings for Toronto and surrounding areas and were therefore not included in the listings obtained from InfoCanada.

The sample frame used for the survey consists of listed residential phone numbers within the boundaries of the survey area defined as accurately as possible by postal codes. Households without phones, or with unlisted phone numbers, were excluded from the sample frame while households with multiple listed phone numbers were included more than once. The extent to which these limitations in the sample frame affect the results of the survey is not known. The 1986 and 1991 surveys produced no evidence of significant bias that could be attributed to this factor. The sample frame for the 2006 TTS also excludes households whose members have specifically requested that they be excluded from any telephone or mailing lists given out for marketing or market research purposes. Concerns arising from the conduct of the 1996 and 2001 surveys include:

1. The increase in use of cell phones as an alternative to land lines.
2. The potential underrepresentation of post secondary students.
3. Poor response rates from households living in apartment units.

The post secondary student concern was addressed, in part, by purchasing two lists for each phase of the survey. The 1st list, used for staff training and initial start up, was purchased in July/August. The 2nd, larger, list was purchased in September/October after most post secondary students had taken up residence.

The above concerns, and the effectiveness of the measures taken, are discussed in the validation report “2006 Transportation Tomorrow Survey: Data Validation”.

2.8 Sample Selection

The 2006 TTS area is divided into two components, surveyed in the years 2005 and 2006 respectively, based on postal codes. In urban areas, the first three characters known as the Forward Sortation Area (FSA) are used. In rural areas, the full 6-character code known as the Local Delivery Unit (LDU) is used. In most cases, each LDU is a rural post office. FSAs and

LDUs are not always mutually exclusive in terms of the geographic area they serve. The exact location of a house cannot be determined from the postal code even in urban centres, particularly where box numbers and general delivery codes are used. The boundary of the two areas surveyed is approximate such that some households inside the GTHA were included in the fall 2005 and others, outside the GTHA in the fall 2006 survey.

The sampling procedure used by InfoCanada was to select every n^{th} record after sorting on postal code and street address. The same procedure was used in selecting the samples for the 1986 and 1991 surveys. The sample listings for the 1996 and 2001 surveys were obtained using random selection from the sample frame. There are advantages and disadvantages to each method of sample selection. Selecting every n^{th} record ensures that the sample is distributed uniformly in proportion to the sample frame across the entire survey area but could potentially result in a biased survey if there is a pattern in the way the sample frame is sorted that coincides with the selection frequency. The difference in sample selection procedure is not expected to affect the survey results in any way.

The information contained in the phone listings maintained by InfoCanada includes a “multi-unit” flag for street addresses that are duplicated in the sample frame. The availability of this flag facilitates analysis of response rates by dwelling unit type and permits the two categories to be sampled at different rates.

Table 2.2 gives details of the 6 sample lists that were purchased from InfoCanada. The total number of records is the number that was obtained from InfoCanada. The usable number excludes duplicate records from the previous sample selections and records deleted because they were known to be outside the survey area. Any records containing less than the 1^{st} 3 characters of the postal code were also deleted.

The definition of the survey area was still being refined at the time of the 1^{st} sample purchase. Households in forward sortation areas L1A, K9A and K0K were included in that sample purchase but those records were subsequently deleted.

The rural delivery areas L0G and L0R straddle the boundary between the 2005 and 2006 survey areas. The 1^{st} sample purchased in 2005 included all of those two FSAs. The records for local delivery units known to be inside the GTHA (from the 2001 survey) were then deleted. The definition of Area A was refined prior to the 2^{nd} purchase to only include the remaining local delivery units in L0G and L0R. Sample purchases 3 and 4 also included all of L0G and L0R. Records in the local delivery units included in phase 1 (sample purchases 1 & 2) were deleted prior to interviewing. This approach was taken to ensure that any newly created rural postal codes were not omitted.

Analysis of the phase 1 interviewing statistics showed that the overall response rate was 16% higher for records identified as “single unit” as compared with those flagged as “multi unit” in the original sample list. A validation check was also performed to determine the consistency between the sample categories and the dwelling unit categories as determined by the respondents in the completed interviews with the following results.

Flag	House	Townhouse	Apartment
Single unit	93%	3%	4%
Multi unit	12%	18%	69%

Based on the above analysis a decision was made to sample "multi-unit" residences at a higher rate in phase 2 in order to compensate for the anticipated lower response from those living in apartment buildings with particular reference to the City of Toronto and the potential for survey bias resulting from the high proportion of apartment units in that City. The difference in sampling varied between sample purchases due to the need to choose discrete values of 'n' for each sample purchase. The average difference is approximately 18.5%. There are small variations by geographic area (in the range 17% to 20%) but those differences are not considered to be significant in the context of the survey results.

Significant differences in response rate, greater than in previous surveys, resulted in the need to purchase additional sample (Purchase #5) for selected areas. The relevant FSAs were stratified into three groups sampled at different rates according to the response rates experienced in the 1st 3 months of phase 2 interviewing.

Two errors in the initial sample selection were discovered during data expansion subsequent to the completion of phase 2 interviewing.

1. Parts of Dufferin County, in the forward sortation area 'L0N', were inadvertently included in the survey areas for both phase 1 and 2. As a result the completed sample for that area contains between 250 and 300 more interviews than were necessary to meet the 5% target.
2. The recently created forward sortation area N3A serving the communities of New Hamburg and Baden in the township of Wilmot, Waterloo Region, was omitted from Area A. To rectify that situation sample list 6 was purchased and additional interviews conducted in May 2007.

Table 2.2 Purchase of Sample Lists

Purchase #	Delivery date		Value of 'n'		Number of records	
		Area	Single unit	Multi unit	Total	Usable
1	9 Aug 2005	A	21	21	34,689	33,001
2	18 Oct 2005	A	11	11	59,407	53,220
3	17 Jul 2006	B	20	17	90,787	88,955
4	29 Sep 2006	B	11	9	175,992	162,981
5	13 Dec 2006	C1	25	21	32,402	19,967
		C2	18	15		2,454
		C3	12	10		5,361
6	4 May 2007	D	10	10	486	486
Total					393,763	366,425

Note – When drawing multiple samples from the same area it is important not to duplicate, or to have an exact multiple of, the sample rate since that could lead to multiple duplication of records thus creating a geographic bias.

2.8.1 Area A – External to the GTHA

All postal codes beginning with the characters

L2 N2 (Except N2Z)

Forward Sortation Areas

K9H K9J K9K K9L K9V
L0A L0K L0L L0N L0M L0S
L3B L3C L3K L3M L3V L3Z
L4M L4N L4R
L9R L9M L9S L9V L9W L9Y L9Z
N0B
N1C N1E N1G N1H N1K N1L N1M N1P NIR N1S N1T
N3B N3C N3E N3L N3H N3P N3R N3S N3T N3V

All local delivery units with the 1st 5 characters

K0L 1B	BAILIEBORO
K0L 1H	BRIDGENORTH
K0L 1J	BUCKHORN
K0L 1K	BURLEIGH FALLS
K0L 1R	CURVE LAKE
K0L 1S	DOURO
K0L 1T	ENNISMORE
K0L 1V	FRAZERVILLE
K0L 2B	INDIAN RIVER
K0L 2C	JUNIPER ISLAND
K0L 2E	KAWARTHA PARK
K0L 2G	KEENE
K0L 2H	LAKEFIELD
K0L 2V	NORWOOD
K0L 2W	OMEMEE
K0L 2X	REABORO
K0L 3A	WARSAW
K0L 3B	WESTWOOD
K0L 3G	YOUNGS POINT
K0L 3H	CENTURY VILLAGE
K0M 1A	BOBCAYGEON
K0M 1B	BOLSOVER
K0M 1C	BURNT RIVER
K0M 1E	CAMBRAY
K0M 1G	CAMERON
K0M 1K	COBOCONK
K0M 1L	DUNSFORD
K0M 1N	FENELON FALLS
K0M 2A	KINMOUNT
K0M 2B	KIRKFIELD
K0M 2C	LITTLE BRITAIN
K0M 2J	MANILLA
K0M 2L	NORLAND
K0M 2M	OAKWOOD
K0M 2T	WOODVILLE
L0G 1A	BEETON
L0G 1B	BOND HEAD
L0G 1L	LORETTO
L0G 1W	TOTTENHAM
L0R 1B	BEAMSVILLE
L0R 1E	CAISTOR CENTRE
L0R 1G	CAMPDEN
L0R 1M	GRASSIE
L0R 1Y	ST ANNS
L0R 2A	SMITHVILLE
L0R 2J	WELLANDPORT
L0R 1S	JORDAN STATION
L0R 2C	VINELAND
L0R 2E	VINELAND STATION
L0R 2N	BEAMSVILLE
N0B 1B	ARISS
N0B 1C	ARKELL
N0B 1H	BALLINAFAD
N0B 1J	BELWOOD
N0B 1P	EDEN MILLS
N0B 1S	ELORA
N0B 1T	ERIN
N0B 1Z	HILLSBURGH
N0B 2C	MORRISTON
N0B 2J	PUSLINCH
N0B 2K	ROCKWOOD
N0C 1M	SINGHAMPTON
N0E 1A	BURFORD
N0E 1B	CATHCART
N0E 1K	MOUNT PLEASANT
N0E 1L	OAKLAND
N0E 1R	SCOTLAND
N0E 1N	St. GEORGE

2.8.2 Area B – GTHA

All postal codes beginning with the characters

M (Toronto) L1 L5 L6 L7 L8

Forward Sortation Areas

L0B L0C L0E L0G L0H L0J L0N L0P L0R
L3P L3R L3S L3T L3X L3Y
L4A L4B L4C L4E L4G L4H L4J L4K L4L
L4P L4S L4T L4V L4W L4X L4Y L4Z
L9A L9B L9C L9G L9H L9J L9K L9L L9N L9P L9T

2.8.3 Area C1

L0P L4B L6J L7R L7S L8N M1B M1E M1G M1H M1J M1K M1L M1M M1R
M2K M2M M2N M2R M3C M3H M3J M3K M3L M3M M4A M4B M4C M4E M4G
M4H M4J M4K M4L M4M M4N M4R M4S M4T M5M M5N M5P M5T M6A M6B
M6C M6E M6G M6H M6J M6L M6M M6N M6P M6R M8V M8W M8X M8Y M9B
M9M M9V M9W

2.8.4 Area C2

M2L M2P M3N M4W M4X M5B M5E M5J M9N

2.8.5 Area C3

M4P M4V M4Y M5A M5C M5G M5H M5R M5S M5V M6K

2.8.6 Area D – New Hamburg & Baden

N3A

2.9 Mailing Plan

On receipt of each sample selection, a random number was assigned to each household record. The records were then sorted on the random number and assigned to mailing blocks. An electronic copy of the address information was provided to a commercial mailing house (Corporate Mailing and Printing) who were contracted to mail the advance letter to each household. The files for each mailing were sent to the mailing house by email at least 3 days before each mailing.

Care was taken when new mailing lists were received to move the remaining sample from previous lists that had not already been included in a previous mailing to the end of the combined sample queue in order to maximise the use of the more current listing. The number of households included in the final mailing for each phase of the survey was based on the estimated number of additional records needed to achieve the sample target set for each individual FSA. The remaining households not yet included in a previous mailing were combined into a single list. A priority rating was then assigned to each record equal to:
(The estimate additional sample required to achieve the completion target for that FSA - The number of households already assigned a priority rating for that FSA) / (The estimate additional sample required to achieve the completion target for that FSA).
The households were then assigned to the remaining mailing blocks in priority sequence.

Through 2005 testing was done on the use of 1st versus 3rd class mail. 1st class was found to be faster. Both were equally reliable. In 2006 3rd class mail was used except in cases where immediate receipt of the letters was essential (first and last mailings as well as mailings during the Christmas period).

Table 2.3 Mailing Plan

Numbers and dates are approximate.

Fall 2005

Mailing	# of Letters	Mailing Date	Mailing Class	
1	980	September 6, 2005	1	Training Sample
2	2,000	September 13, 2005	1	
3	3,000	September 20, 2005	1	
4	3,000	September 29, 2005	1	
5	3,800	October 6, 2005	1	
6	3,680	October 17, 2005	3	
7	3,790	October 21, 2005	1	
8	3,770	October 27, 2005	1	
9	7,800	November 3, 2005	1	
10	7,570	November 8, 2005	3	
11	7,520	November 18, 2005	1	
12	7,640	November 24, 2005	1	
13	1,440	December 8, 2005	1	
14	1,430	December 12, 2005	1	
15	8,590	January 3, 2006	1	
16	5,760	January 10, 2006	3	
17	5,150	January 16, 2006	3	

Fall 2006/Winter 2007

Mailing	# of Letters	Mailing Date	Mailing Class	
1	500	August 10, 2006	1	Training Sample
2	1,000	August 15, 2006	3	Training Sample
3	2,000	August 21, 2006	3	Training Sample
4	6,000	August 22, 2006	3	Training Sample
5	8,000	August 28, 2006	3	
6	10,000	September 1, 2006	3	
7	12,000	September 6, 2006	3	
8	12,000	September 12, 2006	3	
9	12,000	September 22, 2006	3	
10	12,000	September 28, 2006	3	
11	13,000	October 5, 2006	3	
12	15,000	October 13, 2006	3	
13	15,000	October 20, 2006	3	
14	20,000	October 25, 2006	3	
15	20,000	October 30, 2006	3	
*16	20,000	November 6, 2006	3	
17	20,000	November 14, 2006	3	
18	15,000	November 21, 2006	3	
19	8,000	December 5, 2006	3	
20	5,000	December 8, 2006	1	
21	6,000	December 12, 2006	1	

22	10,000	December 27, 2006	3
23		*2000 letters not mailed*	
24	9,400	January 8, 2007	3
25	10,000	January 12, 2007	1

- starting with mailing 16 French letter was sent out with English letter. 100,000 French letters were printed.

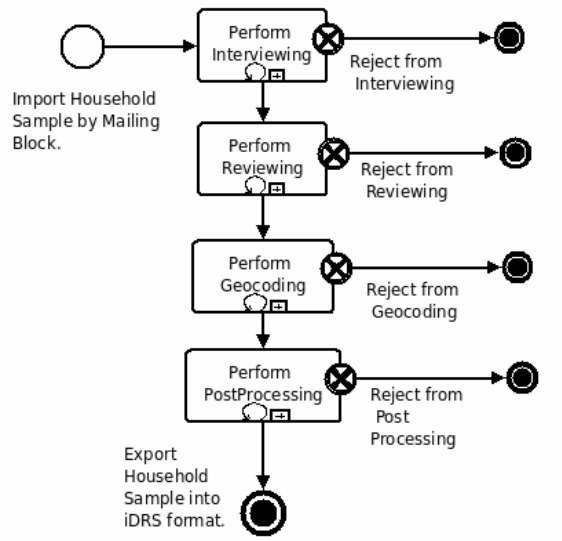
2.10 Sample Management

The 2006 TTS Sample Management System (SMS) unified all aspects of interviewing and the subsequent validation stages within a single environment. This allowed each sample to be identified in full-detail at each step through the interviewing, reviewing, geocoding and post-processing top level stages.

Sample was imported into the SMS prior to each mailing block being sent out. Each record was assigned a unique 6-digit sample identification number, mailing block number and Forward Sortation Address. In 2006 each mailing block was split between the four interviewing team servers according to the relative productivities of each and the number of in progress samples that would be called back during the next shift.

Sample progresses through four top-level stages: interviewing, reviewing, geocoding and post-processing. Figure 2.3 shows the paths sample can follow through the top level stages of the survey. At each top level stage there are three options: the stage is not yet complete, the sample is rejected at that stage or the sample is complete and can be transitioned into the next top level stage.

Figure 2.3 Sample Lifecycle



The Sample Management System (SMS) server software controls access to the sample and invokes a transition process nightly at 2:00 am that transitions samples between the top level stages. Access to sample is controlled through a variety of sample queues for Interviewers and

Geocoders. These queues supply the sample when the interviewer or geocoder requests *any available* sample. Reviewers manually searched for a household to view and Post Processors used a sophisticated search query interface to identify which samples were most in need of additional work.

The Administration Client (AC) was used to apply the management control on the SMS; in addition to the above management features it also allowed:

- Activation/Deactivation of Mailing Blocks.
- Activation/Deactivation of FSA's.
- User creation and role assignment including role specific details like assigned languages for interviewers and geocoding zones for the geocoders.
- Generation of interviewing and geocoding performance statistics for weekly, monthly and arbitrary date ranges.
- Control of which optional batch processes were executed during the nightly rollover process. Only the transition from Interviewing to Reviewing was automatic. The transitions from Reviewing to Geocoding, Geocoding to Post Processing and Post Processing back into Geocoding all required manual Management intervention.

Daily monitoring of the disposition of samples in each stage of the survey using both real-time and daily generated reports was used to determine:

- Changes required in the mailing schedule.
- The appropriate time to activate a new mailing block.
- The number of geocoding samples per GeoZone.
- The appropriate allocation of interview staff to interview stations.
- The de-activation of FSAs that had achieved their completion targets.

2.11 Publicity

Previous surveys indicate three constituents need to be informed about the objectives of the survey and, in varying degrees, about the methods used to conduct the survey. The constituents are the local government and public service officials (particularly the police), the press and households scheduled to be interviewed.

2.11.1 Letter to Local Officials

The best organization to compile and distribute information to appropriate recipients was judged to be the funding agencies. A package of information was compiled by the TTS Management Team. Appendix A contains a sample of this package. The distribution lists were generally made up of the following officials:

- Federal and Provincial Members of Parliament
- Regional Chairpersons
- Mayors, Reeves and County Wardens
- Local Councillors
- Police Departments
- Chambers of Commerce

2.11.2 Press Release

In previous surveys, a press release package was sent to newspapers, television and radio stations in the survey area. In 2005 and 2006 dissemination of information about the survey to the media was left to the discretion of the Funding Agencies and Local Officials.

2.11.3 Advance Letter

The advance letter sent to all selected households was regarded as a critical item in the conduct of the survey as it encourages a high response rate and minimizes the time interviewers need to spend explaining the survey. A copy of the advance letter used for the fall 2005 component of the survey bearing the signatures of the Minister of Transportation and the Regional Chairs (Niagara, Waterloo), City or Town Mayors (Barrie, Peterborough, Orangeville, Kawartha Lakes, Brantford), and County Wardens (Dufferin, Peterborough, Simcoe, Wellington) for the participating agencies outside the GTHA. A copy of this letter is contained in Appendix B. The original letter used for the Fall 2006 component within the GTHA was signed by the Minister of Transportation, the City Mayors (Toronto, Hamilton) and the four Regional Chairs (Durham, Halton, Peel and York). A copy of the advance letter used for the GTHA component of the survey is contained in Appendix C. Starting in November, 2006 both French and English letters were mailed to all selected households. A copy of the letter in French is contained in Appendix D.

Standard Ministry of Transportation envelopes were used for the mailing of the advance letters for all components of the survey. The use of an official government envelope was regarded as important in giving legitimacy to the survey and ensuring that the advance letter not be treated as junk mail.

2.11.4 MTO Info

As in previous years, MTO Info fielded questions from the public regarding the survey. Between August and December, 2006 this amounted to almost 400 calls. MTO provided a weekly summary of these calls which included:

- Questions about the legitimacy of the survey (14%).
- Requests to be removed from the sample base (52%).
- General inquiries and comments (34%).

Inclusion of the survey site phone-in number on the advance letter might have reduced the number of calls received by MTO info.

Section 3 Software Development

3.1 System Design

A total system redevelopment process was undertaken prior to the 2006 TTS. This involved addressing the deficiencies identified in the previous FoxPro based system. A software development process was initiated in late 2003 with the specification of required features followed by an iterative milestone-based development process.

Every 3-4 weeks a new development version of the software was released and distributed to both internal and external testers with subsequent feedback driving the next milestone. This iterative development process worked well in keeping the development effort focused on the specific deficiencies that needed to be addressed at any given time.

Key improvements with the 2006 TTS Sample Management System (SMS) are:

- Client/Server architecture with unified sample management in a server side database.
- Sample allocated one at a time to users versus a shift worth of work being allocated as in the 2001 TTS. This allowed real-time reassignment of active samples requiring a callback if the owning interviewer was unavailable and/or busy for too long.
- Snapshot of complete history of each sample at each point through the survey process allowing management to track when changes occurred and helped improve the quality of the collected data.
- Improved validation with over 150 logic checks shared between all stages of the survey process. All logic checks were run in all phases with each specific subsystem only considering the set of errors that it was interested in.
- The separate server sample database allowed daily and real-time Structured Query Language (SQL) scripts to monitor the survey progress in a multitude of ways. It was much easier and safer to gather information using SQL versus modifying the SMS to provide it.

3.1 Sample Management System (SMS)

The Sample Management System (SMS) is at the heart of the 2006 TTS. It provides the mechanisms to distribute sample from the server database out to interviewers, reviewers, geocoders and post processors. Further, it allows the monitoring of interviewers using the Monitoring Console (MC) and itself using either the Administration Console (AC) or, for special technical cases, the SMS Management Console.

The SMS maintains the knowledge and processing rules to determine what samples a given user has access to (sample check-out) and how returning sample should be categorized based on its present and previous dispositions (sample check-in). The rules for each case are fairly complex and have been tailored based on experience to minimize staff effort while maximizing the probability of each sample being successfully completed.

The SMS is designed to work fairly quickly in distributing loaded samples to requesting users and loading specific samples outside of what is cached. The SMS maintains in-memory sample queues for:

- In progress interviewing samples ordered by callback time, ascending, grouped by interviewing team, and owning interviewer. Ownership is defined as the last interviewer to make substantive changes to the sample.
- Uncalled interviewing samples ordered by sample number, ascending, where: their FSA, mailing block and themselves are presently active.
- Geocoding samples ordered to prioritize processing newer data ahead of existing backlog to facilitate geocoding callbacks occurring as early as possible.

Each sample queue was populated by its own dedicated server thread that would run every thirty seconds and initiate the sample loading process as required. For the uncalled sample and in-geocoding queue this involved checking if the total number of currently queued samples was below 50%; for the interviewing queue there was no limit to the number that would be loaded.

At 2:00 AM each night the SMS automatically launched the nightly transition process that invoked a series of external batch processes and sample transition processes that served to automatically convey eligible samples forward in the survey process and update their state for the next day's shift. For example when a sample without a trip date is checked out by an interviewer, the SMS will assign the current trip date; this date is incremented each night during the transition process.

3.1.1 Sample Check-out Processing

Samples can be checked out from the SMS by either asking the SMS to issue a sample at its discretion or to specifically retrieve the sample by either its unique sample number or 10-digit phone number.

In the specific check-out case only samples that are in the same top level stage as the role of the connecting user can be accessed successfully. For example, a geocoder can only specifically request a household that is presently in geocoding; if the sample is not yet in geocoding or has transitioned into post processing it will not be accessible.

A sample can only be worked on by a single user at a time. The SMS check-out processing infrastructure uses a combination of Java object locking and database locking to ensure that concurrent requests for the same sample will only succeed for one request and fail for the rest. Pessimistic database locks used when adjusting the state of a sample guaranteed this behaviour.

When checked out a sample contains:

1. Sample Details: including the disposition of the sample in the overall survey process and each phase; the callback time for interviewing; the date when it was transitioned into reviewing, into geocoding and into post processing as applicable.
2. Household, Person, Trip and Transit details.
3. Runtime state: for interviewing this was a map of questions that had been asked which was used when locating the previous and furthest question to display; In Geocoding this was a table containing the original location details for all the locations in the household at the point which the household was passed into geocoding.

4. Previous Transaction Details (for each: check-in user, check-in group, transaction time, sample disposition details as they were at this point in time).
5. User comments associated to the previous transactions.

3.1.2 Sample Check-in Processing

The important disposition codes for each survey stage are those that denote completeness of that stage. The editable clients (DDE and GC) were aware of the set of validation errors that were strictly not permitted to exist in a household that was coded as complete. In order to return the sample the user would have to choose one of the other available dispositions.

The details returned to the SMS from the editable client match what was sent except that the Household, Person, Trip and Transit data contain the changes and a copy of all the detected errors.

The check-in processing steps were:

1. Archive the household, person, trip and transit data with the detected errors and runtime state. This part of the process creates a unique number, known as the transaction identifer, which can be used to refer to any part of this data in the future.
2. Determine what the next disposition of the checked-in sample should be and if it should be immediately cached by the SMS or left to be loaded later.
3. Update the sample details in the database to reflect the next disposition.

The main complexity of these processes was contained in step 2 with in-interviewing samples involving the most work due to the need to decrease the work associated to low probability of ever being completed sample. For example checking if the total calls made exceed the total call limit or if the present disposition was a voice mail what the next state and callback time should be.

For the geocoding and post processing cases the complexity was with the boundary cases in which samples were passed back and forward between them. For reviewing there were no similarly complex cases.

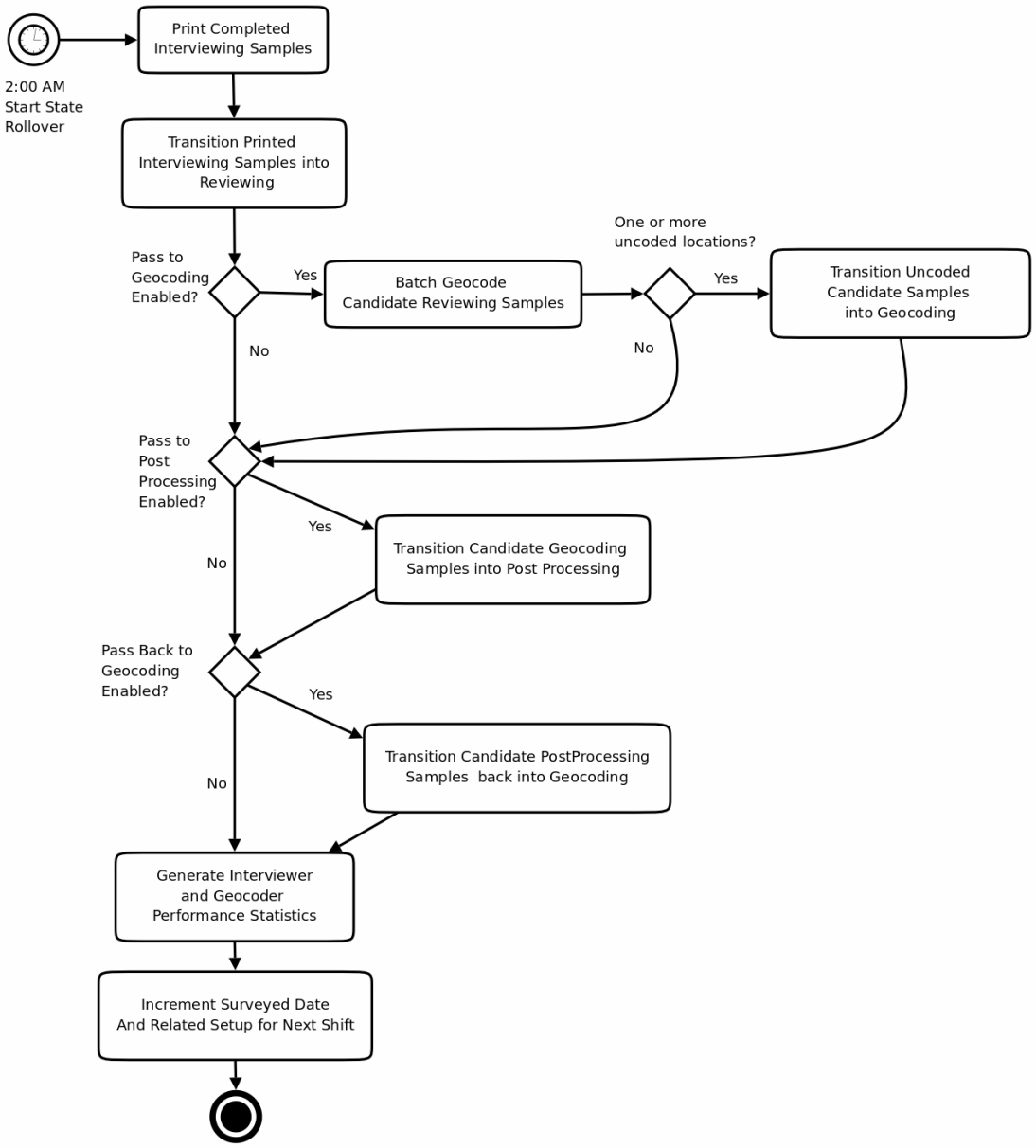
3.1.3 Nightly Transition Process

Figure 3.1 illustrates the Nightly Transition Process launched by the SMS at 2:00 a.m. that would:

1. Automatically transition completed interviewing samples into reviewing.
2. Change the trip date according to preset rules or a specifically defined override defined in the AC. Uncalled samples or in-interviewing but without recorded trips would use this value as the trip date being surveyed about.
3. Generate interviewer and geocoder performance statistics text files.
4. Generate a PDF file containing the household print out for each completed sample ordered by team and interviewer.
5. Generate a PDF file containing the household print out for each sample that is presently in geocoding and needs to be called back.
6. Generate the Transit Error and the School Code PDF reports which display in a table the sample number and page within the nightly printout package where problem sample exist.

7. Execute the Pre-Geocoding Batch on samples being passed into Geocoding if the 'Pass to Geocoding' switch is enabled and the date defined in the AC.
8. Transition the samples that were processed in step 7 into geocoding if they have one or more uncoded locations or directly into post-processing if they do not.
9. Execute the In Geocoding Batch process on samples in Geocoding if the 'In Geocoding Batch' switch is enabled and the date defined in the AC. This is used to automatically apply reference data updates to samples that still require geocoding. The reference data updates are focused on those that occur with the highest frequency and this batch process does reduce the amount of manual work required.
10. Execute the Post Processing Batch process if the appropriate switch in the AC is enabled. Before samples can be worked on by the post processor this batch needs to have been run. The initial batch process is used to detect the post processing specific errors that are used by the post processing interface focusing on the most problematic errors first.

Figure 3.1 Nightly Transition Process



3.2 Direct Data Entry (DDE)

The Direct Data Entry (DDE) is used in Interviewing, Reviewing and Post Processing to enter and edit Household data.

The DDE interface uses a layered approach where previously entered information is visible to the interviewer as they record each descendant piece of information. For example, in the transit case the interviewer is able to see the trip details for which transit was used, the summary details on all persons in the household and the summary details on the household level responses. During development it was determined that a screen size of 1280x1024 would be required to fit the necessary details at a size that would still be visible.

The architecture of how household data is stored and quantified was initially developed within the confines of the DDE. Once it worked it was separated and then shared first with the Geocoding Console and then with the Nightly Batch processes so that there was only a single validation system used by all parts of the system to assess the over 100 distinct consistency, distance and geocoding errors.

3.2.1 SMS Provided Interviewer Features

As interviewing is the most intensive phase of the survey it has the most elaborate SMS provided sample management features.

The SMS utilized several sample queues that were queried to provide each interviewer with sample in order of priority:

1. User specific queue for samples owned by the interviewer with a relevant respondent scheduled callback time.
2. User specific queue for samples owned by the interviewer with a relevant non-respondent scheduled callback time.
3. Team specific queue for samples with a relevant respondent scheduled callback time.
4. Team specific queue for samples with a relevant non-respondent scheduled callback time.
5. New uncalled sample queue containing enabled samples with an active mailing block and FSA.

This unified queue changed how samples were issued to a use-based system. Previously an entire shift's work was issued to each interviewing station whereas in 2006 all the sample resided on the server and was issued one at a time to logged in interviewers.

The unified sample queue allowed for expiry related features to be enabled that allowed samples owned by one interviewer to be issued to another member of the team. This was used when the specified interviewer was unavailable or at the end of the shift when it is very important to make sure that calls specifically scheduled for a callback are made that night.

3.2.2 Management Control Features for Interviewers

The AC provided the management team with these capabilities for interviewers:

- Assignment of additional languages to interviewers. SMS will only issue sample automatically to users that speak the language of the sample.
- Uncalled sample only mode. This mode granted an exemption to a user from the standard queue policy and instead drew all of their samples from the uncalled sample queue.

For additional information on the new Data Direct Entry client please refer to the report: "2006 Transportation Tomorrow Survey Working Paper Series: Interview Manual".

3.3 Geocoding Console (GC)

As part of the overall development of new software to conduct the 2006 TTS, new Geocoding software was developed. The software, called the Geocoding Console (GC), was designed largely on the prior edition of geocoding software but also included some improved features.

Geocoding occurs after the Interviewing and Reviewing stages of the survey process. Typically the Pass to Geocoding transition process will be run on samples within a week of completing the interviewing process.

Before samples are ever examined manually the Pre-Geocoding Batch will have run on each sample being passed in from Reviewing; it will have evaluated each candidate sample in turn and undertaken an automated attempt to geocode as many of the trip locations collected as possible. This is the first point at which monuments will be attempted to be coded.

If there are one or more uncoded locations in a household after the batch it will transition into Geocoding; if there are zero uncoded locations then it will transition directly into the Post Processing stage of the survey.

Within the Geocoding stage samples will be coded interactively using the Geocoding Console. The coders will work through the sample backlog one at a time attempting to solve all of the highlighted geocoding errors using its built-in reference database and other external aids such as phone books, internet search engines and hard-copy maps to find valid coordinates for the locations which need coding.

For a detailed description of the new Geocoding console please refer to the report “2006 Transportation Tomorrow Survey Working Paper Series: Coding Manual”.

3.3.1 Coding Reference Database

The coding reference database contained within the Geocoding console consisted of an address range file, an intersection file, a monument or landmark file, a school file and two place name files.

Additional to these files, lot and concession maps were obtained from both Simcoe County and Dufferin County for assistance in coding some of their households. These files were used for geocoding but were not added directly to the Geocoding Console database.

a. Street Address File

In prior surveys, the street networks and intersection files had been provided for each region by the individual regional agencies involved in the survey. This led to a variety of files being acquired which came without a standard format or standard datum and contained differing sets of information. These files would then have to be given a standard datum, converted to a standard format and combined to generate the street and intersection files for the entire survey area. This required an immense effort to generate the final street network files.

For the 2006 TTS, the street network file was acquired from Land Information Ontario (LIO). LIO is a department of the Ministry of Natural Resources Ontario with responsibility for the

management of geographic information for use in maps and Geographical Information Systems (GIS) technology. In 2005, LIO made available an Ontario Road Network (ORN) file which contained map data on all streets and intersections in Ontario. The availability of this data, containing in one file for the entire TTS area, eliminated some of the processing necessary to create a street file for TTS purposes that had been needed in previous years.

The ORN was obtained by TTS as base user data under an agreement between LIO and the University of Toronto. The ORN street network file was obtained as a Linear Referenced Dataset in the Standard NRVIS Interchange Format (SNIF) which was converted into ESRI shapefile and MapInfo table format for use as the basis of the TTS Street Address files.

The information included in the Street Address files used by TTS included the street name, the cross street name, the number ranges on both sides of the street, the coordinates of the start and end segments of the street and the municipality where the street was located.

b. Intersection File

An intersection is defined as the centre point where two or more streets meet. Intersections are identified by locating all the common nodes in a street network. The Intersection file was generated from the ORN street network file.

c. Monument File

To identify a particular location, it is common to use a monument name instead of a street address. A monument may be a building or landmark, such as the CN Tower or the Eaton Centre. In 2006, a new Monument file was generated. It was based on some of the landmarks listed in the 2001 file but was started from scratch as some of the old landmarks were no longer valid and the TTS area in 2006 had expanded to include areas that weren't surveyed in the 2001 study. Examples of places added to the Monument file included major shopping malls, hospitals, supermarkets, popular tourist attractions, major workplaces, regional and local government offices, sports arenas and big box stores. The addresses for these monuments were located through the use of street maps, internet directories and telephone books.

The collection of landmarks was geocoded and stored in the Monument file. The file strived to be as complete as possible but was by no means exhaustive of landmarks in the TTS area. To be as up-to-date as possible, landmarks which occurred in the interview and geocoding process with some frequency were constantly added to the Monument file as the survey progressed. The Monument file contained such information as the monument name and a special id code, its address, municipality and coordinates.

For some areas, street addresses were not available or respondents only knew the Lot and Concession numbers of their residences. For these incidences, coordinates were calculated using the Lot and Concession maps which were on hand and the locations and coordinates added to the Monument file for input to the household via geocoding.

d. School File

In 2001 a new unique attribute, school code, had been added to the database. This necessitated the use of a new file called the School file. In 2001, the School file contained only the school code

and the name of the school while the Monument file also kept the school name and the relevant address and geocode information. In 2006 all school information including address and geocode coordinates was kept separately in the School file.

The School file was first generated by re-geocoding the existing 2001 school file using address information obtained from the Ontario Ministry of Education's listings and internet listings of private schools.

While from the start this file contained the majority of public, high and tertiary institution locations, this file underwent considerable updating as the survey progressed as new schools especially language schools, new private schools and trade schools were found in the survey and added constantly.

e. Internal Place Name File

The level of geocoding accuracy varied throughout the survey area. The goal was to geocode information to as much detail as possible. Street addresses and monument locations were preferred over street intersections. However there were certain situations where (non-work, non-school) trips were made where the locations could not be ascertained to that level of detail. Attempts, including additional phone calls to the household, were made to get as much detail as possible but failing this, as a last resort, the locations were sometimes coded to the place name.

The Internal Place Name file contained the names and geocode coordinates of places within the survey area. Its use was kept to a minimum during the survey.

f. External Place Name File

Some members of the households which were surveyed occasionally had trips on the trip date which went outside the survey area. Examples of these include trips to the U.S or to places within Canada which were outside the survey area for example Montreal, Quebec, Windsor or Kingston, Ontario. In such cases the location of these trips was coded to the external place name. Names and geocodes for these external places were determined and entered into the External Place Name file.

3.3.2 SMS Provided Geocoder Features

A Geographic Zone was defined as:

- A unique number.
- A set of Forward Sortation Addresses and/or Local Delivery Units.

In 2006 there were 58 distinct Geographic Zones that contained roughly even population; they were used as the basis of delivery of samples to requesting geocoders. With a unified sample queue each geocoder would request *Any Household* from the SMS and be automatically issued a household from one of the Geographic Zones for which they were assigned.

The SMS provided a sample queue for each Geographic Zone of length 10 which would be automatically refilled if the number available fell below 5.

For each zone the sample queue was populated according to these priorities:

1. Samples just passed in from Reviewing and/or passed back from Post Processing were ordered by the date they were passed to geocoding (descending with most recent date first). This allowed the most recently completed households to be geocoded first.
2. Samples that are incomplete and still require more geocoding ordered by their last transaction (ascending with the oldest samples first).

This structure allowed newer data to be processed ahead of the existing backlog and to facilitate geocoding callbacks occuring as early as possible.

3.3.3 Management Control Features for Geocoding

The AC provides two main ways for managing geocoding users:

- Assignment of Geographic Zones to geocoders. The SMS will only issue sample to geocoders assigned to that samples zone.
- Assignment of the 'Geocoder Supervisor' permission with the power to re-geocode an already coded location. This was needed for several cases where the batch geocoding worked but was actually an incorrect location that was determined by a later reference update.

3.4 Monitoring Console (MC)

Monitoring individual interviewers as they conduct the survey is an integral part of the quality control process. In the 2006 TTS two supervisors per team conducted visual and auditoral monitoring during the course of each interviewing shift.

In order to access the MC, a table was presented to the supervisor with a row for each connected user containing:

- Their IP address.
- Their username.
- Time they logged into the system.
- The sample number for the household currently checked out.

By selecting a row and pressing a button it was possible to see through Virtual Network Computing (VNC) exactly what was occurring on the user's screen. Concurrent telephone monitoring allowed a comprehensive assessment of the interview in process.

The monitoring stations used a monitor resolution of 1600x1200 to facilitate viewing the entire 1280x1024 screen the interviewing was being conducted with.

3.5 Administration Console (AC)

The Administration Console (AC) is the client application through which the management team interacted with the SMS. It provides toggles for each stage of the survey process and user role specific properties as described in the previous DDE and GC sections. All actions within the AC interacted directly with the SMS and had an immediate effect.

3.5.1 Sample Management Operations

- Activation/Deactivation of Mailing Blocks. This caused the SMS to schedule a full reload on its next update. Performing this operation during a shift could lead to a temporary *no more samples available* message during the delay it took to reload.
- Activation/Deactivation of Forward Sortation Addresses and Local Delivery Units. As with mailing blocks this operation would cause a reload of un-called samples.
- Creation and Permissioning of Users. Changes were effected in real time with the next interaction of the user with the system.
- Editor for setting an alternate trip date for a specific target date. This allowed management control to, for example, exclude weekday holidays from the period surveyed. A set of alterations could be set at anytime for any future days.
- SMS Queue inspector. This provided a list of all samples presently loaded by logical queue and a mechanism to view the transactional record for any of them.
- User Work History List. This provided a way to see for a specific role and user the work (recorded check-in's) they had performed that day.
- Sample Transaction History Viewer. Allowed a management user to load the transaction history of any sample known to the SMS. This could be used to identify samples that had been in one stage and had since transitioned forward in the survey process.

3.6 External Reporting

The sample database managed by the SMS ran on a PostgreSQL database. This allowed a read-only user to execute queries and build an assortment of real-time as well as daily reports.

3.6.1 Real Time Script Examples

Real time report queries were used by management to track statistics within the current shift like: number of households to be called back before the end of the shift and the current number of samples that have been completed so far in today's shift.

Over 80 scripts were created and used during the 2006 TTS. They allowed monitoring of:

- Survey Stage related: total in all top level stages; breakdown of sub-stage for each top level stage (i.e. last interviewing disposition for in-interviewing samples or in-interviewing with a specific trip date or the number of interviewing samples per disposition that will be callable tonight).
- User specific: total work a user has done tonight by sample disposition; completion statistics for a specific user for tonight; count of sample dispositions for geocoding samples by username.
- SMS specific: samples presently loaded in the SMS; sample available for each Geographic Zone; number of connected users.
- Sample tracking: current sample details to identify where it is; all transactions including user comment for a specific sample number; all scheduler transactions (times where the SMS loaded the sample into memory) for a specific sample number; find a sample based on the phone number contained in any of the comments for any of its transactions (useful where the phone number has changed and respondent calls in and we need to identify the sample number).

3.6.2 Daily Report Examples

Each night at midnight, before the nightly transition process occurred, a set of reports were generated from the sample database. These reports are described below.

a. Summary Reports:

Overview: top level stage disposition totals; disposition of samples in interviewing; disposition of samples in reviewing and disposition of samples in geocoding.

Rejected then complete*:* Samples that were rejected due to the maximum call limit but were subsequently completed.

Complete then Rejected*:* Samples that were completed through interviewing but rejected in the reviewing stage.

b. Interviewing Reports:

Current Uncalled and In-Interviewing: uncalled samples by mailing block; disposition of samples in interviewing; count of samples requiring callback in a language other than English.

Interviewing Voice Mail Three Daily Report(effective in January 2006): non-contacted households that will be called back tonight and left the long voice-mail message.

Interviewing Day Time Callbacks: a report containing the list of samples that have a respondent scheduled callback time for today between 10:00 am and 5:30 pm.

Non-English, Non-Other Language Callbacks: a special report that tried to identify households where the language was not one of the main supported languages. It classified the samples based on the contents of the interviewer comments left.

Out of Range Callbacks*:* a report designed to find households that had been scheduled for a callback after the interviewing data collection phase is scheduled to be terminated by. The DDE would prevent an interviewer from scheduling a callback years in advance but near the end of the survey process this report was useful.

Interviewing With Trip Date: samples that are within the interviewing process and have partial trip data recorded.

Samples Close To Call Limit: list of samples that are one or two calls away from being rejected based on the call limit being reached.

c. Reviewing Reports:

Reviewing Overview: total number of samples that will be held back from the next 'Pass to Geocoding' operation; maximum printed date that has been passed into geocoding as of today; total number of samples by printed date that are eligible to be sent into geocoding. Total number of samples by printed date that are currently being held back from geocoding.

In Reviewing with Critical Errors Present: total number of samples containing a specific 'critical' error (something that needs to be fixed before geocoding); disaggregated list of sample number by error code.

d. Geocoding Reports:

Current and Potential Disposition: total samples by current geocoding disposition (a sense of the work presently available); total number of samples in reviewing that are presently eligible to be passed into geocoding (a sense of the work potentially available if passed through tonight).

Samples Requiring a Reference Update: details on the samples currently awaiting a geocoding reference update to occur showing the last coder, specific geocoding errors and last SMS transaction time.
Geocoding Samples Per FSA: count of geocoding samples per FSA and/or LDU that are not presently passable into post processing, passable into post processing and currently within geocoding.
Geocoding Callbacks: list of all samples that are in geocoding but require a callback to clarify certain details.
Geocoding Incomplete Samples: list of in geocoding samples that have been coded as incomplete (more geocoding required).
Rejected in Geocoding: list of samples that completed the interviewing and reviewing stages of the survey but have been rejected in geocoding.
Uncodeable Location: List of locations in samples presently in geocoding that have been marked as uncodeable. This is used for quality control to make sure all avenues have been explored before a location is marked this way.
Uncoded Locations: a report for each of Monument, Intersection, School and Address that details the number of such locations that are uncoded (indicating a gap in the reference data); report includes: total references, location name and municipality.
In Geocoding Missing School Code: list of persons in geocoding where their usual place of school does not contain a valid school code (indicating a gap in the reference data).
Non-Complete Error Summary: total number of errors for geocoding samples that are not currently passable into post processing.
Complete but Uncodeable Key Locations: Sample disposition is completed geocoding but one or more of its key locations is marked as uncodeable and the sample will be automatically rejected when passed into post processing.

e. Post Processing Reports:
Error Summary: total number of samples by post processing error alias.
In Post Processing Missing School Code: list of samples that have school locations defined but are missing the associated school code.
Home Address is an External Place: list of samples where the home address is geocoded into a place that is external to the survey area.
Home Address is an Internal Place: list of samples where the home address is geocoded to the centroid of a place within the survey area.
In Post Processing with Uncodeable Key Locations*:* list of post processing samples that contain uncodeable locations for any of home address, usual place of school and usual place of work.
Uncoded Locations: a report for each of Monument, Intersection, School and Address that is exactly the same as the in geocoding version except it contains only those errors that occur with samples in the post processing phase.

3.7 Operating System

In the 2006 TTS all of the computers were set-up using the Debian Linux stable distribution. Both had local PostgreSQL database version 7.4 (in Fall 2005) then version 8.1 (in Fall 2006).

On the servers there was a reference database and sample database containing all of the sample.

On the workstations there was a reference database and a sample database designed to hold a single sample record corresponding to what is presently checked out by the DDE or GC that is running.

3.8 Open Source Components

While much of the software created in support of the 2006 TTS was written specifically for the survey many infrastructure related tools that facilitated its implementation used open-source components.

Java (java.sun.com)

- Hibernate Object Relational Management (www.hibernate.org); for managing the persistence of java object data within the database.
- Quartz scheduler (www.opensymphony.com/quartz); for defining time based jobs like the nightly rollover process.
- Joda Time (joda-time.sourceforge.net); advanced time and date functions used when computing callback times.
- iText (www.lowagie.com/iText); PDF reading and writing API used to create the school and transit daily reports that cross-referenced the page numbers from the nightly printout bundle.
- Spring Framework (www.springsource.org/about); Dependency Injection framework for simplifying project setup and enhancing testability. Used to setup the final export process which converted the 2006 TTS production database content into the iDRS database format.
- Eclipse (www.eclipse.org); Java development environment used to develop and debug the system.

Debian GNU/Linux (www.debian.org)

Used as the operating system of both the workstations and servers.

Kiosktool (extragear.kde.org/apps/kiosktool)

Used to create the limited KDE profiles under which the TTS software ran.

SAMBA (www.samba.org)

Used to provide separate network shares for management users and geocoders.

PostgreSQL (www.postgresql.org)

Used as the production database management system.

OpenVPN (www.openvpn.net)

Used to give management users secure access to the Data Management Group internal systems.

TightVNC (www.tightvnc.com)

Used by the MC to handle the actual mechanics of viewing an interviewing station remotely.

Python Bittorrent Headless Client by Bram Cohen

Used to efficiently distribute reference and software updates over the network. Customized to invoke the client-payload installer after it had been 100% downloaded from the swarm.

Bering-UClibc (leaf.sourceforge.net/bering-uclibc)

Linux Embedded application Firewall used to manage internet access.

Dansguardian (www.dansguardian.org)

Used to apply browser site restrictions to geocoders.

Squid Proxy (www.squid-cache.org)

Used with Dansguardian to apply browser site restrictions to geocoders.

Lighttpd Web Server (www.lighttpd.net)

Lightweight web server used to serve nightly generated HTML reports for each SMS.

G4u Disk Imager (www.feyrer.de/g4u)

Used to create client computer images and to image new computers over the network.

Section 4 Equipment

The design and structure of the 2006 TTS network drew heavily on what had been done in the 2001 TTS. The main differences were the new Sample Management System that ran on higher-powered server computers and that all except one management computer used the Debian Linux operating system. Each of the non-server computers had a valid Windows 2000 or Windows XP license which was retained for resale purposes.

4.1 Computer Network

The wiring structure of the computers on the floor was similar to the 2001 TTS. The amount of wiring necessary was minimized by locating switches close to each team and linking only 1 wire from each team to the core switch located with the servers. Teams with multiple switches were accommodated by cascading the switches together.

Two networks were created:

1. The 10.10.0.0/16 main 100 megabit network that contained the servers, printer and client workstations. The main network was primarily used to transfer samples between the client workstations and the server computers.
2. The 192.168.0.0/24 gigabit network linking the four servers together. The server network was primarily used to transfer reference update and backup files between the servers at the end of each shift.

The main network was allocated from the 10.10.0.0/16 network range with the following structure:

Team A 10.10.1.0/24
Team B 10.10.2.0/24
Team C 10.10.3.0/24 [Fall 2006 only]
Team D 10.10.4.0/24 [Fall 2006 only]
Call in 10.10.5.0/24
Team Leaders/Monitoring 10.10.5.0/24 [10.10.2.0/24 in Fall 2005]
Management 10.10.6.0/24
The 4 servers were split between the 10.10.7.0/24 and 10.10.6.0/24 network.

The host part of each IP address was assigned based on the station number from the floor layout drawing which corresponded to the extension number in the telephone call monitoring system. This allowed the team leaders to easily see based on who was presently logged into the system which phone line they could be monitored on.

A Linux Embedded Application Firewall, using Bering-Ulibc, was setup as the firewall/router between the private TTS network and the University of Toronto network. Network access was provided by the Faculty of Medicine at the University of Toronto.

Only management users were able to directly connect to the Internet. All other computers did not have a default route set and were unable to access beyond the local network. Geocoders were allowed to use the Internet but their access was through the DansGuardian Squid based proxy

which allowed us to restrict access, track which sites they viewed and prevent access to non work related sites like Facebook.

The Lexmark T612 printer purchased during the 2001 TTS was used successfully through the 2005 and 2006 survey phases.

Figure 4.1 Main Network Set-up

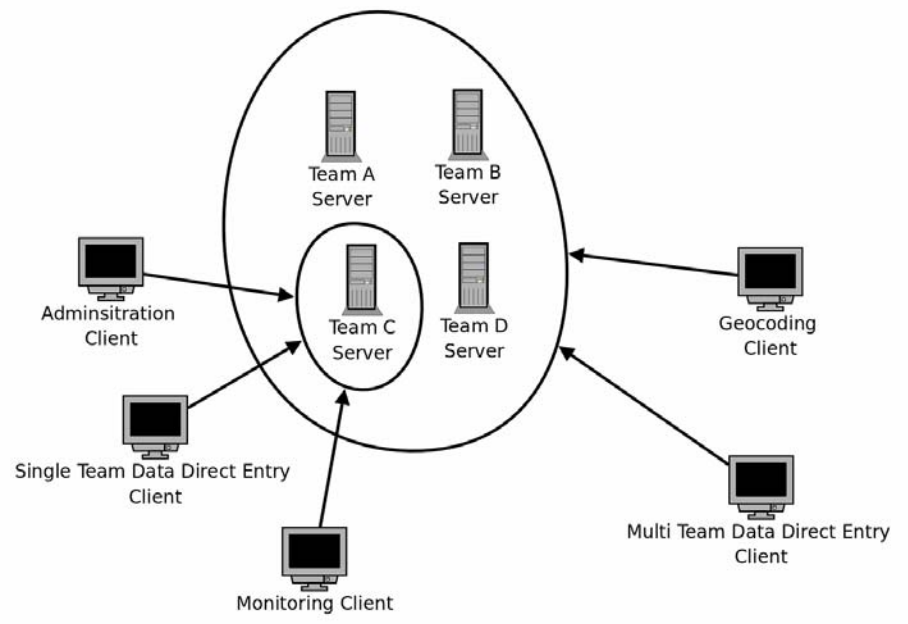
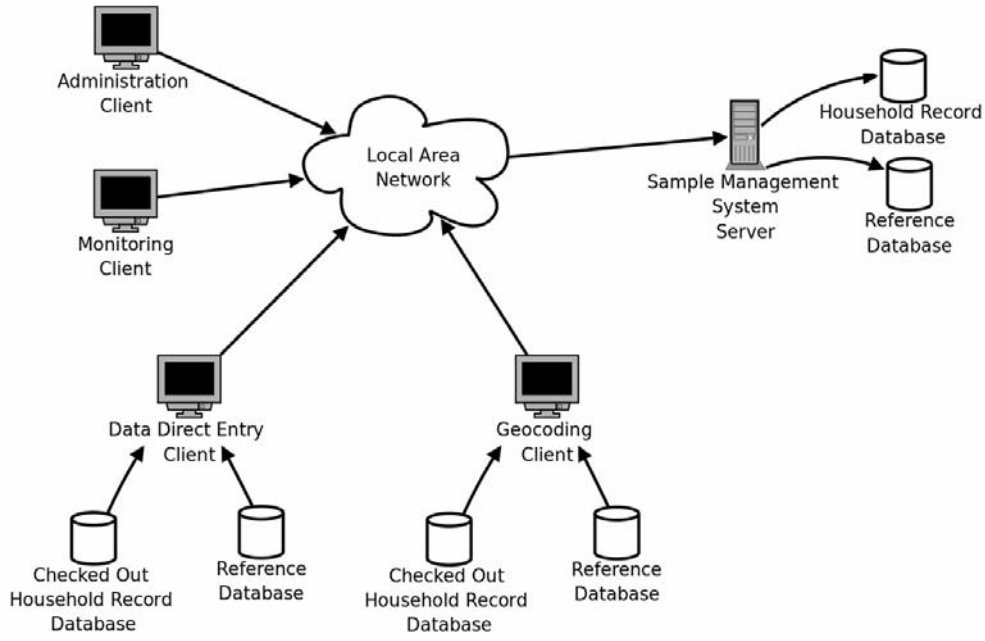


Figure 4.2 Local Area Network Set-up



4.1.1 Servers

2005 Configuration:

Dell Power Edge 1800 configuration:
One 3.0 GHz Pentium 4 processor.
One 2 GB 10,000 RPM SCSI disk.
2.0 GB of memory.
Two Gigabit network interface cards.

There were two servers with this configuration. One of them had an additional Adaptec 2940UW SCSI card through which the Quantum DLT 8000 backup tape was connected. The maximum amount of storage provided by each tape was 80 GB compressed.

2006 Configuration:

Dell Power Edge 1850 configuration:
Two 3.0 GHz Pentium 4 processors.
2.0 GB of memory.
Two Gigabit network interface cards.

Both Power Edge 1800's were given a second processor and the drives reorganized as follows:

tts1 - Dell Power Edge 1800 with two 72 GB 10,000 RPM drives in RAID-0
tts2 - Dell Power Edge 1800 with two 36 GB 15,000 RPM drives in RAID-0
tts3 - Dell Power Edge 1850 with two 36 GB 15,000 RPM drives in RAID-0
tts3 - Dell Power Edge 1850 with two 36 GB 15,000 RPM drives in RAID-0

The tts1 server with its additional storage space was used as the primary file sever (using SAMBA) which provided backed up network partitions for management users files as well as for reference update related files from the Geocoders.

Each TTS server ran the Debian Linux stable version and consisted of these elements:

- Java Sample Management System server application.
- PostgreSQL database for samples.
- PostgreSQL database for reference data.
- Lighttpd web server for displaying the HTML reports generated daily.
- System access for administrators to extract real-time statistics from the sample containing databases.

Over the course of the survey over 60 different SQL queries were encoded into scripts to help better inform decision making.

The Dell Power Edge 2400 from the 2001 TTS was reused as the training server from which all demonstration samples were drawn during each interviewer's initial training period.

4.1.2 Clients

2005 Configuration:

Dell Optiplex GX260:

1.8 GHz Pentium 4 processor.
512 MB of memory.
20 GB disk.
19 inch Trinitron CRT display.

2006 Configuration:
Dell Optiplex GX520.
2.8 GHz Pentium 4.
512 MB of memory.
80 GB disk.
17 inch LCD display.

In August 2005 63 Dell Optiplex GX260's were purchased used from the Engineering Computing Facility at the University of Toronto. They were setup to use Debian Linux primarily because of the requirement of having a local PostgreSQL database on each of them to store the currently checked out sample. This was an essential reliability feature to ensure that information given to us by a respondent would never be lost due to a technical problem like the software crashing or loss of power.

Kiosktool was used to create an extremely limited user profile that locked the user in to only be able to access the TTS software and in the case of Geocoders the Firefox web browser.

A profile was created for the interviewer/reviewer/post-processor, the geocoder and the team leader/monitoring user classes. Their default account was for the DDE and required no password. The geocoder and team leader classes required a password which was only distributed to those authorized to have access. In 2005 a special training profile was created which used a specially configured DDE to talk to the training sample server. In 2006 the training took place on the 5th floor which was an isolated network so the standard client image was used.

In July, 2006 65 Dell Optiplex GX520 systems were acquired new from Dell to provide enough computers for the additional interviewing teams. The necessary setup was configured on one of each of the two kinds of computers and then replicated to all the computers using the G4u disk imaging system. Each computer was then capable of fulfilling any role in the survey. This feature was used to increase interviewing capacity by converting monitoring and reviewing stations into interviewing stations when necessary for the evening shift

In September 2006, 15 Dell Optiplex GX150's with 1.0 GHz processors and 512 MB of memory were lent TTS from the Engineering Computing Facility at the University of Toronto. These systems were used in training allowing the existing training computers to be redeployed onto the call centre floor. At the end of the survey an equal number of Dell Optiplex GX260's were transferred to them as payment.

The 2007 supplementary survey was conducted from within the Data Management Group offices at the University of Toronto using tts3 and a new PostgreSQL database for the new sample. The 2006 TTS computers brought back at the end of the interviewing phase were used to conduct the interviewing.

4.1.3 Backup

Each night at midnight, before the TTS software would transition samples between top level stages, a backup process would run on one of the Dell Power Edge 1800's that was connected to a Quantum DLT 8000 SCSI-2 backup tape drive. A script would be remotely executed using ssh key based authentication to generate database dumps and copy incremental changes to a staging area from which the tape dump could take place.

4.1.4 Resale

Following the 2006 phase of the survey when the leased office space was released most of the computer equipment was re-sold except for the 4 servers and 6 workstation computers to be used for post processing related purposes.

It was easier to sell the newer Dell GX520's with their LCD screens and the balance of a three year warranty than the older Dell GX260's with large CRT monitors. Resale considerations are important as the cost of the computer equipment is based on the net of original purchased price minus final sale price.

4.2 Telephones

In the Fall 2005 survey two Dees CM-30 telephone monitoring units were installed and wired to the 51 analogue telephone lines used by the interviewers. This configuration allowed two supervisors to monitor any of the interviewer lines in each of the two teams. Software was installed on the monitoring station computers to allow the supervisor to visually monitor an interviewer's computer screen at the same time as listening to the interview over the phone. Unlike 2001, the phone lines were not connected through the Province of Ontario's Centrex system. Instead, regular Bell lines were installed which, in addition to incurring additional installation and long distance charges, had the significant disadvantage of not showing the Province of Ontario on the call display when an interviewer called a potential respondent. Instead call display showed "TTS 1-888-xxx-yyyy".

The same telephone set-up was duplicated for four teams in 2006. The interviewer lines totalled 129 with 8 Dees CM-30 telephone monitoring units operating in four banks. A combination of cordless and regular phones was used for monitoring enabling one supervisor per team to move around the room while still performing the monitoring function. There were 141 phone lines in total installed for the interviewing, monitoring, coding and management operations. Again, regular Bell lines were used and the call display showed "2006 TTS 1-888-xxx-yyyy".

Headsets are an important component for interviewers using computers for direct entry of data. The cost of commercial headsets was considered high given the low resale value after only 4 months of operation. Having had previous success using the significantly less costly Plantronics T100 headsets and keypad combination designed for domestic use, a decision was made to populate the floor with them. In previous years each interviewer had been provided with their own headset to plug into the keypad at the workstation. In 2005 and 2006, to keep costs down while still providing for the comfort of the interviewers, each interviewer was provided with their own set of foam ear and mouth pieces for the workstation headset.

Separate phone lines were installed for management functions and to receive call-ins from potential respondents who had been left a voice mail message. These call-in phones were equipped with automatic transfer to another line if the first line was busy or un-answered. With the number of households now using voice mail or answering machines, these call-in responses to messages left at the household were considered very important. Every attempt was made to have these lines answered by a trained interviewer during the day and evening. Otherwise, an answering machine was used to describe the hours of operation and record any message the respondent wished to leave.

Section 5 Conduct of the Survey

5.1 Historical Overview of Survey Statistics

Table 5.1 Historical Overview of Statistics

	1986 TTS	1991 TTS	1996 TTS	2001 TTS	2006 TTS
Number of households in the survey area	1.47 Million	1.71 Million	2.32 Million	2.51 Million	2.87 Million
Target sample	5%	High growth 4.5% Low growth 0.5%	5%	5%	5%
Completed sample	4.2%	1.4%	5.0%	5.5%	5.2%
Sample used (approximate number of letters mailed)	102,606	34,167	158,753	215,000	340,820
Valid contacts	83,764	27,813	139,952	174,000	207,082
Refusal rate (of valid contacts)	25.9%	11.4%	21.8%	21.1%	26.6%
Completion rate (of sample used)	60%	72%	70%	64%	44%
Final Database					
Household records	61,453	24,507	115,193	136,379	149,631
Person records	171,086	72,496	312,781	374,182	401,653
Trip records	313,633	142,453	587,676	817,744	864,348
Transit records	56,615	14,896	70,295	85,095	87,244
Mean household size (expanded data)	2.77 persons	2.77 persons	2.71 persons	2.70 persons	2.68 persons
Trips per person 11 or older	2.35	2.54	2.48	2.54	2.47
Interview stations	86	33	120	120	121
Interviewers & Supervisors recruited	390	75	300	275	370
Coding staff recruited	N/A	6	17	13	14

A household sample becomes a ‘valid contact’ when it has reached the status of complete or refused.

The above interview station and staffing statistics are for the main components of the 1996, 2001 and 2006 surveys.

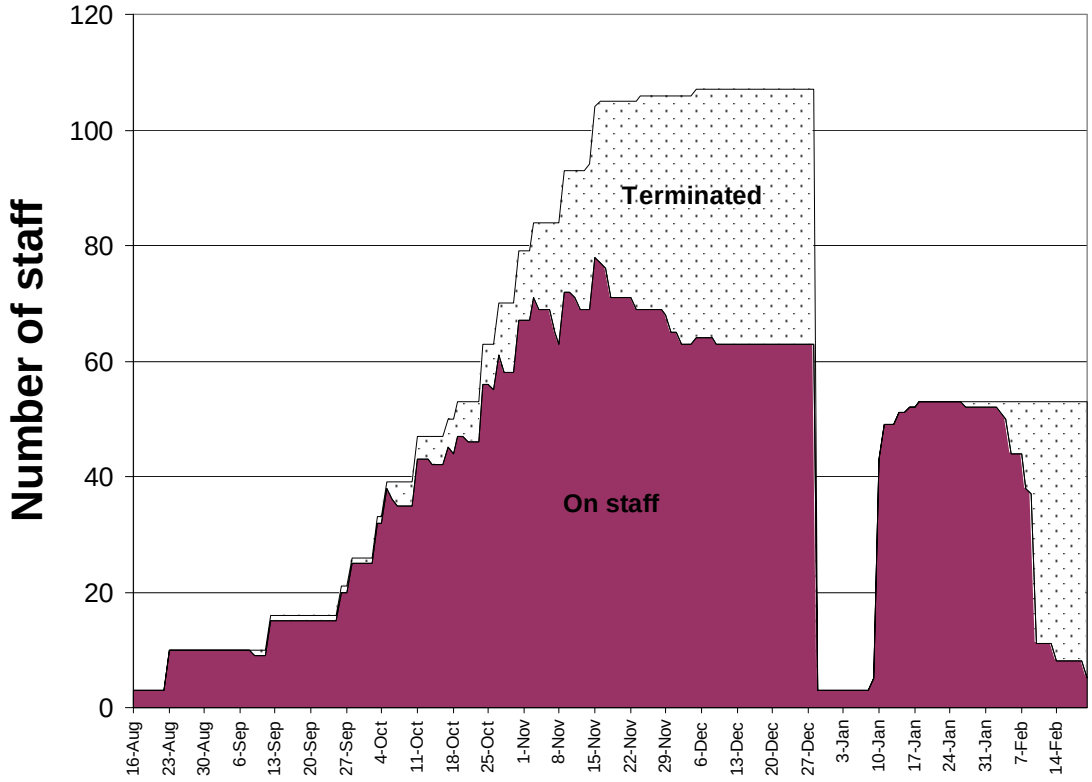
The lower completion rate reflects the increase in households being rejected after multiple unsuccessful attempts to contact them.

5.2 Interview Staffing

The number of interview staff required, together with the need to recruit and train them in a short time, is unquestionably the most challenging aspect of conducting a survey the size of TTS. As in 1996 and 2001, a large number of interviews (more than 37,000) were done in the fall of 2005 thus reducing the target for the main part of the survey to 115,000. The fall 2006 survey was done from the same location (Downtown Toronto) as the fall 2005 component enabling a significant number (28) of the staff hired and trained for the 2005 survey component to be re-hired for the survey in 2006. 12 of these 28 had also been part of the 2001 survey. In addition, another 13 interviewers from 2001 joined the main component of the survey in 2006. Of the total 41 interviewers with previous TTS experience, 30 (73%) stayed on for the duration of the 2006 survey through January 2007, compared to 130 of the 329 interviewers without previous TTS experience (40%). The 4-team leaders for the main survey were selected from the returning staff, as was the chief assistant to the hiring and training manager.

The primary method for recruiting interviewing staff was help-wanted advertisements placed in the Toronto Star newspaper and on "workopolis.ca". An advertisement for geocoders was placed at the University of Toronto's Career Centre. Hiring and training of staff for the fall 2005 component of the survey commenced on August 16th, 2005. A total of 102 interviewers and 5 coders were hired and trained. The maximum number of interviewers on staff in 2005 at any one time was 74 (including team leaders). Figure 5.1 shows how the number of interview staff varied over the course of the 2005 survey.

Figure 5.1 2005 Interview Staff

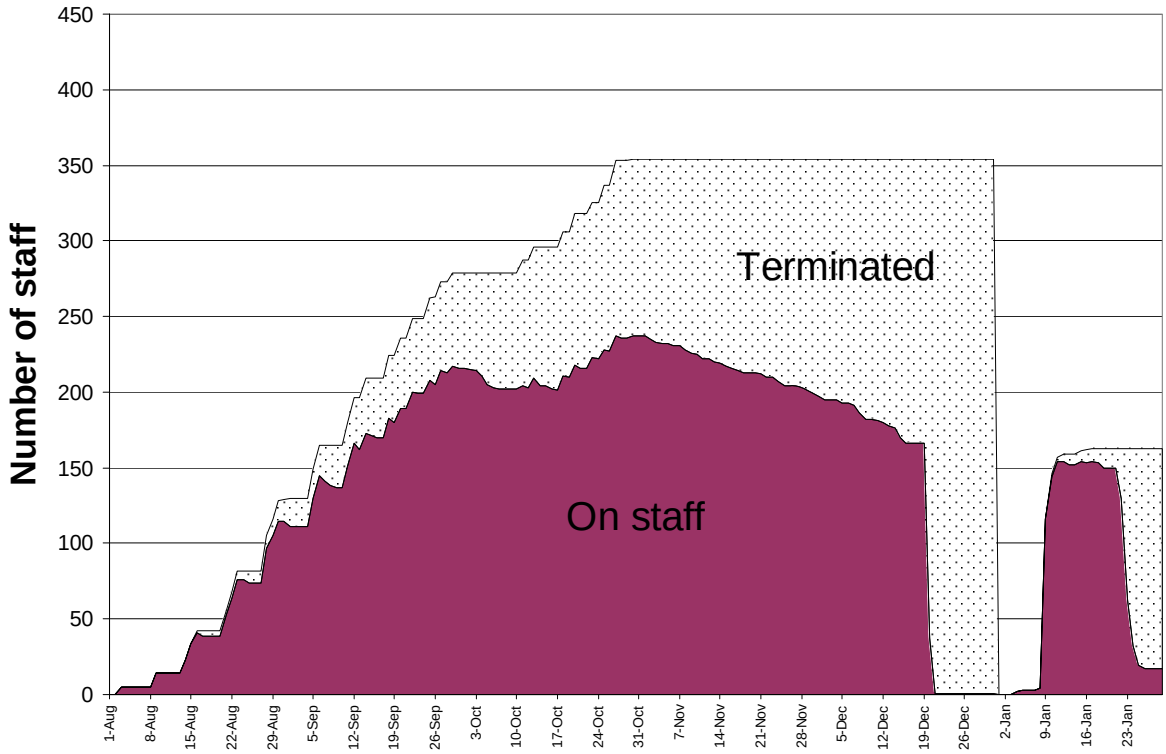


Hiring and training of staff for the main component of the survey commenced August 3, 2006. The availability of the returning staff from the previous year made it possible to have approximately 100 interviewers trained by the time the survey started on September 6, 2006. In total 370 interview staff and 14 coders were recruited over the course of the survey. 21 of the interview staff hired failed to complete the training. The maximum number of people on payroll at any one time was 232 at the end of October. Figure 5.2 shows how the number of interview staff varied over the course of the survey.

The interview staff were organised into four teams. Three of the teams each had a single team leader and the fourth had a pair of team leaders. The leader(s) of each team had the responsibility for the scheduling and supervision of their team. Having an extra leader proved beneficial in easing the load on our less experienced leaders and provided a backup leader in the event that another leader had to be absent, particularly valuable given the 6 days-a-week schedule. A daytime supervisor was appointed with responsibility for ensuring that enough staff was available during the day to carry out functions such as answering the phone and making scheduled callbacks. The scheduling of staff to review the interviews conducted the previous day was the responsibility of the individual team leaders.

Four team leaders agreed to conduct live interviews themselves for the May 2007 supplement for Wilmot. They operated as a single team. Over the course of one week 211 interviews were completed.

Figure 5.2 2006 Interview Staff



5.3 Training

The initial training program consisted of three consecutive evening sessions for each new group of 9 to 16 interviewers (average size 11 people). A maximum of three groups a week were trained. In August and early September training usually starting on Monday, Tuesday and Wednesday evenings which allowed each group to complete training in the same week. In mid-September Monday, Wednesday and Friday starts were implemented to make the best use of the available training space. The Friday group had their second session on Saturday and completed their training the following Monday.

The first evening of training consisted of a detailed demonstration of the software by the Hiring Manager. The demonstration, with appropriate time for questions and answers took 2 to 3 hours. The trainees spent the remainder of the four-hour shift, working in pairs, familiarising themselves with the software.

On the second day of training, the candidates practised interviewing each other. Supervisors were available to answer questions and provide guidance. A review meeting was held towards the end of the evening to provide a recap about certain aspects of the software and to allow questions.

In the third training session, the recruits continued to practice interviewing while the supervisors went around testing each person in turn. Once the training supervisor was satisfied that a trainee was ready to start live interviewing, that person would be moved to the main interview floor. Having the new interviewers come on to the floor one at a time enabled the team leaders and their monitoring staff to pay special attention to each person during the conduct of their first few live interviews. Enhancements to the software allowed new interviewers to be assigned only households that had not yet been contacted. This simplified their work and increased their confidence.

An additional 1-2 hours of training was provided after new employees had been interviewing for a week to review performance reports, the visual review procedure, give more detail on geocoding requirements and provide an opportunity to answer questions and clarify issues interviewers had encountered in their first few shifts. In previous surveys this additional training had occurred on an adhoc basis. Floor supervisors were always available to answer questions and respond to problems throughout regular interview shifts.

5.4 Rates of Pay

Interviewers were paid $10 per hour during training and $12 per hour as soon as they started to conduct live interviews. Rates of pay were reviewed every week with merit increases awarded on the basis of performance. Daily and weekly performance statistics were calculated for each interviewer taking into account 3 measures:

1. Productivity. Both the number of phone calls placed and the number of interviews completed per paid hour of interviewing time.
2. Trip Rate. The average number of trips recorded per person in the households for which interviews were completed.

3. Refusals. The proportion of households contacted where the respondent refused to participate in the survey.

Although performance statistics were used as the primary factor in setting rates of pay, other factors were also taken into consideration. These factors included the number of post interview callbacks required, the general accuracy of their work, their willingness and co-operation. Interviewers who were actively conducting surveys in languages other than English were given increases to compensate for the additional time required to translate the interview on-the-fly as well as the additional complexity these households often presented.

Saturday to Friday was chosen as the pay period permitting the performance reviews to take place on Saturday in time for the payroll to be processed over the weekend. The merit increases were applied to the pay period that justified them so that interviewers received immediate reward for good work and improvements in performance. Pay cheques were dated for the following Friday and were generally distributed during or after the Friday night shift. This provided a significant incentive for staff to attend the Friday or Saturday shift each week.

Staff were given a different rate of pay for non-interviewing time including supervisory duty and visual editing of interviews. The non-interview rates of pay were generally kept lower than the rate paid for interviewing in order to maintain the incentive for putting in as many hours as possible on the phone. The average rates of pay per hour, including incentive bonuses and vacation pay, are shown in the following table. The corresponding 2001 and 1996 TTS averages are also shown.

Table 5.2 Average Rates of Pay

	2006	2001	1996
Trainee	$10.00	$10.00	$9.00
Interviewer	$13.96	$13.23	$11.25
Team leader	$17.15	$16.63	$16.04
Coder	$14.03	$12.83	$11.15

5.5 Hours of Work

Standard evening interview shifts ran from 5:30 to 9:30 p.m. Some experimentation was done with weekday afternoon shifts, the results of which confirmed the rationale for starting at 5:30. The daytime success rate and productivity rate were both low for experienced interviewers during the afternoon period, although having cleared the calling queue, the evening shift did experience a significant improvement in performance. Taking the afternoon and evening shift together the total productivity for the day was not an improvement over a day with a standard evening shift.

Staff was instructed not to start any new interviews after 9:30 p.m. but were encouraged to complete any interviews in progress. They were credited with an extra 15 minutes of interview time if they had a live interview in progress at 9:35 p.m. This encouraged interviewers to dial right up to the 9:30 cut-off, maximizing potential completions for the day. Interviewers who did not want to 'risk' going overtime would opt to do their confirmation callbacks in the last few minutes of the shift instead. On Saturdays, the basic interview shift was from 10:00 a.m to 2:00

p.m. but staff were allowed to continue until 4:00 p.m. on some days if they so wished. On days with a shift until 4p.m. consideration was given to interviewers who preferred to work an alternate 4 hour period (ie. 11a.m. to 3p.m. or noon to 4p.m.). This may have increased the number of interviewers participating in Saturday shifts.

Starting in December a light week-day afternoon shift was brought in to re-contact households that had been positioned as 'answering machine, no message left' the previous evening. This reduced the amount of time evening interviewers spent leaving messages and provided an alternate time to attempt the previously unsucessfully reached household. If the afternoon shift again encountered an answering machine a detailed message was left and the household returned to the regular evening calling queue.

5.6 Incentive Bonuses

Initially a bonus of $2 was paid for each hour of interviewing in excess of 14 hours in one pay period. The purpose of the bonus was to encourage regular turn out thereby reducing the total number of interviewers that needed to be recruited.

Mid-way through the 2005 and 2006 surveys an incentive bonus was introduced to persuade interviewers to stay on until the end of the survey and encourage them to work extra shifts. Between the last week of November and the end of the survey an end of the survey bonus of $.60 per completed interview was paid to qualifying interviewers. In order to qualify they had to:

A) Remain on staff until the end of the survey.

B) Complete a minimum of 12 hours (3 shifts) of interviewing in every pay period during the incentive period.

C) Short falls in one pay period could be made up by working extra shifts in a subsequent pay period on a two for one basis (ie. 2 hours of extra interview time to compensate for each 1 hour missed).

During this period team leaders were awarded $0.03 for each qualifying interview by team members.

Additionally, in October of 2006 the bonus rate paid for each hour of interviewing in excess of 14 hours in one pay period was adapted to a sliding scale such that the better interviewers received a bigger bonus. The sliding scale was set equal to their base rate of pay minus $10 with a minimum of $2 and a maximum of $5 per hour. Supervisory and other non-interview time did not qualify for the bonus. The number of qualifying hours was reduced to 10 for short work weeks resulting from public holidays.

No bonuses were paid during the initial training period in either August, 2005 or in August, 2006.

5.6.1 Other Work Environment Incentives

Over the years various techniques have been used to encourage staff retention, promote increased shift scheduling, ensure quality work and increase job satisfaction. With the bulk of the staff being both temporary and part-time these initiatives are well-received and differentiate the TTS work environment from other similar work environments.

From the beginning each staff member is treated as an important individual within the organization. They are given their own set of tools (notepad, pen, headset pieces) in a permanent folder left in a designated area on-site. All management staff on-site address each interviewer by name. Coffee and filtered water are provided free of charge and a fridge and microwave are available on site. A break room with a phone for local calling is available. Feedback from interviewers is given due consideration and their preferences regarding work hours and station assignment are respected in so much as it is possible. Any problems they experienced are given quick attention. Strong workers with a good attitude are rewarded financially and are given the opportunity to move-up within the organization. Recognition is offered daily for work well-done and feedback provided on how to improve. Daily postings on a large white-board keep all staff current on our progress and provide a quick way to make any announcements.

Weekly team meetings build morale and provide an opportunity to congratulate individual and team successes. Occasional team-based contests encourage performance and provide a bit of fun. Every Saturday donuts are provided before the start of the shift which gets the day off to a prompt start. Every other month (or so) a whole staff event with pizza or cake provides an opportunity for management to re-cap progress to date and make any significant announcements, as well as providing a chance to socialize. Holiday Season and end-of-the survey parties congratulate the success of the Team and help build the foundation of staff that will want to return to future TTS projects. Another key element to building this foundation is the provision of personalized letters of reference to all deserving employees who finished the project. Taken as a whole, these elements have been found to build a real loyalty in a critical mass of interviewers.

5.7 Quality Control

Quality control of the information being collected was assured by the following procedures.

1. Logic checks performed by the DDE software.
2. Monitoring of interviews while in progress.
3. Daily monitoring of interview performance statistics.
4. Visual review of all completed interviews.
5. Callbacks.
6. Feedback from the coding process.
7. Rotation of sample between interviewers.
8. Random quality control audits.

5.7.1 Logic Checks

The DDE software controls the flow of the interview, preventing the interviewer from moving on until a valid response has been entered for each question. At the completion of an interview, the software performs a second series of checks on the consistency and completeness of the information. A list of errors and warning messages appears on the screen prompting the interviewer to go back and make corrections immediately while the respondent is still on the phone. Any errors that are not corrected will appear on the print out of the interview for visual review by a supervisor.

5.7.2 Monitoring

All interview stations were equipped for monitoring, both auditorily and visually, by a supervisor. Newly trained interviewers were monitored more frequently than seasoned interviewers. The team leaders and their most experienced staff carried out monitoring. Any comments were recorded in writing. Minor problems were brought to the attention of the interviewer immediately, particularly if corrections to a just completed interview were required. Serious problems were reported to the team leader for appropriate corrective action. Items of particular concern were the interviewers' telephone manner and their ability to question respondents to ensure completeness and accuracy of information. Interviewers were warned not to lead respondents in their answers, not to make assumptions, and were coached on methods to encourage potential refusals to become completes.

5.7.3 Performance Statistics

The sample control software produced data files that were read into Excel to print comprehensive statistics on interviews conducted by each interviewer, both daily and weekly. Team leaders and management staff could also display or print a historical record of any interviewer's weekly performance statistics. In addition to setting rates of pay, the performance reports served to identify other problems, such as below average trip rates and higher than average refusal rates, so that corrective measures could be taken. A sample report is shown in Table 5.3.

5.7.4 Visual Review

After each interview session, all of the completed interviews were printed out. The software used to print the interviews performed the same logic checks as the DDE software, flagging errors with appropriate messages. A supervisor visually reviewed every interview by looking at the error messages, the consistency and logic behind the information collected, and the manner in which descriptive information, such as trip destinations, was recorded. The printouts were sorted by interviewer within each team and the printing was done overnight so that the visual review could be completed before the next interview session. Problems and corrective actions were noted on the printouts.

A separate visual review was done for transit related errors by a staff person from the TTC. Most problems resulted from missing route descriptions in the look-up database or routes that did not connect. The sample control software was designed to prevent a household from being passed on for geocoding until a valid code had been assigned to every transit route used. Most problems were fixed by using the DDE software to amend the route description. In other cases, new route descriptions were added to the look-up database. Problems requiring callbacks were noted on the printout. The review of transit problems was generally done prior to printouts being reviewed by a supervisor.

Table 5.3 Typical Performance Printout

2-Oct																End Date -		19-Oct					Sorted by Performance Score										
						PERCENT DISTRIBUTION OF CALLS																												
ID	Logged hours	Completions	Persons	Trips	Calls	Call Backs	Non English	No Answer	ans. mach.	Line Busy	Intrpt. Interv.	Invalid	Out of Service	Refused	Succ. Compl.	Edits	Intersections	Persons/hhld	Calls/hr		Comp/hr		Trip Rate		Refusals		Intersections		Perf. Score	
223jb	33.7	157	418	820	885	20	1	29	24	2		2	1	2	18		67	2.7	26		4.7		2.0		11%		8%		8.2	
105ab	6.1	21	53	113	128	23	2	15	34	3		1	2	4	16		4	2.5	21		3.5		2.1		19%		4%		6.4	
129ab	22.8	63	177	369	507	18	2	12	47	4		2	2	1	12		26	2.8	22		2.8		2.1		6%		7%		6.1	
256cj	23.9	69	165	388	498	18	2	15	37	3		2	3	6	14	1	15	2.4	21		2.9		2.4		29%		4%		5.6	
206ma	18.0	43	119	315	308	19	4	17	35	3		3	3	3	14		25	2.8	17		2.4		2.6		19%		8%		5.3	
410vb	33.1	85	250	451	765	18	2	20	41	4		1	1	2	11		33	2.9	23		2.6		1.8		15%		7%		5.2	
397vd	49.7	121	345	784	867	23	4	9	38	3		3	2	2	14	2	48	2.9	17		2.4		2.3		14%		6%		5.2	
333mm	52.7	153	425	852	751	17	1	18	31	4		3	3	4	20		15	2.8	14		2.9		2.0		15%		2%		5.1	
312bb	26.7	76	232	471	541	21	3	13	37	4		2	2	4	14		102	3.1	20		2.8		2.0		21%		22%	*	5.0	
281sd	31.7	76	213	381	542	26	4	17	34	1		2	2	1	14	6	10	2.8	17		2.4		1.8	*	7%		3%		4.9	
261sr	23.1	60	173	347	515	19	2	22	35	2		2	3	5	12		42	2.9	22		2.6		2.0		30%		12%		4.7	
450mb	18.7	46	115	232	265	26	2	13	32	2		2	3	3	17	1	2	2.5	14		2.5		2.0		15%		1%		4.7	
139jm	24.5	53	164	341	560	16	8	27	29	3		2	2	4	9	1	16	3.1	23		2.2		2.1		27%		5%		4.6	
424aj	23.0	71	209	335	441	16	3	12	38	3		2	2	8	16		14	2.9	19		3.1		1.6	*	33%		4%		4.6	
400cb	34.2	72	184	375	508	21	2	22	32	0		3	3	3	14		42	2.6	15		2.1		2.0		19%		11%		4.0	*
339cs	31.4	67	203	479	590	23	3	15	35	2		1	3	6	11	2	78	3.0	19		2.1		2.4		36%		16%		4.0	*
383ah	27.2	72	204	352	552	14	5	18	31	3		2	5	8	13		53	2.8	20		2.6		1.7	*	38%		15%		3.9	*
286hc	29.0	59	186	351	461	22	2	14	40	2		2	1	3	13		49	3.2	16		2.0		1.9		18%		14%		3.8	*
377jr	26.4	59	169	277	633	22	2	13	44	2		1	2	5	9		32	2.9	24		2.2		1.6	*	37%		12%		3.8	*
431jt	48.1	87	251	581	813	21	3	12	41	4		1	2	6	11		16	2.9	17		1.8	*	2.3		36%		3%		3.7	*
Team A	1434	3441	10030	20330	24681	21	3	18	33	3		2	3	4	14	42	1792	2.9	17		2.4		2.0		23%		9%		4.4	
Team B	1361	3887	10921	22852	25018	20	2	17	33	3		2	3	4	16	30	1387	2.8	18		2.9		2.1		21%		6%		5.3	
Team C	928	3632	10371	21436	25605	20	2	16	34	3		2	3	5	14	15	2116	2.9	28		3.9		2.1		25%		10%		7.0	
Team D	1204	3500	9769	19676	24772	19	3	18	33	3		2	3	5	14	7	1254	2.8	21		2.9		2.0		27%		6%		5.2	
Others		915	2566	5767	4693	22	2	15	30	4		2	1	5	19	709	343	2.8					2.2		20%		6%			
Total	4926	15375	43657	90061	104769	20.1	2.4	17.1	33.3	3.2		2.0	2.7	4.6	14.7	803	6892	2.8	21		3.1		2.1		24%		8%		5.7	

5.7.5 Callbacks

Printouts requiring callbacks or clarification were given back to the respective interviewer before the next interview session began. Interviewers were notified, either by the notes on the printout or verbally by the team leader, of areas where improvements to their work could be made. The interviewers were required to make the callbacks during the course of the current shift, and to continue their attempts until the issues had been resolved. Corrected information was written on the printouts, which were then given back to a supervisor. Supervisory staff then made the corrections to the database using the DDE software.

If the original interviewer was not available to work the next session, the printouts were held until the following day. If the callbacks had not been made within two days then a supervisor would arrange for the call back to be made by an alternate interviewer. In some cases, callbacks were made by supervisors which provided an opportunity to check on the quality of the interviewer's work by speaking directly with the interview respondent.

5.7.6 Feedback from the Coding Process

Once all the visual reviews, callbacks and corrections had been made for a given interview date, the data for those households was moved to the coding database for geocoding. A series of computerised logic checks was performed on each household to ensure that the information being passed on was complete. Incomplete interviews, and those containing identifiable errors such as missing transit route codes, were kept in the review database and reprinted for further checking.

If the geographic information in the coding database proved to be insufficient or ambiguous, the coders had the option to flag the record for a new printout to be generated. The following day these printouts were returned to the interview teams for geocoding callbacks. Once callbacks were completed and the information clarified, the corrected printouts were given back to the geocoders for entry into the geocoding database.

Problems encountered in the geocoding process were monitored continuously and reported to the team leaders so that corrective action could be taken with respect to future interviews. The survey procedures were set up with the expectation that the geocoding would take place within 3 days of the interview. For the most part coding was able to keep up with the information being passed to it but there were delays in the review and edit process which sometimes resulted in a time delay much greater than 3 days.

5.7.7 Rotation of Sample Between Interviewers

In previous surveys, once a particular household was assigned to a computer workstation, all future contact with that household had to be from that station. By rotating interviewers at a particular workstation it was possible to observe problems in the way that a given interviewer had previously recorded information and how households had been dispositioned. Of particular concern was an interviewer scheduling callbacks for households instead of accepting refusals.

Improvements to the sample control software in 2006 specified 'ownership' of a household by interviewer ID, not by workstation. Once a household interview was initiated the same

interviewer followed up with that household until it was completed, unless that interviewer was unavailable within the acceptable window of time that a repeat contact was scheduled. This allowed an increased efficiency in having the same interviewer complete all contact with a household with which they were familiar, but removed the check and balance of the previous workstation rotation framework.

Releasing sample into the general team queue could be forced by setting any given interviewer to 'fresh only' mode, whereby he/she only received previously uncontacted households. Interviewers were still instructed to report to their supervisors any problems in the way that previously collected information, or call disposition, had been recorded, however the new protocols greatly reduced the instance of this as no single interviewer was ever forced to wade through a collection of work from another single interviewer in the same way. The ability to assign one interviewer's pending work to another, single, interviewer would replicate the check and balance that was previously available in the work-station dependant model, and might be considered in future TTS.

Some of the increase in refusal rate observed in 2005 and 2006 might be attributed to interviewers more readily accepting refusals given the near certainty the household would return to them for callback.

5.7.8 Random quality control audits

Upper level management conducted adhoc quality control audits at several levels during the interview process:

- Adhoc real-time monitoring of interviewers including callbacks for additional information.
- Periodic review of team monitoring sheets to assess consistency of monitoring overall, ensure monitoring of each interviewer on a regular basis and identify re-current issues.
- Assessment of visual reviews for each team and for each reviewer to assess quality of work produced by each team and ensure completeness and correctness of comments provided by reviewing staff.
- Occasional supervisor callbacks to confirm and/or supplement data originally collected.
- Occasional confirmation of completeness of information entered by supervisors following requests for interviewers to gather additional information on paper.
- Duplicate assignment of adhoc households to multiple geocoders to check for consistent coding methods.

5.7.9 Paper Management

The amount of paper generated in the processing and validating of households through the various stages of the survey is not insignificant. Great care is taken in tracking and organizing this paper, both as a means of being able to step-back through additional information and edits made to individual records, and for the purpose of maintaining the confidentiality of our respondents. Every page of every printout is collected, changes entered into the database and then re-sorted by team and interview date. Only when all the pages have been accounted for and relevant changes made are the households for any interview date passed to the next stage

of the process. At the completion of the data collection portion of the survey, all of the printouts are shredded.

5.8 Answering Machines (Voice mail)

The terms answering machines and voice mail in this section, and elsewhere in this report, are used inter-changeably and refer to either answering machines or voice mail.

By the end of October, 2006 statistics showed a 6% drop in the likelihood of getting a completion in the course of regular evening interviewing, compared to 2001. 40% more calls needed to be placed to get the same number of completions as in 2001 and almost twice as many as in 1996.

By the end of November, 2006 statistics showed the mean probability of encountering voice mail for all calls to be 45%. The probability was lowest (28%) between 7:30 and 8:30 p.m. and highest (61%) in the early afternoon. The number of previous calls made to the same number did not significantly change the probability.

The procedure for handling answering machines was modified to address the following concerns:

- Interviewer productivity (long messages left in the evening seriously impact productivity).
- Quantity of messages left at a household (viewed by some to be an unwarranted invasion of privacy).
- Quantity of previously contacted households returning on any given shift (a two or four day cycle on returning calls spreads the load most uniformly).
- Content of the messages (allowing additional time for leaving detailed first messages was critical in cases where an advance notice letter may not have been received such as in apartment buildings).
- Maximizing the probability of making live contact on one of the first two calls.

CALL 1 – no message left, callback scheduled for next available week-day between 7:30 and 8:30 p.m. or during a Saturday shift if CALL 1 occurred on a Friday.
CALL 2 (assuming CALL 1 encountered an answering machine) – no message left, callback scheduled for next week-day at 2pm.
CALL 3 (assuming CALL 1 & 2 encountered an answering machine) – leave a detailed message with similar content to the advance letter. Advise the recipient that an interviewer will call that evening or the next day. Leave a phone number that the recipient can call to do the survey at their convenience.
CALL 4 – same as CALL 1
CALL 5 (final attempt) – message left stressing importance of recipient's participation in the survey with a request to call in to complete the survey.

A household was removed from the active calling queue under the following circumstances:

1. After the 9th call in 2005 and after the 8th call in 2006
2. After 5 consecutive no answers
3. After 4 voicemail if the interview had not begun (no persons entered), CALL 5 above (call 2 was dispositioned as a callback, not a voicemail).

These households were still available for completion if the household called in to complete the survey. Any household that reached this state and had any trip information was printed for review by a supervisor who could decide to further pursue gathering the data to make it complete.

5.8.1 Call-in From Voice Mail

In previous surveys, when a household called in it was necessary to take their phone number and have an interviewer call them from the particular workstation that contained their sample information. Improvements to the software for the 2006 TTS allowed respondents calling in to be interviewed immediately. Most of these calls were in response to the answering machine message. The call-in phones were staffed from 10 a.m. to 9:30 p.m. each day and from 10 a.m. to 2 p.m. on Saturdays. At other times a voice mail message was provided asking the respondent to either call back between those hours or, if the call was in response to a request for a specific piece of information, to leave that information on the voice mail. In 2005, a total of 8 bounce lines were used at 8 regular interviewing stations. In 2006, 3 bounce lines were used in a dedicated call-in room. In both cases a supervisor carried a cordless telephone for the last bounce line ensuring someone was always available to answer an incoming call during regular interviewing hours.

While this improvement streamlined the process from the perspective of the respondent, the interviewers lost the additional incentive to leave proficient messages in the hopes that the complete would come back to them by way of a respondent returning their personalized message. In future surveys returning call-ins to the interviewer who last made contact with the household in cases where the interviewer is present and the household is willing to wait to be called back would be advisable to increase morale on the floor and enthusiasm for leaving effective messages. Another option would be to track who left the last message and offer recognition through an increase in performance score (and possible resultant pay increase) or a fixed bonus amount.

5.9 Survey Interruptions

The only system-wide disruption to normal interviewing was a result of Election Day in Canada on January 23rd, 2006 when somewhat abnormal trip behaviour by households in the survey area might be expected, hence, no trip data was collected for that day. Instead, on Tuesday, January 24th 2006, trips were collected for Friday, January 20th instead of for Monday, January 23rd.

Aside from this disruption there were a few specific localized problems in various regions which necessitated turning off the interview sample for those areas at different times in the survey.

- In the 2005 phase of the TTS interview sample for Orangeville & Dufferin County were turned off on February 6th and 7th of 2006 due to severe weather in that region.
- In 2006, there was a long running transit strike by Durham region transit authority which necessitated turning off all Durham sample for a full month, from October 6th to November 7th.
- Sample for the Scarborough area of the City of Toronto was also turned off from October 27th to 29th due to a major power outage in the area.

5.10 Non-English Callbacks

The Direct Data Entry (DDE) software allowed the interviewers to schedule a call back to be made in a language other than English. The languages that could be specified were selected based on the frequency with which they were used in the 1986, 1991 and 1996 surveys. Those languages (and the total interviews in the 2006 TTS) were Cantonese (1703), Mandarin (678), Italian (892), Portuguese (515), Spanish (224), Greek (167) and French (85). The category "Other" could be selected for other languages or if the appropriate language could not be identified. Interviewers were instructed to specify the 'other' language, where known, in the comments. In the last half of the survey period a report was generated sorted by the language specified in the comments. Where possible, this report was distributed to interviewers proficient in the relevant language and in many cases the interview could be completed in the respondent's language of choice. The interviewers conducting non-English interviews did their own translation from the Standard English script. Households in the other category, where the required language was not identified or not spoken by one of our interviewers, were contacted by an experienced interviewer who would attempt to conduct the survey in English, in most cases with another member of the household from the one which was originally contacted. There was limited monitoring of non-English interviews.

A total of 2198 interviews were completed in 'other' languages including: Arabic, Bengali, Bosnian, German, Gujarati, Hindi, Hungarian, Lituanian, Polish, Punjabi, Romanian, Russian, Serbo-croatian, Somali, Tamil, Ukranian and Urdu.

Households coded as non-English were available from any work-station within the team from which it was initially contacted, or from any work-station operating in 'call-in' mode. No special efforts were made to recruit a sufficient number of interviewers with non-English language skills, although early attempts were made to identify and encourage other language skills. In the 2006 survey we would have benefited from another interviewer proficient in Italian (we had only 1). We were fortunate to have a proficient Portuguese interviewer return from 2001 as well as a new Portuguese-speaking interviewer. One interviewer in Greek was sufficient and we had 7 interviewers able to do Spanish. With 5 Mandarin speaking, 5 Cantonese speaking and 6 French speaking interviewers we were able to stay on top of those households from the beginning of the survey period.

Section 6 Completion Statistics

Table 6.1 shows the number of completed interviews in the final database for the areas represented by each of the agencies. The table also includes dwelling unit and population counts from the 2006 Canada Census. The Census dwelling unit counts include seasonal residences and vacant buildings and are therefore not directly comparable with the TTS data. The mean expansion factors shown are those used in the final database for expansion of the survey data to represent the universe of households in the survey area. The expansion factors have been calculated by postal areas, which do not necessarily match municipal boundaries hence, neither the expanded dwelling unit nor household totals match the census data exactly. The expanded survey population is generally slightly less than the census number due to the exclusion of nursing homes, hospitals, prisons and other collective homes from the survey. The 5% sample target was exceeded in the County of Dufferin where the forward sortation area 'L0N', was inadvertently included in the survey areas for both phase 1 and 2.

Table 6.1 Completed Interviews by Agency

	2006 Census		TTS Records		Expanded Totals		Mean	Mean
	Dwelling Units	Population	Dwelling Unit	Person	Dwelling Unit	Person	Expansion Factor	Sample Rate
City of Toronto	979323	2503017	51613	129276	979319	2445990	18.97	5.3%
Region of Durham	194639	561186	9527	26404	194641	535841	20.43	4.9%
Region of York	275673	892712	14221	44046	275677	852826	19.39	5.2%
Region of Peel	360990	1165810	17946	56050	360990	1096137	20.12	5.0%
Region of Halton	156931	439204	9207	24736	156930	422672	17.04	5.9%
City of Hamilton	194475	504559	9972	25184	194479	487012	19.50	5.1%
GTHA	2162031	6066488	112486	305696	2162036	5840479	19.22	5.2%
Region of Niagara	169425	427421	8884	22052	169425	415302	19.07	5.2%
Region of Waterloo	177996	478121	8731	23525	177992	468159	20.39	4.9%
City of Guelph	44706	114943	2331	5968	44707	112954	19.18	5.2%
Wellington County	19761	55952	1076	2876	19761	53097	18.37	5.4%
Town of Orangeville	9429	26925	549	1525	9429	26018	17.17	5.8%
City of Barrie	46533	128430	2076	5550	46537	124567	22.42	4.5%
Simcoe County	97903	263478	5388	14280	97900	260512	18.17	5.5%
City of Kawartha Lakes	29509	74561	1707	4092	29514	70751	17.29	5.8%
City of Peterborough	17094	45919	920	2327	17095	43143	18.58	5.4%
Peterborough County	31204	74898	1733	4059	31204	72642	18.01	5.6%
City of Orillia	12238	30259	631	1482	12235	28736	19.39	5.2%
County of Dufferin	9368	27511	762	2147	9370	26457	12.30	8.1%
City of Brantford	35609	90192	1739	4353	35608	87890	20.48	4.9%
Brant County	12238	34415	618	1721	12236	34076	19.80	5.0%
Total exc. GTHA	713013	1873025	37145	95957	713011.3	1824304	19.20	5.2%
Total survey area	2875044	7939513	149631	401653	2875047	7664783	19.21	5.2%

Preliminary comparisons made between the 1996 TTS and Canada Census data suggested that the survey underrepresented people in the age range of 18 to 22 years by 8%. The same age

group was underrepresented by about 11% in the 2001 TTS. In the 2006 survey, the age group of 18 to 27 was underrepresented, based on comparison to Canada Census data, by an average of 20%, and the age range of 28 to 37 were underrepresented by an average 10%. The reason for the increasing underrepresentation is not known. Possible explanations include:

1. The increasingly widespread use of cell phones. Most cell phone numbers are not listed and are therefore excluded in the sample selection. Their exclusion is not a problem for those cell phones which are used in addition to a household's regular land line. But if they are used as a substitute for land lines it could result in an underrepresentation of some segments of the population in the survey results.
2. It is not known to what extent the phone listings from which the sample was drawn are completely up to date with respect to students moving into new homes or residences at the start of the school year.
3. People who are frequently out in the evenings are harder to contact and are therefore less likely to be surveyed than those who remain at home.

Unlike previous surveys, in 2006 there is an overrepresentation of people in the age group of 58 to 87 with the highest overrepresentation between the ages of 68 and 77. The response rate is generally better for people from this age group and they are more likely to have a listed residential phone line to be included in the sample time frame. This is another proof of the effect of the exclusive use of cell phones.

The under or overrepresentation of one age group creates the potential for bias in the survey results to the extent that the travel patterns and behaviour of that age group differ from that population as a whole.

Table 6.2 gives a summary of the combined completion statistics for all 3 components of the 2006 TTS. The numbers shown for the 1996 TTS are not exactly comparable because of the change in procedure with respect to answering machines. Starting in 2001 the inclusion of most answering machines in the "sample used" sub-total is done to give a better measure of contact and completion rates but leads to an overstatement of the difference in those rates relative to the 1996 rates.

Table 6.2 Completion Statistics

Total sample		351,828				
Not attempted		9,389				
Incomplete		8,500				
Sample used		333,939				
Out of service		26,487				
Invalid		23,046				
Rej. No Answer		9,556				
Rej. Uncontactable		18,978				
Rej. Voicemail		48,790	2006	2001	1996	
Valid Contacts		207,082	62.0%	81.2%	88.0%	of sample used
Refusals		54,314	26.2%	21.1%	21.8%	of valid contacts
Completed interviews		152,768	45.7%	64.1%	68.9%	of sample used
Rejected in review*		3,137	844	857	849	
			2006	2001	1996	
	Households	149,631				
	Persons	401,653	2.68	2.74	2.72	
	Trips	864,348	2.15	2.19	2.13	
	Transit records	87,244				

*includes 2293 households done in training and subsequently discarded

Refusal rate, calculated as:

Households who refuse/(households who refuse+households who complete)

increased to 26.6% in 2006 from 21.1% in 2001 and 21.8% in 1996.

Additionally, the increase in households rejected after multiple unsuccessful attempts to contact them produced a considerable reduction in the completion rate from 70% in 1996, 64% in 2001 to only 44% in 2006.

The low response rate from multiple unit dwellings appears to be the primary reason for the lower completion rate in Toronto relative to the other areas. Where the average completion rate outside Toronto in 2006 was 48%, the average within Toronto was only 41%. In 2001 the completion rate between Toronto and outside Toronto varied 4%, from 62% to 66%.

Of the 222 forward sortation areas included in the survey, 24 of the 30 areas with the lowest completion rates were within Toronto, 3 bordered the airport in Mississauga (L4V, L5T, L5S), 1 was in central Hamilton (L8N), and 2 in south-central Burlington (L7S, L7R). Figure 6.1 illustrates the completion rate by FSA within Toronto.

Figure 6.1 Completion Rates for Toronto Postal Areas

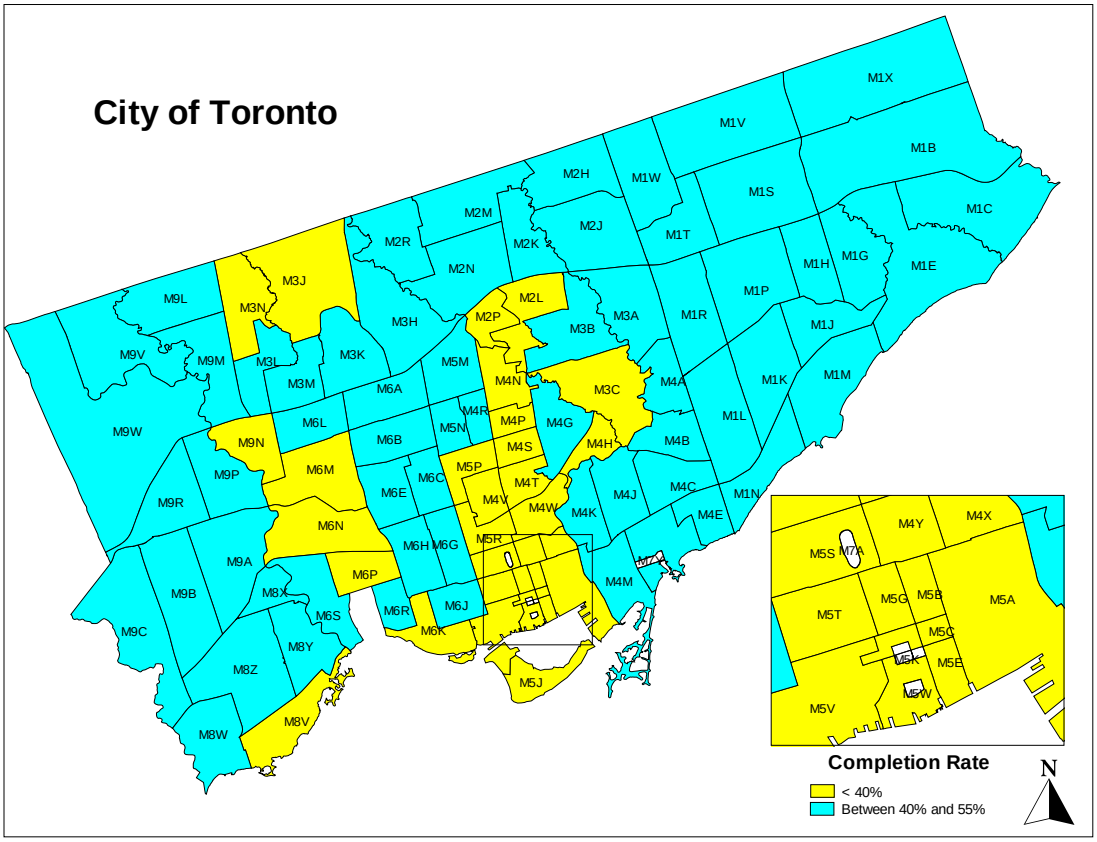


Table 6.3 shows the outcome of all of the phone calls that were made during each of 2006, 2001 and 1996. The most significant trend is in the number of calls that resulted in no answer or contact with an answering machine. The combined total of these categories increased from 42% of the calls placed in 1996 to 49% in 2001 and then to 52% in 2006. A substantial increase was also noted in 2006 for line busy (3%, up from 1% in both 2001 and 1996) and out of service (2%, up from 1% in both 2001 and 1996). The increase in line busy may be attributed to the use of the phone line for extended periods for internet services. The increase in out of service may be indicative of a less up-to-date sample source than previous years. The number of callbacks, both English and non-English has levelled out at 20%. The net result is that the average number of calls that had to be placed to obtain each completed interview in 2006 was 47% more than in 2001 and 87% more than in 1996.

Table 6.3 Disposition of Phone Calls

Phone calls	2006 TTS		2001 TTS		1996 TTS	
Out of service	25171	2%	5,543	1%	4,527	1%
Invalid	22261	2%	8,877	2%	9,279	2%
Line Busy	33094	3%	7,080	1%	5,487	1%
No answer *	172749	17%	128,529	27%	80,271	20%
Ans. Machine						
message left	228545	22%	104,025	22%	90,315	22%
no message left	135051	13%	*		n/a	
Call back						
English	184202	18%	89,680	19%	68,270	17%
Non-English	22871	2%	10,716	2%	6,742	2%
Interrupted			184	0%	464	0%
Refused	51024	5%	25,231	5%	31,260	8%
Complete	147154	14%	101,568	21%	109,204	27%
Total	1022122		481,433		405,819	
Calls per completion	6.95		4.74		3.71	

- *The 2001 No answer count includes an estimated 50,000 to 65,000 answering machines that were recorded as no answer and are not included in the answering machine count. The 2001 totals are based on the fall 2001 component only.
- The 2006 totals are based on the main survey periods of Sep 12 '05 to Feb 9 '06 and Sep 6 '06 to Jan 24 '07

Table 6.4 shows the number of completed interviews by trip day of the week. Trip data for Fridays were collected on both Saturday and Monday except on the occasional Saturday when Thursday trip data were collected to limit the overrepresentation of Friday trips. Trips for Mondays were slightly underrepresented due to public holidays and the starting of the survey on a Wednesday. The uneven distribution of completed interviews by day of week results in an overall trip rate that is slightly higher than if all 5-week days were weighted equally.

Table 6.4 Completed Interviews by Trip Day

Trip Day		**Trip rate**
Monday	17.42% (17.3%)	2.10 (2.15)
Tuesday	19.35% (18.4%)	2.13 (2.15)
Wednesday	19.18% (19.4%)	2.13 (2.17)
Thursday	21.10% (19.6%)	2.14 (2.17)
Friday	22.94% (25.2%)	2.22 (2.26)

(2001 rates as published in 2001 Design and Conduct of the Survey Report are displayed in brackets)

Figure 6.2 shows the number of interviews completed by day and compares it with the corresponding day in the 2001 survey. Figure 6.3 shows the completed interviews per logged hour in 2006 compared to 2001. Interviewer productivity in 2006, at 2.83 completed interviews per paid hour of interview staff time was significantly lower than the 3.42 interviews per paid hour in 2001 and markedly lower than the rate of 3.7 achieved in 1996. The difference can be attributed to the escalation in the average number of calls required to achieve each completed interview.

Figure 6.2 Completed Interviews by Day

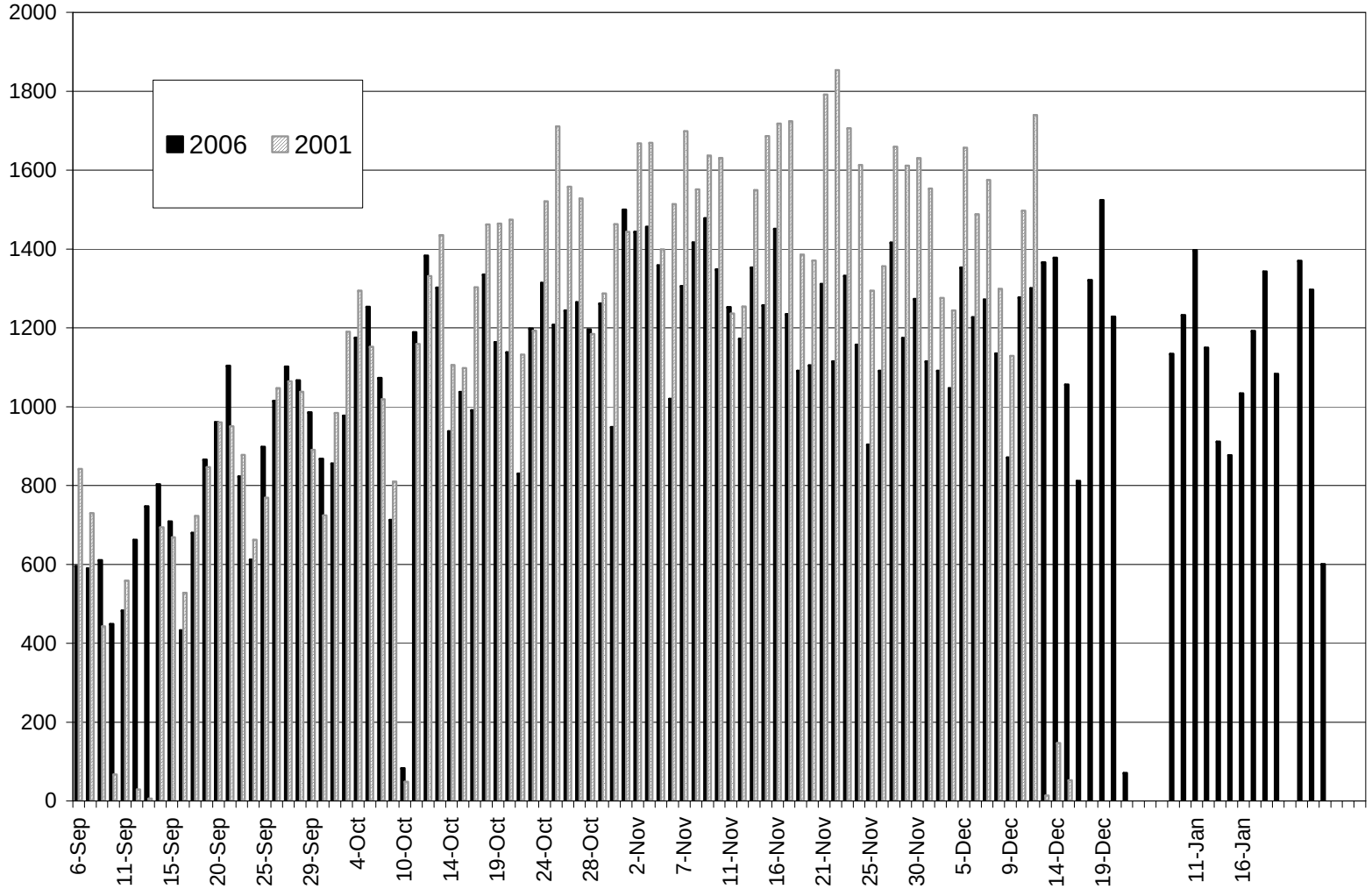
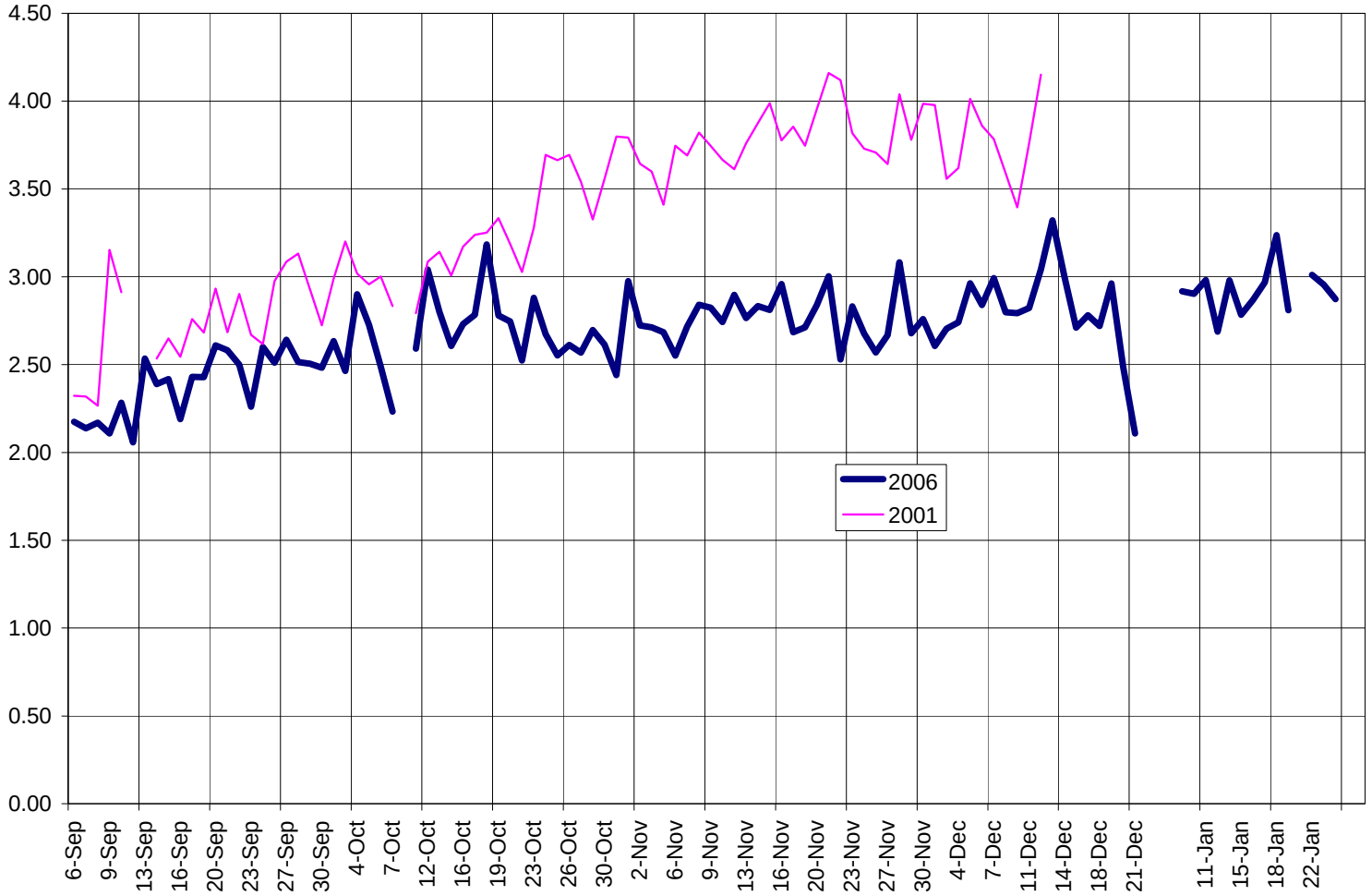


Figure 6.3 Completed Interviews per Logged Hour



Section 7 Coding

7.1 Staffing and Training

In 2005, recruiting of geocoders did not start until after interviewing was well underway as the geocoding software was not yet completed. In 2006, recruiting started about three weeks prior to the actual start of the survey. Coding positions were advertised through the University of Toronto and Ryerson University's employment placement centers with emphasis on computer and geography knowledge for applicants. Approximately 30 applicants were interviewed for the two surveying periods by the coding supervisor. 16 were retained. Nearly all of the coding staff had a University education with the majority coming from Geography, Engineering and Computing backgrounds. In addition, two interviewers joined the geocoding team during the survey.

Training for the coders took 2 days, with a formal ½ day session at the beginning where coders were introduced to the project and what was required of them. The coders were introduced to the geocoding console program and trained to use reference material such as telephone books, internet search engines and paper maps. This was followed by more training where the coders were allowed access to the geocoding console and worked on interviews collected during the interview training period. During this period the trainee coders were supervised by one or more of the senior coders.

Some of the coders were also trained to perform visual review and edit corrections in the early stage of the 2006 survey in order to reduce the load of the non-interviewing component on the day-time interviewing staff. Since coding was the last part of the survey process, extra effort was placed in stressing accuracy of information. The pay rate for coders started at $12 per hour and was increased to $13 per hour in November 2006 for most coders, with two senior coders making $14 per hour. The highest paid coders assisted in setting up the geocode reference database and some administrative and site computer work.

The coding staff was hired in stages throughout October and November of 2005 and September and October of 2006. While turnover was not great, those coders who did leave were not generally replaced on the floor as the staggering of the initial hirings allowed us to increase output from the remaining staff and maintain a relatively stable level of productivity without increasing staff. On average over the last two months of the survey a compliment of 8-10 coders was available daily.

7.2 Coding Activity

7.2.1 Coding in 2005

The coding plan was to geocode survey records within three days of the interview. The shorter the turn around time the better it would be for callbacks if households had to be contacted again to clarify information.

In the 2005 segment of the survey however it was not possible to fully geocode the interviews within three days of their completion. The Geocode console software was not yet completed and

hence coding took a different course. During the interview phase, instead of using the Geocode console during the actual interview phase, coders worked on hardcopies of the previous night's interviews. Using a combination of GIS software, paper maps and internet search engines to check locations on the interview sheets, an emphasis was placed on ensuring that home, work, school and other trip locations were codable. If they were not, the households were sent for callback. This paper coding process was done before the interviews were passed on to the daytime Review Edit staff and hence was done in a limited time frame usually until 1-2 p.m. the day after the interview although as more interviews were completed daily the time frame to complete took longer.

When the geocode console was ready in February 2006, the completed interviews were then completely coded using this software. The completed interviews were divided into 58 geozone areas and coders were assigned to specific zones allowing them to develop better knowledge of their section of the survey area. No significant problems were encountered using this method of coding as most of the issues had been addressed in the initial paper coding process.

Overall 5 coders were hired for this portion of the survey, three of whom remained on staff after the interview phase was completed. These three coders plus two interviewers who joined the group after the interview phase were responsible for the final coding using the geocoding console.

In 2005 interviews were only conducted on households external to the GTHA but initial geocoding reference databases were created at that time for the entire survey area and updated frequently as both the interview and geocoding processes continued. Updates were performed by the geocoding supervisor.

7.2.2 Coding in 2006

In 2006, the geocoding staff started to code a few days after the interviews were completed. The goal again was to geocode within three days of the interview. However, the review and edit stage of the interview process at times took longer than anticipated because there was a large volume of work to process and many of the interviewers did not work consecutive days.

Completed interviews were assigned to one of the 58 geozones and coders were assigned households to work on based on the geozones, number of households available for coding in the specific geozone and the update status of the available interviews (household just passed in from reviewing first, then the oldest in the backlog). This was done to allow newer data to be processed ahead of the existing backlog to facilitate geocoding callbacks happening as early as possible. Given this structure and the fact that some geozones had more households than others, some balancing of coding resources was necessary to ensure the stategy was adhered to.

In 2006, the survey focused on the GTHA area. Geocoding was done between 9a.m. and 5p.m. daily. This allowed sharing of the machines between geocoding and interviewing staff. Interview completions did not reach a maximum until sometime in mid-October. At this time coding was required for a large volume of households on a daily basis. For the most part the coding staff was

able to keep up with the required schedule without too much difficulty. A few coders worked part time. Part-time coders were assigned to geozones where a full-time coder was also working.

Updates were done on an almost daily basis to the monument and school reference database files. A procedure was put in place where coders made a list of new monuments and schools which needed to be added and these were then passed on to one of the senior coders who double checked the information before adding the locations that were necessary to the database.

At the end of the interviewing portion of the survey, three coders stayed on to assist with additional coding and other post-processing clean-up and validation work.

7.3 Post-Processing

Once geocoding was completed, the households in the TTS database were passed to a final post-processing phase. In this phase, checks were performed to search for miscoded locations, uncodable locations were removed and extensive logic checks were performed on the locations and information contained in the database to make sure that the data was correct. This process was used to identify any errors that may have gotten past the previous stages of data processing.

The first step in this process was a batch process run on all completed households in the database to check for errors in logic or geocoding errors. If potential errors were found in the household they were flagged and the households sent for manually checking. If no errors were found the data was placed into a final state in the database.

Some of the logic checks performed on the data during this batch process include (but weren't limited to) checks for:

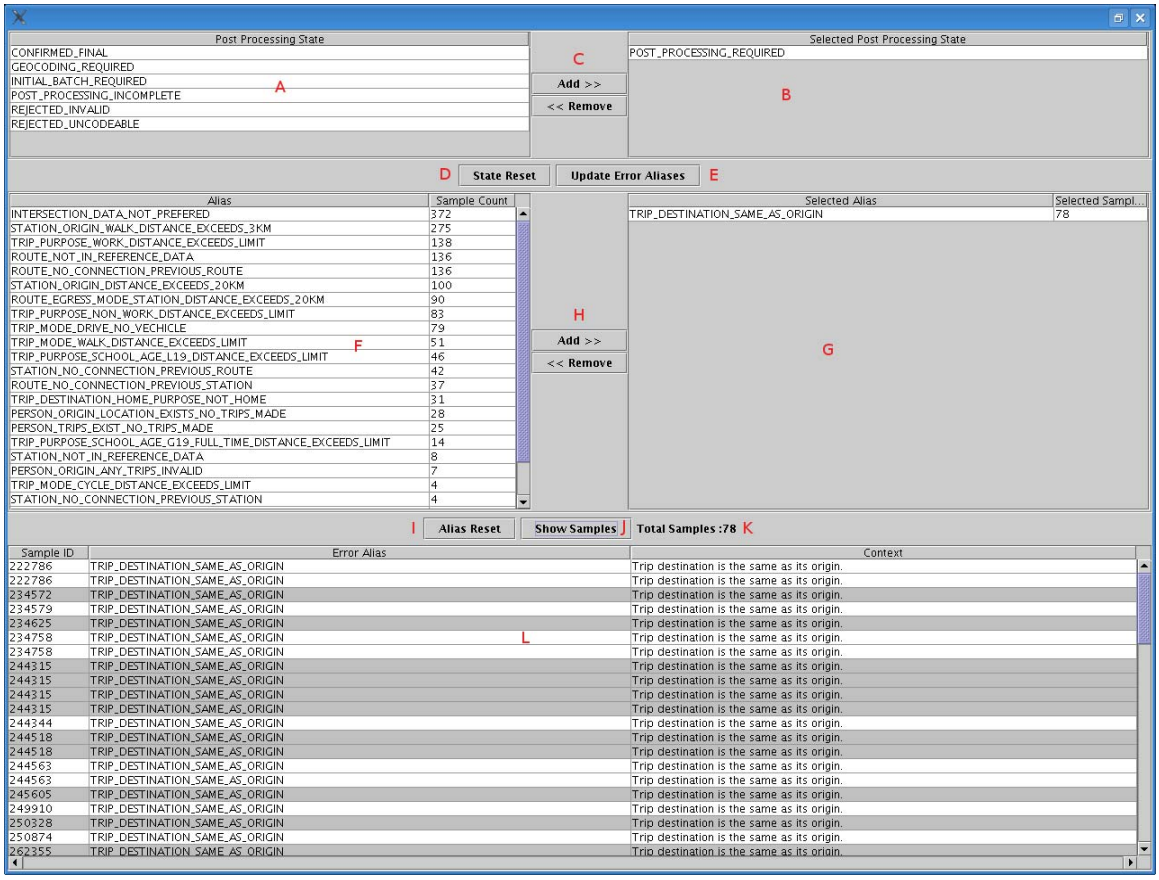
- Walk or cycle trip distances which were longer than thought to be valid
- Trip speeds which were excessive
- Lengthy access or egress distances from transit transfer points
- Extremely long school and work trip distances
- Transit routes not connecting

This process produced a list of potential errors to be manually reviewed and recoded as necessary. Figure 7.1 illustrates how post-processors used the DDE to identify households to work on:

A Post-processing states available to search.
B Selected post-processing states. Only samples in these states will be shown in the sample summary table (L).
F List of error aliases and their frequency (count of sample occurrences) that exist in the selected post-processing states (B). If the error alias is moved into the selected error alias list (G) then it will not appear in this list.
G Selected error aliases that samples will be displayed for in the sample summary table (L).
K Shows the total number of samples that match the assigned post-processing state list (B) and the assigned error alias list (G).

L Lists the sample summary information for the samples that meet the requirements of the selected post-processing state (B) and error alias data (G). An alternating colouring pattern is used to differentiate between different household samples. Selecting a row will provide the option to review the history of the sample and to check-out a specific version of the household.

Figure 7.1 Post-Processing DDE Screen



7.4 Statistics

A location was geocoded by one of three methods:

1. Cross referenced to another location field (i.e., trips to home, usual place of work or usual place of school)
2. batch processing or
3. interactive geocoding.

Table 7.1 is a breakdown of coding method (i.e. address type) for different surveyed information (i.e. location type).

In 2006, unlike 2001, no records were coded to Traffic Zone. Overall less than one percent of the records were coded to the more general Internal and External Place Name address types and

75% of the records were coded to Street Address which is the type of accuracy which was strived for. This was a significant increase from the 65% recorded in 2001.

Table 7.1 Location Types verses Address Types

Location Type	Street Address	Intersection	Monument	Internal Place Name	School	External Place Name	Total
Home	144,889	664	4,051	26	1	-	149,631
	97%	0%	3%	0%	0%		
Work	134,872	36,934	8,267	230	8,080	1,049	189,432
	71%	19%	4%	0%	4%	1%	
School	685	2	29	3	65,167	-	65,886
	1%	0%	0%	0%	99%		
Destination	610,242	111,420	43,668	1,287	94,107	3,624	864,348
	71%	13%	5%	0%	11%	0%	
1st Origin	266,202	2,409	9,668	115	113	922	279,429
	95%	1%	3%	0%	0%	0%	
Total	1,115,890	151,429	65,683	1,661	167,468	5,595	1,548,726
	75%	10%	4%	0%	11%	0%	

Section 8 Survey Budget and Costs

The total budget for the survey was $3.09 million including: software development, conduct of the survey, preparation of the final database, production of a series of Working Papers and production of the following three Reports:

Conduct of the Survey
Data Guide
Validation

The Data Management Group undertook the preparation of:

An Information Bulletin
2006, 2001 and 1996 Summary of results for the entire survey area
2006, 2001, 1996 and 1986 Summary of results for the GTHA

The original budget estimate for all aspects of the survey up to the presentation of results was:

$2.00 million for the areas within the Greater Toronto and Hamilton Area (GTHA),
$0.81 million for the areas outside the GTHA but within the survey area.

The cost sharing agreement in the GTHA was for GO Transit to cover 3% of the GTHA budget, and of the remainder the Ministry of Transportation would cover 75% with the remaining 25% covered by the Regions in proportion to their 2001 population. Outside the GTHA the participants were to be charged on the basis of the number of successful completions with the Ministry of Transportation covering 75% of that cost. In addition, all participants were to be invoiced in three equal billings in 2005, 2006 and 2007. Based on billing from an approved budget rather than on actual expenses, the Steering Committee responsible for the 2001 TTS agreed to allocate a carry forward of $51,000 for software development in preparation for a survey in 2006.

The survey management team realized during the interviewing phase of 2006 that the survey could not be completed by the end of December 2006. The options presented to the GTHA funding partners were to accept a smaller sample than 5% or provide the additional funds necessary to continue the interview phase into January and February 2007. The funding partners in the GTHA agreed to increase the budget by $250,000 with the same cost sharing arrangement. As a result of all the above, the final budget/expenses for the complete survey was:

$2.25 million for the areas within the Greater Toronto and Hamilton Area (GTHA),
$0.79 million for the areas outside the GTHA but within the survey area.
$0.05 million carry forward from the 2001 TTS

The marginal cost of completing a household interview increased from $12.37 for the 2001 survey to $15.80 for the 2006 survey. A 30% increase in telephone interviewing is attributed in a large part to the increased incidence of ‘call screening’ (where a household chooses not to answer the telephone after reviewing a call display).

8.1 University Overhead and Taxes

The overhead charged by the University of Toronto was 40% of University staff staffing costs and 2% of other expenditures. These overhead charges helped cover the cost of providing the Data Management Group office facilities, general supplies and secretarial services. University staffing costs includes the fees charged by the Project Manager but excludes the interviewers, coders

and supervisors hired specifically for the survey. The survey qualified as a University research project. Most equipment purchases were therefore exempt from Provincial Sales Tax. The University also qualifies for a refund of 2/3 of the net amount paid in Federal Goods and Services Tax (GST). University staff costs, excluding the Project and Site Managers, were exempt from GST. The appropriate amount of University overhead and net taxes has been included in the individual itemised costs in the following sections.

8.2 Cost Summary and Comparison with Previous Surveys

Table 8.1 provides a summary of expenditures incurred in the conduct of the 2006 Survey together, for comparison, with the same information for the 1996 and 2001 Surveys. The costs incurred for interview and coding staff in the 2006 Survey are the net of payroll expenditures including fringe benefits and payroll taxes. The staff were hired and paid by Peter Dalton Consulting, who invoiced the Data Management Group for the net amount of the payroll cost plus 4% to cover the cost of administration and interim financing.

8.2.1 Software Development

The computer software used to support the activities of the 1996 and 2001 Surveys was developed in 1990 with updates and improvements for subsequent surveys. The computer language used by the software was no longer supported and the procedures were in need of improvement. With the approval of the Steering Committee, the Data Management Group undertook the task of coordinating the development of a new suite of computer programs. The task began in 2003 with funds remaining from an approved total budget for the 2001 TTS. The expenditure item for software development in the 2006 Survey does not include the $100,000 incurred prior to the beginning of the 2006 Survey in 2005.

8.2.2 Interview Staff and Training

The productivity of interviewing staff, in terms of the number of completed interviews per interviewer hour, never reached the level of previous surveys. The most likely cause was a significant increase in ‘call screening’ where a household does not answer the telephone based on the information contained in a call display. Because the household is contacted at least 8 times, with voice messages left when appropriate, the result is a significant increase in the average number of calls required to complete an interview. The issue was so severe that the interviewing period, anticipated to be from September to December of 2005 and 2006, had to be extended into January of 2006 and 2007. Adding to the expense of the extensions was the labour law requirement to provide statutory holiday pay for all returning interviewers.

8.2.3 Coding Staff

The increased cost of coding staff in 2006 was partly the result of an increase in the number of completed interviews and partly due to changes in procedures. The coding staff participated in an increased effort to avoid the use of street intersections as a location and in an increased effort to shorten the time between the completion of an interview and geo-coding. The result was a more rapid and thorough request to clarify incomplete or in-accurate information in the original interview. Also, for the first part of 2005 geocoding was done on paper, and then subsequently in the new Geocoding Console once it was brought online, which also contributed to increased geocoding costs overall.

Table 8.1 Actual Expenditures for TTS's in 1996, 2001 and 2006

	1996	2001	2006
Development and Testing			
Software Development	**233,000**	**21,000**	**163,000**
Interviewing			
Interview Staff and Training	**714,000**	**1,076,000**	**1,369,000**
Coding			
Coding Staff	**132,000**	**143,000**	**223,000**
Equipment			
Computer Hardware and Software	198,000	42,000	87,000
Telephones (Equipment and Charges)	24,000	94,000	183,000
Sale of Equipment	-75,000	-31,000	-22,000
Subtotal	**147,000**	**105,000**	**248,000**
Other Direct Expenses			
Printing and Mailing	73,000	120,000	208,000
Office Space and Furniture (Security in 1996)	86,000	187,000	266,000
Sample	19,000	31,000	34,000
Office Expenses and Supplies	25,000	26,000	16,000
Subtotal	**203,000**	**364,000**	**524,000**
Management and Coordination			
Management	**636,000**	**414,000**	523,000
Total Expenses	2,065,000	2,123,000	3,050,000
Post Survey Processing			
Reports and Analysis	309,000	300,000	101,000
Total Cost	**2,374,000**	**2,423,000**	**3,151,000**

8.2.4 Equipment

The combined cost of computer hardware and sale of equipment in 2001 was unique as the purchases occurred just at the time agencies were disposing of hardware in anticipation of a problem when the date changed to the year 2000. This combined cost in 2006 is a reflection of the true cost of purchasing and disposing of computer hardware, in particular, the personal computers used by the interviewers. Approximately half of the computers, which satisfied the needs of the first phase in 2005, were purchased as used equipment from a University of Toronto computer laboratory. After two years of use (2005 and 2006) on TTS these computers had limited resale value. The remainder of the personal computers were purchased new from Dell and account for most of the recovered cost.

The cost of telephone equipment in 2006 was reduced somewhat by recycling some telephones from the 2001 TTS. However, many of these sets encountered an unacceptable failure rate and had to be replaced. The telephone connection and charges were organized through the University of Toronto's Communications Office and reflect market rates.

8.2.5 Other Direct Expenses

The increased cost of printing and mailing in the 2006 Survey reflects two things. The first is that more pre-interview letters were required because of the call screening mentioned above and because of difficulty reaching apartment dwellers, particularly occupants of large apartment complexes. Due to a restriction imposed by the Canadian Radio-television and Telecommunications Commission, apartment numbers were not included in the sample detail. In an effort to overcome the poor response rate from apartment dwellers, a larger sample was used for dwellings in this category. More sample, hence more letters, were required per completed interview. Secondly, the management group made a decision to use the slightly more expensive third class postage rather than bulk mail and for some mailing blocks where the timely delivery of a pre-interview letter was essential to use the even more costly first class postage.

The cost of office space and furniture reflect the cost of renting commercial office space. In 1996, the Metropolitan Toronto Planning Department provided office space and furniture as part of their contribution to the cost of the survey. The amount shown is the net amount of the credit they received under the cost sharing agreement with the other agencies. The cost of office space in the 2001 survey reflected a reduced cost of occupying space at 500 University Avenue that was available during a change of use. The cost in 2006 reflects the true cost of commercial space in central Toronto.

8.2.6 Management

The increase in management cost from the 2001 to 2006 Surveys can be attributed in part to the more complex management structure used in 2006 and in part to the increased duration of the interviewing phases.

8.3 Unit Cost Comparison with Previous Surveys

Table 8.2 gives a comparison of the per interview 2006 survey costs with the 1986, 1991 and 1996 surveys after the previous survey costs have been adjusted for inflation. Inflation factors of 66.4%, 31.9%, 22.7% and 11.6% have been applied to the 1986, 1991, 1996 and 2001 survey costs, respectively, to make them comparable to the 2006 values.

The unit cost of conducting the interviews has been increasing since 1991 when the first survey was conducted using direct data entry to a computer file. The biggest savings of changing from pencil and paper to computer files has been in the cost of coding the data once the interview has been completed. Improved software design, more comprehensive and up-to-date reference databases, the use of direct data entry and the networking of computers have, together, resulted in a significant reduction in the unit cost of coding survey records since 1986.

The relatively high unit cost of the 1991 TTS can be attributed to the development costs associated with the writing and testing of the original DDE software being spread over the relatively small number of interviews that were conducted in 1991. The absence of any

significant development cost associated with the 2001 TTS contributed to the low unit cost of that survey. The low fixed cost, primarily management and co-ordination, associated with the 2001 survey resulted, to a large extent, from the use of tried and tested procedures, continuity of staffing from previous surveys and the effective staging of the survey over 2 years. Some of those cost savings were unique to the situation in 2001 and were not carried forward to the 2006 survey.

Ignoring the smaller survey in 1991, the growing cost of conducting an urban travel survey using a retrospective telephone interview is evident. These cost increases are mainly attributable to the difficulty in obtaining telephone contact with households.

Table 8.2 Unit Cost Comparisons for TTS's in 1986, 1991, 1996, 2001 and 2006

	1986 TTS	1991 TTS	1996 TTS	2001 TTS	2006 TTS
Assumed Inflation factor included	66.40%	31.90%	22.70%	11.60%	0.00%
Number of completed Interviews	61,453	24,507	115,193	136,379	149,631
Interviewing Cost	$318,000	$208,035	$886,700	$1,207,000	$1,617,000
Interviewing Cost/Interview	$8.61	$11.20	$9.44	$9.88	$10.81
Coding Cost	$333,000	$49,649	$132,200	$143,000	$223,000
Coding Cost/Interview	$9.02	$2.67	$1.41	$1.17	$1.49
Other Variable Cost	$113,000	$53,460	$177,300	$338,000	$524,000
Other Variable Cost/Interview	$3.06	$2.88	$1.89	$2.77	$3.50
Total Variable Cost	**$764,000**	**$311,142**	**$1,196,200**	**$1,688,000**	**$2,364,000**
Total Variable Cost/Interview	**$20.69**	**$16.75**	**$12.74**	**$13.81**	**$15.80**
Fixed Cost	$190,000	$180,400	$721,900	$414,000	$523,000
Fixed Cost/Interview	$5.14	$9.71	$7.69	$3.39	$3.50
Development Cost	$38,000	$172,900	$146,500	$21,000	$163,000
Development Cost/Interview	$1.03	$9.31	$1.56	$0.17	$1.09
Total Cost	**$992,000**	**$664,500**	**$2,064,600**	**$2,123,000**	**$3,050,000**
Total Cost/Interview	**$26.86**	**$35.76**	**$21.98**	**$17.37**	**$20.38**

Section 9 Conclusions and Recommendations

Despite the problems with a low response rate for apartment dwellers, the validation results indicate that the overall travel data in the 2006 database is of a quality similar to previous surveys. There is, however, a growing trend for young people in the work force to be underrepresented in the results. The most likely cause of this trend is the growth in cell phone use and the increasing number of households without conventional telephone service. Telephone directory listings are the basis for the TTS sample and cell phones are not listed.

9.1 Data Quality

Every TTS has used the same basic survey instrument, which uses a telephone interview to record a retrospective reporting of travel by all members of a household on the day prior to the interview. The interview is conducted with the person answering the telephone unless that person is unaware of the travel by other members of the household. In which case, an attempt is made to interview the other household member(s) either during that interview or a subsequent 'callback'. A retrospective survey has an inherent bias resulting from forgotten trips by the respondent, which could be compounded by the respondent reporting trips taken by another household member. The impact of these forms of instrument bias is reflected primarily in underreporting of discretionary trips (trips taken for purposes other than work or school). The underreporting is understood and appears to be of the same magnitude for all TTS, including the 2006 TTS.

The change in survey instrument from a pencil and paper recording of a telephone interview to direct recording to a computer file resulted in an improvement in the recorded number of trips per person over the age of 11 years. The improvement is more likely the result of aids provided to the interviewer using lookup files than to the method of recording. The rate dropped slightly (3%) from 2001 to 2006 but has remained reasonably constant over the last four TTS. Comparisons with 2006 Cordon Count and transit ridership data from several sources reveal no evidence of any underreporting of morning peak period, work trip or school trip data. Analysis by trip purpose indicates that the differences are primarily in the amount of discretionary travel recorded. Care should therefore be exercised in drawing any conclusions as to trends in trip rates. Comparisons with the 1986, 1991, 1996, 2001 and 2006 survey data reveal a high degree of consistency in the distribution of trip rates, modal splits, trip lengths and many other factors.

Validation of the expanded survey data included demographic comparisons with data from the 2006 Canada census. Two significant differences were identified:

1. An underrepresentation of apartment units relative to houses and townhouses. Precise estimates of the degree of underrepresentation are not possible due to differences in definition between the census and TTS. Statistics Canada has made changes to the enumeration process used to classify dwelling unit type since the previous census and it would appear that this has led to the reclassification of a significant number of dwelling units in some areas, most notably the City of Toronto.
2. In the TTS the population in the 20 to 30 age range is underrepresented by 20% relative to the census with a corresponding overrepresentation in other age groups. These discrepancies in age distribution are much larger than in previous surveys.

The nature of the above discrepancies is consistent with the problems previously identified but there may well be other factors that contribute to hidden bias. The comparisons do not identify the "cause" and "effect" needed to estimate the impact of each problem factor.

The fact that there are discrepancies between the census and TTS with regard to demographic data does not necessarily mean that there are similar problems with the travel information. In fact comparisons made with cordon and transit ridership counts suggest that the 2006 TTS data is at least as good, and possibly better, than previous surveys with respect to aggregate travel patterns especially public transit use.

The concern with the low response rate is that there could be other hidden biases that are not revealed in the validation. In addition the underlying problems can only be expected to get worse in future surveys.

9.2 Software

The 2006 TTS was the largest travel survey conducted to date and utilised the technological developments that were implemented in previous surveys.

1986 The 1986 TTS was a pioneer in the use of automated geocoding.

1991 The 1991 TTS was the first to use Direct Data Entry. Although the information was compiled without the aid of a computer network, it was the first application of recording interviews directly on a computer file.

1996 The most significant new development for the 1996 survey was the on-line networking of the interview computers.

2001 No significant changes were made to the software for the conduct of the 2001 TTS. While significant cost savings were realised the software limitations became evident.

2006 The entire data entry, sample control and geocoding process was reviewed and a complete re-write of the software was undertaken for the 2006 TTS. The process began in early 2004 and improvements were implemented through the entire survey period.

9.3 Hardware

Very few computer hardware problems were experienced during the conduct of the survey. The decision to have only two different personal computer hardware models made rapid updates possible. The purchase, and subsequent resale, of used name brand equipment is recommended as the most cost effective and efficient way to equip a survey of this magnitude. The fileserver is central to most operations. "Over" purchasing, in terms of its performance, reliability and back up capabilities, is recommended.

9.4 Supervisory Staff

Finding an adequate number of staff with the experience and background necessary to act in a supervisory role is a significant challenge in the conduct of each TTS. The quality of first level supervision is probably the single most important aspect in overall quality control. Early in the recruiting process in 2005 and 2006 previous supervisors and interviewers in good-standing were contacted with an offer of employment. We were fortunate to have a significant number of past employees return. The team leaders for the main survey were selected from returning staff, as was the chief assistant to the hiring and training manager and the daytime manager. The other

supervisory positions were filled from the early ranks of the interview staff (some number of whom also had previous TTS experience).

Supervisory responsibilities include:

- The training of new interviewers.
- Supervision of and assistance to interviewers.
- Selective monitoring of interviews in progress.
- Visual review of completed interviews.
- Review of call back information.
- Entry of corrections.

Efforts to build the foundation of staff that will want to return to future TTS projects should be continued and contact lists and employment details of previous employees should be maintained for future TTS projects. Returning employees understand the scope and intent of the project, reach production targets more quickly and have nearly twice the retention of staff hired without TTS experience. Conducting a smaller scale survey in the year prior to a full-scale survey provides an essential opportunity to pre-train a critical mass of interviewers and provides a pool of trained staff from which to select supervisory personnel for the main component of the survey.

9.5 Interview Site

The central site location in Toronto with convenient subway access proved to be extremely good. There was no shortage of applications for interview and coding staff positions. As mentioned previously, there were relatively few people with the maturity and experience needed for supervisory positions. The use of space in the same building for both the 2005 and 2006 components of the survey was an added convenience although not as important as the downtown location and subway access.

Site costs were significantly higher than previous surveys due to the need to rent commercial office space.

9.6 Advance Letter

The advance letter has always been regarded as a critical item in reducing respondent refusals. Having the advance letter increases interviewer's confidence and provides respondents with a measure of the survey's validity. While it has been shown that experienced and competent interviewers can achieve the same degree of respondent compliance with or without the letter, the reality of the varied skill levels of the interviewers, and short time frame in which interviewing is done, dictates the necessity of the letter. Households where respondents report having received the letter usually require less explanation from the interviewer, are completed more quickly and often have more detail.

Approximately 45% of respondents in 2006 claimed not to have received the advance letter, a 10% increase from 2001 but approximately equal compared to the 1996 TTS. In 2001 it was felt that the use of Government of Ontario envelopes aided in the higher reporting of letter receipt. Non-government envelopes were used for the 1996 TTS. The continued use of official Government envelopes is recommended for all future surveys. Households reporting receipt of

the letter have been fairly consistent at approximately 50% since 1996. In the 1991 TTS, when complete address information was available for all households, including apartment buildings, and Government envelopes were used, 65% respondents reported that they had received the letter.

Receipt of advance letter (not in 1986):

	2006	2001	1996	1991
Unknown	0.6%	7.7%	5.9%	2.4%
No	46.5%	36.9%	45.2%	33.1%
Yes	52.9%	55.4%	48.9%	64.5%

Receipt of the advance letter significantly reduces the refusal rate, probably by about 15% (consistent with previous experience when there has been a problem with the mailing). The fact that residents of apartment buildings are less likely to receive the advance letter, due to the exclusion of apartment numbers from the address information, produces a measurable bias in the survey results due to apartment units being underrepresented.

Control letters to survey staff members were included in each mailing as a check on the timing. Based on previous experience, bulk mail was not used. Canada Post offers no guarantee for bulk mail as to how long delivery will take. The cost of third class postage is slightly higher but there are savings in mail preparation costs since the letters do not have to be pre-sorted. Testing was done in 2005 to compare the use of first and third class mail services. First class mail was used in 2006 only at the start and end of the survey when prompt delivery was essential. The commercial mailing house was cost effective and efficient in preparing the mailings, as was the case in 1996 and 2001.

In future surveys, the advance letter should include the site's phone number to allow potential respondents to call-in directly. This requires a sample control software modification to allow access to households that may not have been released into the calling queue yet.

9.7 Productivity

Table 9.1 shows two measures that are factors in determining both productivity and the quality of the survey results.

Table 9.1 Productivity and Quality Measures

	Calls per completed interview	Overall Response Rate
1986 TTS	not available	60%
1991 TTS	not available	72%
1996 TTS	3.71	70%
2001 TTS	4.74	64%
2006 TTS	6.71	45%

The average number of phone calls made per completed interview in conducting the 2006 TTS was 40% higher than in the 2001 TTS and 80% higher than in the 1996 TTS. More calls per completed interview translate into the need for more interviewers, more equipment, more training

and more supervision. Quality control inevitably suffers due to production pressures and the finite resources available.

Overall response rate is the number of completed interviews divided by the number of households where contact was attempted. The lower the response rate the greater the potential for hidden biases in the survey results in addition to any bias that might be present in the original sample frame.

All of the potential measures of interviewer productivity have steadily deteriorated since 1986. These measures include: number of calls per completed interview, number of answering machines encountered and number of refusals after contact is made. The situation became so severe in 2006 that interviewing in both phases had to be extended into January and February. In addition, the budget had to be increased in order to meet the target of a 5% sample.

It is unfortunate that the call display could not identify an agency such as the Ministry of Transportation as it was evident that call screening is a growing phenomenon. It is expected that any attempt at using telephone interviews in the future will encounter more difficulty in making contact, and likely experience more refusals.

9.8 Student Population

Student travel is an important component of total daily travel patterns with distinct characteristics. Two problems exist in capturing information on that component. The first problem is in obtaining a representative sample that includes the student population. The second is the method of expansion given that the Canada Census is not done during the post secondary school year. It is not known to what extent on campus residences are represented in the sample. It is clear from comparisons with post-secondary enrolments, that this section of the population is underreported in TTS.

9.9 Sample Selection and Management

The problems in sample selection that were experienced in the 2006 component of the survey indicate the need for a review of the alternative sources of sample lists prior to the next survey and the need for rigorous checking of sample lists to the extent that it is possible prior to having the results of the interviews. Anecdotally, there is a growing problem with households using cell phones as their only telephone service. These tend to be young people in the work force; a demographic that has been underrepresented in the TTS and is a growing trend.

In particular, a sample list needs to contain complete address information, including apartment numbers, and must contain households not listed in the telephone directory (households using cell phones exclusively). It would be beneficial for the listing to be current, which would include post-secondary students renting accommodation for the school term, and include an identifier for apartments. A rigorous checking of the sample list needs to be undertaken to ensure complete and equitable geographic coverage.

Although the original sample information did not contain apartment numbers, those records with an address that was repeated 6 or more times in the complete database from which the sample was drawn were flagged as multi-unit addresses. During phase 1 (external areas) of the 2006

TTS it was noted that the response rate for those “flagged” records was 20% lower than for non-flagged records. In phase 2 (GTHA) “flagged” records were sampled at a 20% higher rate than non-flagged records to compensate for the expected difference in response rate. Subsequent analysis showed that in some areas the difference in response rate was significantly greater than 20%. Within the City of Toronto it was about 35%.

9.10 Geocoding

Duplication of street and municipal names within the vast survey area made coding especially difficult. For example, there are 52 Church Streets in the survey area without accounting for variations such as Church Road, Church Lane and Church Street East and West. Coding small towns and hamlets in rural areas were also more difficult because of the lack of commercial street maps and reference materials.

Also some street names used and reported to interviewers by locals tended to be different from the official names found on maps and in reference materials.

Overall, coding productivity (quicker turn-around and more accurate locations) improved since the 2001 TTS. The improvement is attributed to several factors:

- The quality of the reference street network was better than in previous years.
- The new Geocoding Console was easier to use than the previous version and allowed easier searching of the reference databases. It also allowed the coders to see historical changes to the household which could give further hints as to locations that were difficult to code.
- There was more interaction between the coding staff and the interview and daytime review staff than in previous years. This allowed the daytime review staff to be more aware of what was unacceptable for coding purposes and hence to pre-screen some of the more difficult to code locations before they ever reached geocoding.
- The use of search engines such as google, google.maps and 411.ca provided advancement in the use of the internet for search purposes which allowed coders to be more efficient. This saved both time and effort in looking up addresses for uncommon monuments recorded in the interviews.
- Grouping the coders into units by large geographic areas enabled the coders to gain experience in particular areas while allowing them to assist one another in solving problem records. It is worth noting that there were no partitions between coding stations as there were with interview stations. This was to allow coders to freely communicate with one another and share reference materials.

9.11 Coding Reference Databases

Coding of most of the street and intersection databases was easier in 2006 due to the street base map for the entire area being obtained from one organization, Land Information Ontario (LIO). This eliminated much of the processing to consolidate the data which had occurred previously when the files were being obtained from multiple sources.

Coding of the monument files began a few months before the survey’s start. For future surveys it is recommended that development work on the reference databases start even earlier. It is also

recommended that the procedures for updating of the reference databases during the actual interview phase of the survey be reviewed and streamlined.

Section 10 Recommendations for 2011

10.1 Background

The basis of all five previous TTS was a retrospective of trips taken during the previous day by all members of a household. The information was collected from a telephone interview. A 5% sample of households was the target and the universe of households and estimates of total travel were based on the number of households reported in the national census.

Applications of the TTS data by a wide variety of users has evolved over the years to assume a content and level of accuracy that is possible with a large sample using a consistent set of questions during the interviews. However, if a TTS is to be conducted in 2011 and the decision is taken to maintain as much consistency as possible, several issues should be addressed:

- A growing number of households do not have a listed telephone number as they use a cell phone exclusively and these households are not equally distributed over the universe of households.
- A growing number of households use call-screening.
- Telephone listings do not include the unique address of apartment units.
- Post-secondary students are underrepresented in the sample.

The TTS Management Team recommends a set of changes to a possible TTS in 2011, while still maintaining the same basic survey instrument. The concept is to maintain consistency with existing data while, at the same time, testing some alternate data collection procedures.

10.2 A Feasible Approach

Using the standard telephone directory as a sample source is no longer effective. Any alternate sample source representing a cross-section of all households is unlikely to contain complete information for each household. One possible procedure is to obtain a sample from the Municipal Property Assessment Corporation (MPAC). The sample would likely contain the complete address, including the apartment number, but would not contain the occupant's name. A reverse telephone lookup on all households that have a unique street address should yield a unique telephone number for 50% to 60% of the sample in the GTHA and more in external areas.

10.2.1 Survey Method 1

Households which were successfully matched with a telephone number would be sent a pre-interview letter and be interviewed by telephone in the same manner as previous surveys. It is important that the telephones at the call centre be installed through the Province of Ontario exchange, as was the case in 2001. The call display would then indicate 'Province of Ontario', which should help reduce the incidence of call screening.

10.2.2 Survey Method 2

Households not successfully matched with a telephone number would be sent a letter to their unique address with a request to complete a survey either by calling in or via the Internet. A call centre would be set up that would be specifically designed to receive calls and conduct the interview immediately. A browser based web site would be established to complete the survey

questions. The respondent could complete the survey independently online, and would have a telephone number to call with any questions. The call centre could display the current status of the household completion and guide the respondent through to completion over the phone. If the sample was not completed within a given time period, a follow-up letter would be sent.

10.2.3 Survey Method 3

Post-secondary students pose a unique problem and a unique opportunity. Anecdotal evidence suggests the incidence is very high of these students relying entirely on a cell phone. At the same time, they tend to be very computer literate with ready access to the Internet. A sample of students would be solicited from all the post-secondary institutions in the study area, which would contain their email address. The students would be contacted by email and asked to complete the survey by telephone or the Internet. The results of these interviews would then be integrated into the estimates of travel with consideration given to the possibility of double counting.

10.3 Issues Requiring Early Attention

10.3.1 Sample Selection

A contact with MPAC is likely to be more effective if initiated by the Ministry of Transportation and perhaps some regional municipalities. If a sample from MPAC is not possible, other possible methods of sample selection need to be investigated as soon as possible.

10.3.2 Browser Based Interview

Large scale prototype testing is necessary in late 2009 and early 2010. The development of the necessary software is underway and needs to be continued if this deadline is to be met.

10.3.3 Development of Cost Estimates

It should be anticipated that the cost per completed interview will be significantly higher than for previous TTS's. Factors contributing to higher costs include:

- Development costs associated with the on-line component and other software modifications.
- Continuation of the downward trend in productivity associated with the telephone components.
- Higher per unit costs associated with the mail only component.
- Additional sample, pre-processing and post-processing costs associated with the increased complexity of the survey.

Appendix A Letter to Local Officials

NEWS RELEASE FOR IMMEDIATE RELEASE
August, 2006

Transportation Tomorrow Survey To Include More Than 140,000 Households

TORONTO — Twenty regional, county and local municipal governments are participating in a major travel survey of more than 140,000 households designed to help municipalities meet their future needs for roads and transit services.

"The 2006 Transportation Tomorrow Survey will examine the travel habits and preferences of residents of the Greater Toronto Area as well as the extended area from St. Catharines to Barrie and Peterborough," announced Gerald Steuart, the project director of the survey. "It will help in making decisions about road and transit improvements, and provide information for long-term planning."

The first phase of the survey took place in the Fall of 2005 when survey staff will contacted over 35,000 household in areas surrounding the Greater Toronto Area, including Niagara Region, Brantford, Waterloo Region, Wellington County, Guelph, Dufferin County, Orangeville, Simcoe County, Barrie, City of Kawartha Lakes, County and City of Peterborough. The second phase will occur in the Fall of 2006 when over 105,000 households will be contacted in the Regions of Durham, Halton,, Peel and York, and the Cities of Hamilton and Toronto.

"This survey will help us better respond to each community's needs," said Mr. Steuart. "The population of the survey area is expected to grow to well over seven million people in the next 20 years. We need to assess how this will affect our transportation system and ensure that it can meet the increased requirements."

This is the fifth Transportation Tomorrow Survey. The first was conducted in 1986, a second in 1991, a third in 1996 with a fourth in 2001. Information

More ...

gathered in previous surveys has been used to plan a wide range of transportation initiatives in the Greater Toronto Area.

The survey consists of a telephone interview of randomly selected households. In addition to trip information for each household member (i.e., origin, destination, time, reason for travel, mode of transportation), interviewers will also ask about the number of vehicles available for personal use and where each family member works or attends school.

The University of Toronto's Data Management Group, hired to develop and carry out the survey and gather the results, is conducting the survey.

Used for statistical purposes only, all information related to individual households will be kept strictly confidential. Once the study is complete, the survey results will be collated and released early in 2007.

For further information, please contact:
Gerald Steuart
Project Director
Transportation Tomorrow Survey
(416) 978-5979

BENEFITS OF A COMPREHENSIVE TRANSPORTATION SURVEY

1. **Helps Identify Transportation Needs and Impacts**

   - Estimation of transportation implications of short and medium term land use changes, particularly in high growth areas;
   - Identification of cross boundary needs;
   - Monitoring effectiveness of existing transportation systems;
   - Travel behaviour change; and
   - Assessment of local transportation impacts.

2. **Provides Much Needed Data**

   - Capture changing travel patterns in a rapidly changing urban environment;
   - Build on existing time series data (particularly important in high growth areas);
   - A reliable means of capturing cross boundary data;
   - Important data on changing transit use;

3. **Provides Valuable Information For Many Agencies**

   - Planning and Development Departments;
   - Engineering Departments;
   - Finance Departments;
   - Transit Departments;
   - Federal Government (Airport access);
   - School Boards;
   - Social Agencies;
   - Emergency Service Planning Coordinators;
   - Housing Industry;
   - Ministry of Transportation;
   - Ministries of Energy, Housing, Treasury & Economics and Treasury Board;
   - GO Transit;
   - Consultants;
   - Developers.

4. **Enables Cost-Effective Transportation Improvements**

   - Design of transit services;
   - Identification of low ridership areas and strategies to improve ridership;
   - Structuring of routes to serve non-central destinations;
   - Monitoring cross boundary travel;
   - Phasing of highway improvements;
   - Monitoring of transportation for both official plans and individual developments;
   - Input to development proposals;
   - Determining need for GO Transit improvements;
   - Development and calibration of travel forecast models;
   - Determining need for road improvements.

Appendix B Advance Letter GTHA

TRANSPORTATION TOMORROW SURVEY

We are conducting an important travel survey on behalf of your municipality, other municipalities in central Ontario, and the Province of Ontario. Every five years for the past 20 years, we have conducted this survey so that we may keep up with changing transportation needs. The purpose of this survey is to collect information on the travel choices and preferences of people in the area. We need your help to provide this information so we may continue to plan transportation services to meet your area's future needs.

Here is how it works. You will be telephoned at home by a professional interviewer and asked to spend about 10 minutes answering questions. A sample list of the questions to be asked is shown on the back of this letter. The interviewer will call sometime in the next two weeks. On weeknights, the calls will be made between 5:30 p.m. and 9:30 p.m. If the interviewer calls on a Saturday, it will be between 10:00 a.m. and 5:00 p.m.

Please inform other members of your household that you have received this letter and to expect our telephone call.

Most of the questions will be about travel on the weekday before the call, for you and those members of your household who are over 11 years old. We would like to know specific information about where and when trips were taken by each member of your household. This information from approximately 150,000 households will give us an accurate picture of travel needs to plan improved transportation services and facilities in your area.

All information will be kept strictly confidential. No information will be released in any way that could be traced to your household. Your answers will be combined with other responses in your area. This information will be used to forecast travel patterns and recommend future transportation plans.

If you have any questions, please call the Ministry of Transportation at 1-800-268-4686, or visit our web site at www.TransportationTomorrow.on.ca <http://www.TransportationTomorrow.on.ca>

We would like to extend our personal thanks for your assistance in this project. Your help will mean better transportation services in the future.

Regards,

John Oosterhof, Warden
County of Dufferin

Peter Partington, Chair
Regional Municipality of Niagara

Robert Hamilton, Mayor
City of Barrie

Kate M. Quarrie, Mayor
City of Guelph

Ken Seiling, Chair
Regional Municipality of Waterloo

Sylvia Sutherland, Mayor
City of Peterborough

Drew Brown, Mayor
Town of Orangeville

Neal Cathcart, Warden
County of Peterborough

Dennis Roughley, Warden
County of Simcoe

Barbara A. Kelly, Mayor
City of Kawartha Lakes

Brad Whitcombe, Warden
County of Wellington

Mike Hancock, Mayor
City of Brantford

Harinder Takhar, Minister
Ministry of Transportation Ontario

Survey Questions

A. About your household

- Type of building (house or apartment)
- Number of people
- Number of vehicles available for personal use

B. About each person

- Their age
- Do they have a driver's licence?
- Where do they work or go to school (street address)

C. About each trip made by each person the previous day

- From where, to where (street address preferred or otherwise building name)
- Reason for making the trip (e.g. shopping)
- Start time of the trip
- Type of transportation (bus, car, bicycle, etc.)

We will only be collecting trip data for persons 11 years of age or older. A trip is a one-way journey from one location to another by any form of motorized transportation or bicycle. We will request some information on walking, but only for trips to and from work or school.

Authority for collection of this information has been obtained from each of the Regional and Local governments participating in this survey. Confidentiality of this information is protected under the Freedom of Information and Protection of Privacy Act.

Appendix C Advance Letter Areas External to the GTHA

TRANSPORTATION TOMORROW SURVEY

The Transportation Tomorrow Survey is an important travel survey, conducted in partnerships with your municipality, other municipalities in central Ontario, and the Province of Ontario. This survey has been conducted every five years for the past 20 years, so that we may keep up with changing transportation needs. The purpose of this survey is to collect information on the travel choices and preferences of people in Central Ontario. We need your help to provide this information so we may continue to plan transportation services to meet your area's future needs.

Here is how it works. You will be telephoned at home by a professional interviewer and asked to spend about 10 minutes answering questions. A sample list of the questions to be asked is shown on the back of this letter. The interviewer will call sometime in the next two weeks. On weeknights, the calls will be made between 5:30 p.m. and 9:30 p.m. If the interviewer calls on a Saturday, it will be between 10:00 a.m. and 5:00 p.m.

Please inform other members of your household that you have received this letter and to expect our telephone call.

Most of the questions will be about travel on the weekday before the call, for you and those members of your household who are over 11 years old. We would like to know specific information about where and when trips were taken by each member of your household. This information from approximately 150,000 households will give us an accurate picture of travel patterns to plan improved transportation and transit services and facilities in your area.

All information will be kept strictly confidential. No information will be released in any way that could be traced to your household. Your answers will be combined with other responses in your area. This information will be used to forecast travel patterns and recommend future transportation plans.

If you have any questions, please call the Ministry of Transportation at 1-800-268-4686, or visit our web site at www.TransportationTomorrow.on.ca <http://www.TransportationTomorrow.on.ca>

We would like to extend our personal thanks for your assistance in this project. Your help will mean better transportation services in the future.

Regards,

Larry Di Ianni, Mayor
City of Hamilton

David Miller, Mayor
City of Toronto

Roger Anderson, Regional Chair
Regional Municipality of Durham

Joyce Savoline, Chairman
Regional Municipality of Halton

Emil Kolb, Regional Chair
Regional Municipality of Peel

Bill Fisch, Regional Chair and CEO
Regional Municipality of York

Donna Cansfield, Minister
Ministry of Transportation Ontario

2006 Transportation Tomorrow Survey
Survey Questions

A. About your household

- Type of building (house or apartment)
- Number of people
- Number of vehicles available for personal use

B. About each person

- Their age
- Do they have a driver's licence?
- Where do they work or go to school (street address)

C. About each trip made by each person the previous day

- From where, to where (street address preferred or otherwise building name)
- Reason for making the trip (e.g. shopping)
- Start time of the trip
- Type of transportation (bus, car, bicycle, etc.)

We will only be collecting trip data for persons 11 years of age or older. A trip is a one-way journey from one location to another by any form of motorized transportation or bicycle. We will request some information on walking, but only for trips to and from work or school.

Appendix D Advance Letter In French

Le Sondage pour les systèmes de transport de demain est un important sondage sur le transport, effectué en partenariat avec votre municipalité, avec d'autres municipalités du Centre de l'Ontario et avec la province d'Ontario. Ce sondage est effectué tous les cinq ans depuis 20 ans, afin que nous puissions nous tenir au courant des besoins changeants en matière de transport. Le but de ce sondage est de recueillir des renseignements au sujet des choix et des préférences en matière de déplacements des personnes habitant dans le Centre de l'Ontario. Nous avons besoin de votre aide pour obtenir ces renseignements, afin que nous puissions continuer à planifier des services de transport qui répondront à la demande future de votre région.

Voici ce qui se passera. Un intervieweur professionnel vous téléphonera à la maison et vous demandera de répondre à des questions pendant une dizaine de minutes. Une liste d'exemples de questions qui seront posées figure au verso de la présente lettre. L'intervieweur vous appellera au cours des deux semaines à venir. Les soirs de semaine, les appels seront faits entre 17 h 30 et 21 h 30. Si l'intervieweur vous appelle le samedi, il le fera entre 10 h et 17 h.

Veuillez aviser les autres membres de votre ménage que vous avez reçu la présente lettre et que vous recevrez un appel à ce sujet.

La plupart des questions portent sur les déplacements effectués le jour de la semaine qui précède l'appel, pour vous et pour les membres de votre ménage qui ont plus de 11 ans. Nous souhaitons obtenir des renseignements précis sur le moment et la destination des déplacements effectués par chaque membre de votre ménage. Les renseignements recueillis auprès d'environ 150 000 ménages nous donneront une représentation précise des déplacements, ce qui nous permettra de planifier des services de transport routier et de transport en commun dans votre région.

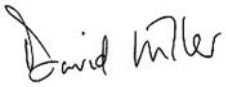
Tous les renseignements demeureront strictement confidentiels. Nous ne divulguerons aucun renseignement qui puisse servir à identifier votre ménage. Vos réponses seront combinées à celles d'autres ménages de votre région. Ces renseignements seront utilisés pour prévoir les tendances en matière de déplacement et pour recommander des plans pour les systèmes de transport à l'avenir.

Si vous avez des questions, veuillez appeler le ministère des Transports au 1-800-268-4686; vous pouvez également consulter notre site Web à l'adresse www.TransportationTomorrow.on.ca <http://www.TransportationTomorrow.on.ca>

Nous souhaitons vous remercier personnellement pour l'aide que vous apportez à ce projet. Votre aide nous aidera à vous fournir de meilleurs systèmes de transport à l'avenir.

Veuillez agréer l'expression de notre considération distinguée.

Larry Di Ianni, maire
Ville de Hamilton



David Miller, maire
Ville de Toronto

Roger Anderson, président régional
Municipalité régionale de Durham

Joyce Savoline, présidente
Municipalité régionale de Halton

Emil Kolb, président régional
Municipalité régionale de Peel

Bill Fisch, président régional et PDG
Municipalité régionale de York

Donna Cansfield, Ministre
Ministère des Transports de l'Ontario

Sondage 2006 pour les systèmes de transport de demain
Questions du sondage

A. Au sujet de votre ménage

- Type d'immeuble (maison ou appartement)
- Nombre de personnes
- Nombre de véhicules disponibles pour usage personnel

B. À propos de chaque personne

- Son âge
- Possède-t-elle un permis de conduire?
- L'endroit où elle travaille ou où elle étudie (adresse civique)

C. Au sujet de chaque déplacement fait par chaque personne la journée précédente

- Origine et destination (adresse civique préférée, sinon le nom de l'immeuble)
- Raison du déplacement (p. ex. faire des achats)
- Heure de départ du déplacement
- Mode de transport (autobus, voiture, bicyclette, etc.)

Nous ne recueillons des renseignements sur les déplacements que pour les personnes âgées de 11 ans et plus. Un déplacement est un trajet dans une seule direction, d'un endroit à l'autre, utilisant n'importe quel mode de transport motorisé ou à bicyclette. Nous demanderons certains renseignements au sujet de la marche, mais seulement dans le cas de déplacements entre la maison et le travail ou l'établissement d'enseignement.

L'autorité pour la collecte de ces renseignements a été obtenue de chacun des gouvernements régionaux et municipaux qui participent au présent sondage. La confidentialité de ces renseignements est protégée en vertu de la Loi sur l'accès à l'information et la protection de la vie privée.