UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING
Transportation Research Institute

UTTRI

EVALUATING FACE-TO-FACE SURVEYS TO AUGMENT REGIONAL TRAVEL SURVEYS

2017-12-15

Transportation Tomorrow Survey 2.0

Chris Harding, Siva Srikukenthiran
Khandker Nurul Habib & Eric J. Miller

EXECUTIVE SUMMARY

The methods employed for carrying out the Transportation Tomorrow Survey (or TTS) are currently under revision. The survey was conducted in much the same way every five years since 1986, using a landline-based sampling frame and telephone interview method. In response to issues in representation and emergence of new data collection regimes, a research and development project named TTS 2.0 was launched to investigate alternative approaches to collecting data on travel demand in the region.

The motivation for this project is simple: declining rates of landline penetration and increasing reticence to participate among households who do still have a landline at home. To address this problem, a next-generation core-satellite framework for data collection is being designed. Where the TTS has historically been carried out as one monolithic effort, the next generation TTS is likely to include a variety of data collection tools, ranging from smartphone apps to web surveys. Passive data, multiple sampling frames and data fusion will all also likely be components of future TTS efforts.

One method of data collection which has not been given much consideration in the region of late is the possibility of making use of the oldest form of travel survey: in-person, or face-to-face interviews. While carrying out a household travel survey for a region as large as Toronto using exclusively in-person interviews would be a very complex and costly task, there may well be value in using targeted cluster sampling to augment other methods of data collection in parts of the region where other methods of respondent recruitment have proven to perform poorly.

The cost-effectiveness and data quality implications of such an approach are what this project set out to assess.

Seven census tracts within the region where response rates were deemed problematic for certain demographic groups in the 2011 TTS were selected for our study. Invitation letters were delivered an average of 6 days beforehand to advise residents that interviewers would be coming by to collect travel information. Interviewers with Internet-connected tablets were dispatched to these tracts to conduct computer-assisted personal interviews (CAPI), going door-to-door to preselected locations. Survey codes were left behind for persons not reached, meaning residents could also later answer the survey if they were not available when interviewers were present.

A detailed log of labour hours and other costs was kept, to allow the cost-effectiveness of the data collection effort to be assessed. From both a cost-effectiveness and demographic perspective, the effort was a success. Without having an official government seal on invitations to increase the credibility of the effort, 985 completed surveys were obtained from approximately 6,100 listed addresses – an overall completion rate of 16%. Including survey materials, incentives, transportation and labour (interviewers), but excluding the time put in by the lead research assistant, the cost per completed survey came in at around $13 per survey. Calculations made to assess the potential cost per survey were the approach to be employed on a larger scale and without any free student labour, indicate a cost per survey of $22 is a reasonable expectation – on par with current TTS costs. The method may not be generalizable to the entire region, but for certain areas, CAPI may be an effective means by which to reach prospective respondents.

Travel behaviour of those who filled out the survey with one of our interviewers would appear to be slightly different from the travel behaviour of residents who were not interviewed in person, with these slight differences explored. The respondents who did not open the door when interviewers visited, but rather used the survey code left behind and answered online on their own, reported a slightly higher daily trip rate (3.2 to 2.8 trips per day) but were also more likely to be retired. This could indicate a number of things; it would be reasonable to interpret this higher reported trip rate as a sign of greater dedication to providing accurate data, given that these respondents chose to answer the survey, not because of pressure felt by having an interviewer in front of them, but because they have an interest in the topic, or simply more free time. It could also indicate, however, that unlike the respondents who were interviewed directly, they ignored the instructions and reported a 'typical' day instead of reporting the travel episodes of the day prior.

While the summer CAPI survey employed had an equivalent respondent burden to a single person household responding to the TTS, it was not a replication of the TTS. Where the TTS asks for all trips made by all household members over the age of 11, this survey asked considerably more questions about the household, but then only asked for one travel diary per household. A TTS-replication survey was tested in the last week of data collection with similar results, but the sample size is insufficient for robust conclusions to be drawn.

Related to this issue of one versus many travel diaries per household, one of the lessons learned was that counting on an initial household respondent to pass along an email invitation to the other members of the household for them to in turn answer a travel diary was not an effective approach: only 30 travel diaries were collected in this manner.

The most interesting finding, however, may well be that there were only a handful of incompletes when residents spoke with interviewers face-to-face. Whether respondents were asked to answer a long survey about the planning process or provide travel diaries for multiple members of a given household, in only a few instances did the respondent ask to end the survey – less than 0.5% of cases. Considering how long and detailed the surveys were, this is an important finding that indicates household travel surveys such as the TTS may be well suited to face-to-face/door-to-door CAPI methods.

Issues of language barriers (inability to follow-up with non-English speakers) and physical access to residents in condo and apartment buildings still need to be investigated, but preliminary analysis of this data shows there is potential for including CAPI in the TTS core and satellite framework..

The report ends with a discussion of improvements that can be brought to the CAPI approach if scaling it up is to be considered.

Evaluating Face-To-Face Surveys to Augment Regional Travel Surveys

TRANSPORTATION TOMORROW SURVEY 2.0

Table of Contents

Executive Summary 1

1 Introduction 7

2 Advantages and disadvantages of CAPI 8
- **2.1 Advantages of CAPI and face-to-face methods** 8
- **2.2 Disadvantages of CAPI and face-to-face methods** 9

3 Data collection 12
- **3.1 Interviewer selection** 12
- **3.2 Sampling strategy** 12
- **3.3 Initial and follow-up contact** 14
- **3.4 Presentation** 14
- **3.5 Incentive structure and survey design** 14

4 Assessment approach 16
- **4.1 Cost-effectiveness of door-to-door CAPI work** 16
- **4.2 Demographics** 16
- **4.3 Reported travel** 16

5 Results and discussion 17
- **5.1 Costs and cost-effectiveness** 17
- **5.2 Effect of delay and delivery day on completion rates** 18
- **5.3 Respondent demographics** 21
- **5.4 Reported travel** 23
- **5.5 Differences between 'in-person' and 'on their own' responses** 26
- **5.6 TTS versus optional additional household diaries** 29

6 Conclusions 30

7 Scaling up 31

8 Problems faced, errors made and options to explore for future data collection work 33
- **8.1 Design** 33
  - 8.1.1 Response alternative issues 33
  - 8.1.2 Loops 34
  - 8.1.3 Completion time 34

8.1.4 Website issues 35
8.2 Field 35
8.2.1 Apartment and condo residents 35
8.2.2 Survey code 35
8.2.3 Address list problems 35
8.2.4 Sunlight, Safety, and Scheduling 37
8.2.5 Quality assurance 37
8.2.6 Mailers 39
8.2.7 Proximity to transit 39
9 References 40

List of Tables

Table 1 - Main advantages and disadvantages of personal interviews, as per TravelSurveyManual.org (Chapter 3), reproduced 8
Table 2 - CAPI 2017 non-labour costs 17
Table 3 - Effect of the number of days (delay) between initial invitation letter delivery and follow-up visit by interviewers, leaving mailers with code if no one answers or if no one wants to answer the survey at that moment 18
Table 4 - percent completes by source and delay between initial and follow-up contact 20
Table 5 - Day of 2nd contact made and completion rates 20
Table 6 - Demographics of Census 2011, TTS2011, CAPI2017 and Census 2016 respondents 21
Table 7 - Breakdown of gross household income responses, CAPI2017 22
Table 8 - Travel reported by main household respondent and proxy respondents in CAPI2017 and TTS2011 23
Table 9 - Trip rates and distance by purpose, main respondent-only 25
Table 10 - Detailed breakdown of trip rates by purpose, main respondents only 26
Table 11 - Demographics and reported travel of in-person and on-their-own CAPI respondents – not all household members, but only the survey taker 27
Table 12 - Completion rates for surveys online 29
Table 13 – Approximate Cost estimates for a scaled-up CAPI survey 32

List of Figures

Figure 1 - 2011 TTS PERCENT ERRORS DISTRIBUTION OF YOUNG ADULTS (18-34) BY CT IN TORONTO - Source: Chen et al. (2016) 13
Figure 2 - Percent difference in representation of 18-34 year olds, TTS2011 and census 2011 13
Figure 3 - Mode frequency matrix 33
Figure 4 - Map showing the prevalence of 'unknown' dwelling type addresses (white stars) in the City's Open Data 36
Figure 5 – Monitoring interviewer location using Find My Phone 38

1 INTRODUCTION

The Transportation Tomorrow Survey (TTS) has employed telephone and mail-invitation methods for household recruitment since its inception in 1986. Where the prevalence of collecting travel data in-person decreased significantly in Canada as a result of a perceived high cost, changes in technology may have caused the approach to become cost-effective once again, at least in certain contexts. The factors leading to a shift in relative cost-competitiveness would be i) the rising costs associated with telephone surveys resulting from the declining ability to reach landlines (because of voice over IP (VOIP), a shift to non-publicly-listed cell phones, and increased adoption of the national Do Not Call List), and; ii) the potential for use of inexpensive internet-connected tablet computers when in the field. The use of internet-connected tablets reduces processing costs and decreases the likelihood of erroneous data being recorded.

In-person survey methods may not be realistic options for the rapid collection of data from hundreds of thousands of households spread over a large area. They may be appropriate, however, for targeted data collection efforts in neighbourhoods with low landline penetration rates, and for data collection projects where the complexity of the data entry task is high and the physical presence of an interviewer may prove beneficial. The potential for face-to-face methods of data collection are in fact discussed in section 10.2 of the 2011 TTS Conduct report - Recommendations for 2016, A Feasible Approach (Data Management Group, 2014, p. 75).

In an effort to quantify the cost-effectiveness of in-person data collection, a door-to-door survey[1] was executed in the summer of 2016. The aims of the field test were to i) quantify the costs associated with carrying out research in-person using a small team or recruiters equipped with tablets and ii) assess the quality of responses provided to see if in-person survey methods can help reduce the prevalence and severity of reporting errors.

This report first briefly discusses the advantages and disadvantages of in-person interviews, with a focus on household travel diaries. The data collection effort is then described, with analysis of the collected data following. Based on the cost-effectiveness of the effort, as well as the impact on demographic representativeness and travel recorded, the potential for CAPI moving forward is then discussed. Finally, also presented are an estimate of costs and implications for scaling up, problems faced in summer 2016 and solutions to address these going forward.

[1] For the purposes of this report, door-to-door survey, computer-assisted personal interview (CAPI) and face-to-face will all be treated as synonyms of personally administered in-home interviews.

2 ADVANTAGES AND DISADVANTAGES OF CAPI

There is a significant body of work comparing both the cost and data implications of choosing face-to-face, web or telephone survey methods. Within the travel survey field, much of this research emerges from Europe and Australia, where face-to-face surveys are still employed (Duffy & Smith, 2005; Bayart & Bonnel, 2012) – this in contrast to North America where there are very few cases of face-to-face methods employed since the 1970s (TRB Travel Survey Methods Committee, n.d.).

The Travel Survey Manual, published by the Transportation Research Board's Travel Survey Methods Committee (ABJ40), succinctly presents the main advantages and disadvantages of Personal interviews – see Table 1.

Table 1 - Main advantages and disadvantages of personal interviews, as per TravelSurveyManual.org (Chapter 3), reproduced

Advantages	Disadvantages
Probably the most effective way of enlisting respondent cooperation	Likely to cost more than the other alternatives
Interviewer can answer respondent questions and probe, if necessary	Likely to be the most labour-intensive method
Interviewer can administer a reasonably complex instrument, with special sequencing; skip patterns, and difficult instructions; particularly if CAPI used	Requires a trained staff of interviewers that is geographically near the sample
Visual cues and aids can be used	Fieldwork is likely to take longer than with the telephone survey method
Can easily combine a self-administered section of the survey	Method is the most susceptible to disruptions and crime problems
The best method for developing a rapport with respondents and to build respondent confidence	
When conducted in homes, long and very detailed interviews are possible	

2.1 Advantages of CAPI and face-to-face methods

The conclusions which emerge from the literature are that for lengthier surveys, CAPI or face-to-face methods have benefits with respect to maintaining the interest of the respondent, clarifying questions, as well as monitoring whether responses are being provided without concern for the veracity of the response (quality assurance) (Doyle, 2005). This leads to a greater share of travel being reported, which in turn means a reduced concern over trip under-reporting (Stopher & Greaves, 2007).

Face-to-face methods also allow for graphics or visual aids to be employed (Doyle, 2005). Visual aids are possible in printed and web surveys as well, but the advantage in a CAPI context is that technological literacy is not required. Even if the respondent is unfamiliar with the technology employed to display information, the interviewer can help them manipulate the information, which in

turn allows for much more flexible uses of visuals such as dynamic representations, pin drop, and pinch and zoom maps.

Interviewers in a CAPI context can also clarify what is meant by a given question, in real time, as well as gauge whether or not a respondent is answering a question correctly. Incorrect answers may be provided by respondents because questions are genuinely misunderstood, or willfully to save time. Face-to-face interviewers can pick up on non-verbal cues that indicate something has been misunderstood and help avoid erroneous data being recorded. An in-person interviewer can also more closely observe whether a respondent is concentrating on the task, or is not paying attention and multi-tasking (Holbrook, et al., 2003).

Face-to-face methods can also be employed to build legitimacy and trust: "Contact by a stranger over the telephone always involves a degree of uncertainty, so people who are most socially vulnerable because of a lack of power or resources may feel they have the most to lose by taking the risk of answering and may, therefore, be reluctant to participate in telephone interviews" (Holbrook, et al., 2003, p. 94). When face-to-face with an interviewer who can explain the purpose of the project, there is an increased likelihood that a potential respondent will choose to participate.

In Grenoble, face-to-face surveys have also been found to collect more thorough information on short trips, trips made on the route (complex trip chains), as well as more trips overall. Part of this increase in trip rates comes from the ability to "motivate […] respondents to devote effort to the cognitive processing required for generating optimal answers", reducing the effect of weak satisficing (Holbrook, et al., 2003, p. 83). Weak satisficing, in the context of a lengthy travel survey, could take the form of reporting only the most basic trips (commuting) and disregarding others to more quickly finish the survey. This stands in contrast, however, to examples from the Netherlands, where the opposite has been found (Bonnel, 2001).

With telephone and in-person methods, it is also easier to control the survey day. This is very important in the context of a single-day household travel survey where a balance is sought between the travel reported on each day of the week (Bonnel, 2001). 'Announce in advance' can be used to try and get a more even distribution of travel reported on each day of the week, but this comes at the cost of potentially losing some respondents as a result of the multi-stage recruitment process. In contrast, telephone and in-person methods make it possible to ask about 'yesterday's' travel while reducing the potential for a bias on the user's side with respect to cherry picking what day to report – they were asked about their travel when reached, not at their leisure.

2.2 Disadvantages of CAPI and face-to-face methods

Moving on to disadvantages, the most oft-cited disadvantage of CAPI and face-to-face surveys is that they have been found to be considerably more expensive than other survey methods. This is not uniquely a recent conclusion, but something which has been stated since the 1970's, when telephone surveys became low enough in cost and mail-back surveys still elicited high enough response rates to make in-person surveys unattractive (Siemiatycki, 1979). The length of a survey, as well as the location and time period, however, can have a tremendous influence on costs, and as such exact figures will not be compared.

One of the reasons CAPI and other in-person data collection methods are so expensive is the costs involved in multiple visits to randomly selected address points (Bonnel, 2001). The particular way in which contact is first established and follow-ups carried out plays a significant role in determining the extent to which in-person methods become costly; the number of visits, calls or letters mailed, as well as the potential inclusion of scheduled surveys.

Multiple visits to a given location, as with multiple mailings or telephone calls, are carried out not only to increase response rates but also because not all potential respondents are equally likely to be at home when an interviewer visits. The difficulty in reaching highly mobile individuals is common across CATI and CAPI methods and may potentially make these data less statistically representative (Bayart & Bonnel, 2012). If this were to be ignored, trip rates may be under-estimated, less mobile individuals being more often reached (Stopher & Greaves, 2007). Some surveys, like the summer work presented in this report, make use of different methods of data collection at different stages to reduce this potential bias – for example, the telephone being complemented by mail and in-person complemented by mail or phone as well (Siemiatycki, 1979).

Ultimately, the problem of one or more physical visits with CAPI and in-person surveys is that they involve a large amount of costly travel. In an effort to have as random a subset of the population in a region like the Greater Golden Horseshoe (GGH), a researcher may choose to randomly pick a few thousand addresses and mail surveys to those households. Such an approach with simple random sampling or stratified random sampling is, however, not compatible with in-person survey methods. To allow for randomness while selecting households, a cluster sampling approach is the closest thing to random sampling that can be carried out using in-person methods (Holbrook, et al., 2003). A few locations can be drawn at random, with surrounding address points (every address point or every X address point) then being selected for interviews.

Moving away from cost and travel logistics, another important consideration with in-person survey methods is the potential for introducing interviewer bias (Doyle, 2005). In the realms of mail-back and web surveys, there is no interaction with a human being, and as such there can be no interviewer bias. With telephone surveys and in-person data collection methods, however, the interviewer may inadvertently affect the responses provided. This may be because the respondent wishes the approval of the interviewer (social desirability bias) or because, in rephrasing questions, the interviewer may introduce new or incorrect pieces of information that may shift the perspective of the respondent regarding how to answer. In the context of a travel survey related to the retrospective recall of travel, this should not be a concern but should be kept in mind. This is often why market research and other firms conducting telephone surveys will require their interviewers to stick to a very rigid script.

Another negative aspect of face-to-face or CAPI work is that it is more difficult to obtain accurate information from respondents on topics deemed sensitive in nature or where respondents are being asked to self-report “undesirable behaviour”; this, as with social desirability, is more often discussed in the social sciences or public health fields (Siemiatycki, 1979) (Duffy & Smith, 2005). As travel diaries like the current TTS do not ask questions whose responses could be interpreted as socially desirable or not, this problem is not believed to be of concern.

Finally, while face-to-face methods may lead to higher response rates, they are not necessarily *preferred* by respondents. Many respondents prefer mail-back to telephone and in-person surveys, as this allows them to answer the survey at their leisure (Bonnel, 2001). While important to keep in mind,

it does not, however, change the fact that mail-back surveys have a low overall response rate in Canada. Hybrid approaches, like the one used in this work, can alleviate issues related to response preferences by providing options – mail, web or telephone as alternatives to answering in-person, for example.

Overall, there is no clear consensus on whether CAPI, CATI, CAWI or mail-back lead to the best possible data being recorded per dollar spent. Just as importantly, whether CAPI *can* be carried out cost-effectively with modern methods in Canada is unclear and is the focus of this study.

3 DATA COLLECTION

Data were collected from June 18 to August 20 in-person, with the web survey remaining available online for a month thereafter. There were 985 completed household surveys collected from residents in 7 different census tracts.

This section explains the process by which interviewers were selected, followed by the selection of census tracts for data collection, a description of the initial and follow-up contact methods and schedule, a description of interviewer incentive structure employed, and finally a brief description of the data points collected in the different versions of the survey.

3.1 Interviewer selection

Getting people to answer an up to 15-minute survey at their home is not an easy task. Because the strength of the interviewers was of such critical importance, over 20 in-depth interviews were carried out before selecting which persons to offer jobs to. Experience in customer-side sales, door-to-door work (sales or political canvassing), market research or surveys, and finally transportation planning or engineering experience were sought.

4 applicants were offered a job, 3 were hired and 1 quit soon after beginning work. The remaining 2 interviewers stayed on until the end of data collection.

3.2 Sampling strategy

In order to determine which households would be contacted, it was decided that we would focus on census tracts where the previous TTS (2011) had been lacking in its accurate representation of the demographics of residents. One group in particular which we chose to focus on were 18-34 year olds. This group was considerably under-represented in the 2011 TTS. This was an issue both in the City of Toronto, as well as more broadly throughout the Greater Golden Horseshoe.

Figure 1 shows the differences between the percent of individuals in the 18-34 age bracket interviewed in the 2011 and the percent of individuals within the same age bracket, as per the 2011 census (Chen, et al., 2016). While the entire area may seem in need of improved data collection methods to reach this particular demographic, it was decided that we would try and focus on the areas where TTS 2011 did most poorly. As the 2016 TTS and census data were unavailable at the time of survey site selection, an assumption was made that the lackluster representation in 2011 would continue in 2016. Selection of tracts was done using the numbers from the 'Sampling Frame for Household Travel Surveys' report by Chen et al. (2016).

Figure 2 shows a more detailed breakdown of the issue of low demographic representativeness. Tracts in green represent those where the issue is least pronounced. Tracts in yellow or red are areas where the issue is more pronounced.

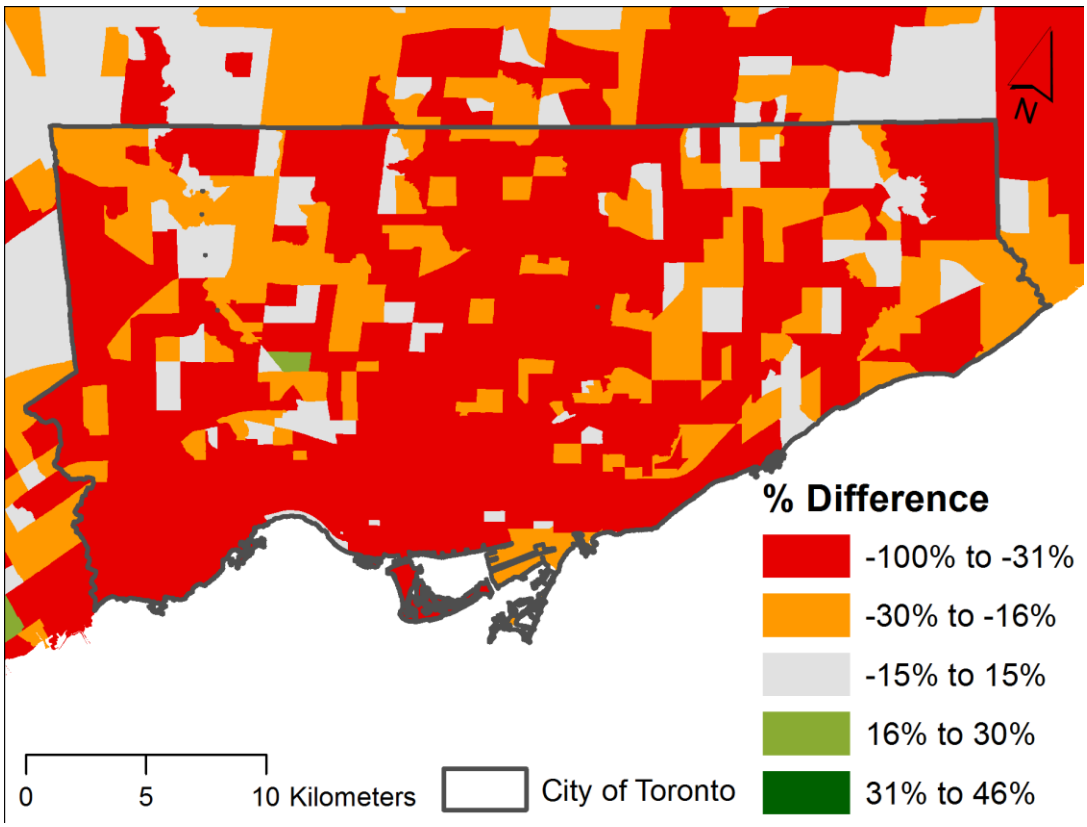

FIGURE 1 - 2011 TTS PERCENT ERRORS DISTRIBUTION OF YOUNG ADULTS (18-34) BY CT IN TORONTO - SOURCE: CHEN ET AL. (2016)

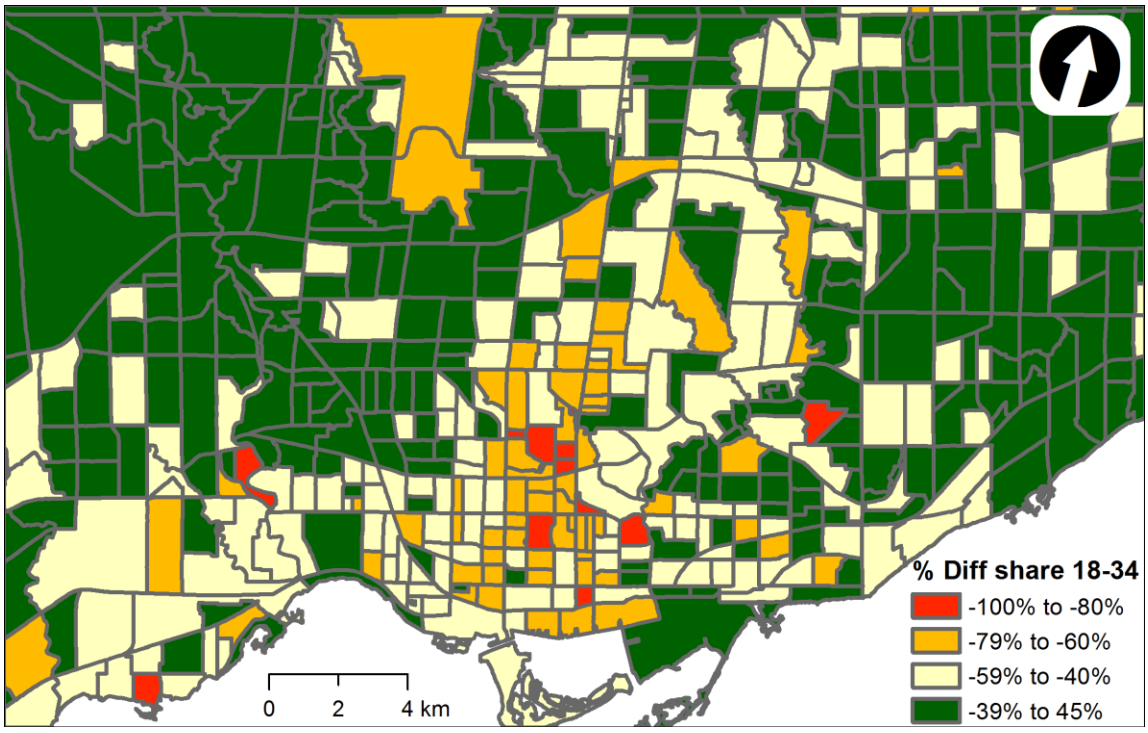

FIGURE 2 - PERCENT DIFFERENCE IN REPRESENTATION OF 18-34 YEAR OLDS, TTS2011 AND CENSUS 2011

To simplify the data collection process, it was determined that tracts near subway stations or easily accessible via TTC should be the first case studies. As the interviewers being hired were students without cars and weekday shifts were to be only 4 hours in length, there was a conscious effort made to ensure transit proximity was high to reduce commuting time.

It was then determined that tracts with a high proportion of 'Low-Density Residential' address points would make ideal test locations. These are locations where getting access to both a mailbox and doorbell for a majority of residents are not complicated by condo or apartment tower access. With regard to the street grid layout, tracts with more gridded patterns were chosen to simplify travel through the tracts as these would also contain fewer dead ends.

Finally, to make analysis simpler, we decided to focus on tracts whose boundaries aligned most closely with TTS traffic analysis zones (TAZs). This allowed for comparison of TAZ reported travel and demographics with CT reported travel and demographics.

3.3 Initial and follow-up contact

> *"Sending a letter prior to the survey is essential in order to inform subjects and motivate them to take part in the survey (all studies show that this has a quite marked effect on response rates."* (Bonnel, 2001, p. 8)

Once addresses were selected, invitation letters were delivered to the homes to be visited – these letters would serve as an advance notice, but did not contain a survey code. Delivery of the initial invitations was carried out anywhere from 1 to 15 days before visiting the address for a follow-up. More even 5-12 days beforehand would have been preferred, but was difficult to ensure – this is explored in section 5.2.

After this notice was left, prospective respondents were then approached at their homes. This was done between 4:30 and 8:30 PM Monday through Friday, and 11 AM to 7 PM Saturday and Sunday. If no one answered the door, or the person who answered the door indicated that they did not want to answer the survey at the present moment, the interviewers would leave the second letter with the survey code in the mailbox or hand the letter to the resident.

3.4 Presentation

The interviewers, when out in the field, presented themselves as carrying out a data collection project in association with the University of Toronto – which comes with a certainly implied seal of quality – and wore UofT t-shirts to convey this. These shirts could easily have been replaced by shirts and badges for the City or Ministry of Transportation, but what's important to draw attention to is that this signalling was an important part of getting people to open the door and let the interviewers in.

If outsourced to a firm, instead of handled by students, it is our opinion that there should be some effort made to visibly align the work with something residents can feel good about.

3.5 Incentive structure and survey design

The incentive structure employed to motivate interviewers was $5 per completed household survey, where demographic and mobility tool information was collected on all household members, but only

one person was asked to provide a travel diary. For additional members of a given household, there was a shorter version which only included the diary. Interviewers were incentivized to push for additional household members to answer this survey through a $2.50 incentive per additional travel diary filled.

Near the project's end, a test was run where the main survey was replaced by a more straightforward replication of the TTS. In this version, all members of the household above the age of 11 were asked to provide a travel diary. The incentives were combined in this case, with a household survey being worth $5 plus $2.50 for every additional household member above 11 reporting travel on the survey day.

These $5 and $2.50 performance incentives were applied whether the interviewers convinced someone to answer a survey face-to-face or spoke to the resident and convinced them to answer later on their own. This approach was chosen to ensure that interviewers would spend the time required to properly explain to residents how they can fill out the survey on their own if the current moment was not ideal.

The logic in testing the single-travel diary version was to see what proportion of households could be convinced to have all members fill out diaries if given the option to do so, as opposed to being explicitly required to do so. Investigation of the potential for individual travel surveys, as opposed to household travel surveys, was also something we wanted to explore. Finally, the effect on completion rates of collecting one diary instead of multiple diaries was also something we wished to explore.

4 ASSESSMENT APPROACH

The following section briefly describes the dimensions of the survey effort that were looked at to assess its strengths and weaknesses. No modeling work was carried out using the data. Instead, simple summary statistics are employed to compare TTS and CAPI data.

It was originally planned that three interviewers be kept, enabling some form of comparison of performance, but given that only two were kept on, in the end, there is no value in trying to associate their attributes to performance. Completion rates for interviews were within 1% one from the other, while trip rates recorded by each were also essentially the same (difference of 0.12 trips per respondent before accounting for any demographic bias in respondents).

4.1 Cost-effectiveness of door-to-door CAPI work

In order to measure the cost-effectiveness of the effort, detailed logs were kept of which interviewer visited each address, on what day, as well as whether a resident was spoken to, a household survey was completed and finally how many 'additional household member' surveys were completed.

4.2 Demographics

To see if door-to-door work addresses any issues of demographic bias of survey respondents, the demographics of the 2011 census, 2011 TTS, this field test and the 2016 census are compared (results in section 5.3). As census tracts were not redrawn between the 2011 and 2016 census, demographics can also be compared across years.

4.3 Reported travel

As a proxy for determining whether travel was reported accurately and completely, trip rates and distances traveled, by mode and overall, are looked at in section 5.4. Trip underreporting is a common problem in travel surveys, made all the more common in denser, more urban-type environments (Harding, et al., accepted). Trip rates found to be comparable or higher in the CAPI survey when compared to the 2011 TTS, as well as fewer individuals reporting no travel being made on the survey day, would be indications of travel being reported more carefully.

5 RESULTS AND DISCUSSION

Examining key statistics, the survey led to a completion rate (combined in-person and online with survey code) of 16.1%. This is very close to the overall completion rate of the 2016 TTS, where a mix of mail-only, telephone and random digit dialing led to a 17% completion rate – random digit dialing playing a negligible role in the overall completion statistics (Malatest, 2017). 2.6 household surveys were completed per labour hour, while 15.9 homes were visited per recruiter per hour.

Comparing the travel reported in the 7 tracts in CAPI 2017 and TTS 2011, trip rates were higher in the CAPI experiment. This is interesting given that summer trip rates have been found to be slightly lower overall, with particular decreases in the number of trips taken on transit and by personal vehicle (30% and 10%, respectively) (Cerdá, 2014, p. 16). The percent of individuals reporting no travel episodes on the survey day was also lower in CAPI when compared with the 2011 TTS.

Finally, there were only 2 or 3 instances of interviews begun with respondents and then abandoned when carried out face-to-face. Given the length of the survey, such a low number of incompletes is an important finding.

A few points to clarify before delving further into the specifics of the quantified results. There were 27 surveys carried out by an interviewer who stayed on the team only 4 shifts. The total number of labour hours calculated for the project, as well as the overall number of completed surveys, includes their work. A handful of surveys were also completed by the head RA. Accounting for 14 cases where surveys were indicated to be completed in person but did not appear on the server, there were 985 completed surveys where records were complete.

5.1 Costs and cost-effectiveness

The CAPI effort proved to be surprisingly cost-effective, with completed surveys coming in at a cost of approximately $13 apiece, materials and labour included, but ignoring the head RA's time. Best estimates of cost per completed survey with the head RA's labour paid for (included at a rate of pay similar to that of 'team leaders' working on the TTS and obtained from the TTS Conduct document (p. 37)), would be $24-25. This is the same cost as the 2011 TTS when converted to 2017 dollars.

The largest non-labour expenses incurred are listed below, wages and incentives having been explained in section 3.5.

TABLE 2 - CAPI 2017 NON-LABOUR COSTS

Website	$30
Tablets	$264
SIM cards	$24
Data plans	$276
Transportation	$609
Mailers	$728
Raffles	$300

Despite having to buy tablets for the project, the largest costs remain transportation to and from survey locations (tokens given to interviewers) and invitations (paper, envelopes, and toner). In 2017, printing, folding and delivery were handled by the graduate research assistant, but in future roll-outs, should be assigned tasks. Paying someone to fold and place letters in envelopes would add an estimated $650 to the mailer costs.

5.2 Effect of delay and delivery day on completion rates

Table 3 presents summaries generated from the logs of the delivery date of initial mailers and follow-up visits by interviewers. By 'delay' (first column), we mean the number of days between mailbox delivery of the initial invitation and follow-up contact (ringing on the doorbell and speaking to the resident or leaving the survey code and instructions).

TABLE 3 - EFFECT OF THE NUMBER OF DAYS (DELAY) BETWEEN INITIAL INVITATION LETTER DELIVERY AND FOLLOW-UP VISIT BY INTERVIEWERS, LEAVING MAILERS WITH CODE IF NO ONE ANSWERS OR IF NO ONE WANTS TO ANSWER THE SURVEY AT THAT MOMENT

	% Complete				
Delay Days	**Interviewer C**	**Interviewer M**	**Not spoken**	**Average**	**Addressed visited**
0	*50%*	*33%*	*11%*	*24%*	*46*
1	33%	46%	4%	12%	485
2	40%	43%	5%	14%	496
3	57%	50%	5%	16%	556
4	50%	44%	6%	16%	613
5	52%	59%	6%	18%	633
6	61%	51%	3%	18%	748
7	47%	42%	4%	13%	520
8	52%	49%	4%	15%	565
9	57%	58%	4%	18%	182
10	50%	38%	7%	17%	181
11	73%	43%	8%	17%	256
12	52%	48%	7%	18%	184
13	56%	53%	5%	16%	268
14	53%	35%	4%	13%	189
15	64%		4%	22%	37
Total	**52%**	**48%**	**5%**	**16%**	**5959**

Note: the "0" delay days row should be ignored, as collecting data the same day as initial invitations were sent was an error. This only occurred for a half shift on the first day of data collection, with only 46 addresses visited and the head RA along for the entirety of the shift. Also, the reason the sum of addresses does not add to 6,100 is that the third interviewer's addresses visited are not shown.

Columns 'Interviewer C' and 'Interviewer M' show the percent of household surveys completed by our two interviewers if they had the chance to speak to a resident. The next column ('Not spoken') indicates the percent of households who filled out the survey even if they did not speak to one of our

interviewers. An average is also presented, combining the two interviewers and the addresses where no one answered the door.

Our interviewers were told by many respondents that they had found and read the initial letter in the mailbox and would not have answered the survey without it. What can be read from the table is that below 3 days of delay, households may not have had the time to look through their mail and read the letter, leading to lower completion rates. Whether 3, 5, 7 or 10 days go by between initial invitation being dropped and follow-up contact, however, completion rates differ little. This is an interesting finding, in that it implies there is no hard and fast rule on the number of days that can be allowed to go by, which in turn makes delivery of invitations and scheduling of interviewers more flexible.

Next, approximately half the residents our interviewers spoke with completed an interview, whether on the spot or on the web after the fact using the survey code provided. If on the other hand, interviewers did not speak to anyone in the household and merely left the second mailer behind with a unique survey code and URL (tts2.ca/inperson), then the completion rate falls to approximately 5%, a tenfold decrease. While this is a sizeable difference, it must also be kept in mind that answering the door is a choice, not an obligation. As such, it may well be that persons who are not interested in speaking to strangers at the door are also less likely to take the time to answer a survey online.

Table 4, a complement to Table 3, shows which proportion of the overall completed surveys were carried out in-person (68%, on average), versus the percent carried out online, either after having spoken to one of the interviewers (7.9%) or simply after having received the second invitation with the survey code (23.6%). The numbers do not vary in any systematic manner in relation to the delay between initial and follow-up contact.

Finally, Table 5 shows the completion rates based on the different days of the week on which the 2nd contact occurred (interaction at door or survey code left behind). While Thursday and Friday lead to slightly lower overall completion rates (15% compared to an overall average of 16.1%), the differences are smaller than originally anticipated. This is also a reassuring finding that implies shifts can be distributed rather uniformly, as data collection does not only work on the weekend or on particular weekdays.

TABLE 4 - PERCENT COMPLETES BY SOURCE AND DELAY BETWEEN INITIAL AND FOLLOW-UP CONTACT

			Overall responses			
Delay Days	**Addressed visited**	**Complete in-person**	**Complete total**	**% In-person**	**% After spoken**	**Not spoken**
0	*46*	*7*	*11*	*64%*	*9.1%*	*27.3%*
1	485	37	58	64%	6.9%	29.3%
2	496	48	70	69%	4.3%	27.1%
3	556	57	88	65%	10.2%	25.0%
4	613	69	101	68%	5.0%	26.7%
5	633	80	117	68%	7.7%	23.9%
6	748	97	133	73%	13.5%	12.8%
7	520	47	69	68%	7.2%	24.6%
8	565	61	86	71%	7.0%	20.9%
9	182	24	33	73%	12.1%	15.2%
10	181	18	30	60%	10.0%	30.0%
11	256	26	44	59%	4.5%	36.4%
12	184	23	33	70%	3.0%	27.3%
13	268	30	44	68%	9.1%	22.7%
14	189	18	24	75%	0.0%	25.0%
15	37	6	8	75%	12.5%	12.5%
Total	**5959**	**648**	**949**	**68%**	**7.9%**	**23.6%**

TABLE 5 - DAY OF 2ND CONTACT MADE AND COMPLETION RATES

Second contact day	**In-person completes**	**In-person + Online completes**	**Letters delivered**	**In-person completion rate**	**Total completion rate**
Monday	72	104	617	11.7%	16.9%
Tuesday	148	203	1,238	12.0%	16.4%
Wednesday*	*29*	*44*	*234*	*12.4%*	*18.8%*
Thursday	153	220	1,432	10.7%	15.4%
Friday	114	162	1,090	10.5%	14.9%
Saturday	60	93	555	10.8%	16.8%
Sunday	107	156	927	11.5%	16.8%
Weekday	516	733	4,611	11.2%	15.9%
Weekend	107	156	927	11.5%	16.8%
Total	683	982	6,093	11.2%	16.1%

** The lower number of 2nd letters/contacts made on Wednesdays are an anomaly, attributed to weekly project meetings, with shifts scheduled less frequently as a result.*

5.3 Respondent demographics

In order to ensure apple to apple comparisons, in both sections 5.3 and 5.4, when comparing respondent demographics or reported travel across survey efforts, data presentation and analysis is limited to the 7 visited tracts, whether CAPI, TTS or the 2011 and 2016 census.

TABLE 6 - DEMOGRAPHICS OF CENSUS 2011, TTS2011, CAPI2017 AND CENSUS 2016 RESPONDENTS.

Source	Census2011	TTS2011	CAPI2017	Census2016
Households	11,550	585	975	11,499
Av household size	2.43	2.46	2.56	2.44
1-person households	31%	23%	17%	31%
2-person households	31%	39%	41%	31%
3-person households	15%	15%	19%	15%
4-person households	23%	23%	22%	23%
Mean age	40.6	45.1	42.7	40.2
% males	48%	50%	47%	48%
% population aged 0 to 14	17%	18%	15%	16%
% population aged 15 to 19	5%	5%	5%	5%
% population aged 20 to 29	13%	4%	8%	14%
% population aged 30 to 39	16%	7%	10%	15%
% population aged 40 to 49	17%	18%	18%	15%
% population aged 50 to 59	13%	17%	17%	14%
% population aged 60 to 69	10%	16%	15%	11%
% population aged 70 +	9%	16%	10%	10%
Full time employed		35%	44%	
Av number vehicles/household		1.22	1.27	
*% Houses**	47%	71%	97%	53%
% Townhouses	5%	4%	1%	5%
Average gross income	$184,618		*High*	$203,890

* Includes single family, semi-detached and other single-attached houses

Table 6 shows that the average household size is slightly larger in the 2017 CAPI dataset than in the TTS or census. This difference in average household size is explained by the sample frame used and its difference from the true census tract housing mix – see % Houses and % Townhouses rows. While the response 'House' was chosen by 97% of respondents (the address points visited were labeled as low-density residential in the City's Open Data address list), only 53% of the units in the same area are 'Houses' according to the census (including single family, semi-detached and other attached houses). As apartment residents tend to live in smaller households than households living in single-family and other houses, it is thus normal to have collected information on larger households. This is also reflected in the smaller share of single person households.

Examining the remaining list of variables of interest in Table 6, the mean age among CAPI respondents is closer to the census than was the 2011 TTS. The representation of population groups 20-39 also more closely resembles the census, as does that of seniors, which is a positive finding given

that oversampling of older individuals is one of the problems the TTS typically faces. All this indicates that, at least for these 7 tracts with a majority of low-density residential homes, a CAPI approach appears to more closely reflect the true population age distribution.

The percent of respondents who are full time employed is also higher in the CAPI, which correlates with the lower number of retirees interviewed. Census information on the percent of residents who are full time employed is not available for comparison in the same format, however. The number of cars per household is slightly higher in CAPI than TTS as well; given the prevalence of larger households, living in houses, this is to be expected.

A final indicator of the type of household reached is reported income. The census average for gross (pre-tax) household income in the tracts visited is nearly $204K according to the 2016 census. While it is not possible to calculate an average income from the responses provided in CAPI – response alternatives being categorical and capped at "$125,000 or more"-, what we see is that 61% of those households who reported an income, reported bringing in over $125K – see Table 7.

TABLE 7 - BREAKDOWN OF GROSS HOUSEHOLD INCOME RESPONSES, CAPI2017

Income category	Percent	Percent (reported)
$15,000 to $39,999	2.8%	4.6%
$40,000 to $59,999	4.0%	6.6%
$60,000 to $99,999	9.2%	15.2%
$100,000 to $124,999	7.6%	12.6%
$125,000 and above	36.9%	61.0%
I decline to answer/ I don't know	39.5%	N/A

Comparison of TTS and CAPI aside, the prevalence of high earning individuals in these tracts indicates that the response alternatives, while based on census categorization, may not be appropriate for the study region. From a modeling perspective, there is an expectation that members of households making $225,000 will exhibit significantly different travel behaviour and have different values of time that households making $125,000. If there are large enough shares of such high earning households, it may be worth asking for income in a manner which better allows for household categorization.

5.4 Reported travel[2]

The following section examines the reported travel of respondents to the CAPI main (one travel diary) and TTS-replication (diaries for all persons over the age of 11) surveys, and contrasts these with the 2011 TTS; 2016 disaggregate data were not available to the researchers at the time of publication. As there is a long history of trip under-reporting related to proxy responses, the differences highlighted the focus on primary main CAPI and TTS respondents (Proxy = No).

TABLE 8 - TRAVEL REPORTED BY MAIN HOUSEHOLD RESPONDENT AND PROXY RESPONDENTS IN CAPI2017 AND TTS2011

Proxy	No		Yes	
Source	**CAPI2017**	**TTS2011**	**CAPI2017**	**TTS2011**
Respondents	824	584	45	679
Mean age	52	57	47	45
Nb Trips reported⁺	2,464	1,513	101	1,396
Trip Rate*	2.99	2.59	2.24	2.06
Sum Dist (km)	12,802	8,199	667	7,235
Av. Dist. Traveled (km)	15.5	14.0	14.8	10.7
Av.Trip Dist (m)**	5,196	5,419	6,602	5,183
Av.Motorized trip dist (m)	7,688	5,881	11,456	6,161
Av.Transit trip dist (m)	6,435	5,712	5,866	5,652
Av.Bike Dist (m)	2,873	2,234	2,365	2,089
Av.Walk Dist (m)	1,305	1,221	873	922
Trip percent Motorized	49%	68%	46%	59%
Trip percent Transit	13%	22%	16%	24%
Trip percent Bike	9%	5%	8%	5%
Trip percent Walk	30%	6%	31%	13%
Distance percent Motorized	72%	74%	79%	70%
Distance percent Transit	16%	23%	14%	27%
Distance percent Bike	5%	2%	3%	2%
Distance percent Walk	8%	2%	4%	2%

⁺ ***'Other' mode trips removed***

**** Includes non-traveling individuals***

***** Trip distances over 25 km counted as 25 km max***

The differences in reported travel presented in Table 8 are notable and informative. One important point to keep in mind, however, is that the travel was recorded during the summer as opposed to the

[2] Please note that this is not an apple-to-apple comparison. The TTS is a fall survey, but the CAPI was a summer survey

fall. This is highlighted throughout the section, where relevant. Also, to reiterate, persons who responded to the survey directly are distinguished from proxy respondents in both the CAPI and TTS. This is an important distinction, as trip under-reporting is considerably higher among proxy respondents.

To begin, both the average trip rate and average distance traveled (all modes summed) are higher among CAPI respondents – 2.99 to 2.59 trips and 15.5 to 14.0 kilometers per day, respectively. Average distances traveled per trip using each mode (motorized, transit, bike, walk) are also greater, with only the *overall* average trip distance decreasing slightly. This is in part due to a greater share of non-motorized trips being reported (39% instead of 11%), pulling down the average trip distance.

One particularly interesting finding relative to trip distances is that average walk trip distance increases in the CAPI survey. The CAPI interviewers were explicitly instructed to probe for walks and discretionary travel in an effort to reduce under-reporting, but the hypothesis beforehand was that this probing would lead to greater number of walk trips with a lower average distance reported. The issue of walk trip under-reporting in the TTS is typically assumed to lead to fewer short, discretionary trips not being reported, so the increase in average distance traveled per walk trip is notable. Seasonality may well play a role here, as the threshold of acceptable distance for travel on foot may well increase in more pleasant weather.

Getting back to the average distance traveled per person increasing from 14.0 to 15.5 km, it should be specified this increase is not affected by inter-city travel, maximum trip lengths being set at 25km to avoid outlier effects on the mean. As such, it appears that irrespective of there being fewer students in school and a greater number of persons on vacation, the number of trips and overall distance reported increase with the CAPI method.

Regarding the mode shares expressed as percent trips made using each mode, what we observe is a markedly lower percentage of travel being reported in private vehicles in CAPI (49% vs. 68%) and a substantial decrease in transit (13% to 22%). The latter is explained in part by commuting and school trips not being made, but the sharp decrease is greater than what one would expect given that the percent of main respondents who report being students is near 4% - see

Table 11. The shares of trips made by bike (9% vs 5%) and on foot (30% to 6%) is also greater than TTS reported travel. The low walk mode share in TTS can be explained by a methodological decision in terms of what data to probe for or record.

The areas visited are within walking distance to a subway station and, for every census tract but one, have commercial arteries with a variety of shops, services and other amenities nearby. Our results indicate that in such environments, walk trips may well replace trips made using other modes of transportation.

Finally, the fact that the overall average distance traveled by motorized modes increased in CAPI (7,688 vs 5,881 meters) despite a greater share of overall trip-making being by active modes (walk and bike), provides evidence that the CAPI respondents are not qualitatively different from TTS respondents (environmentally inclined or car-averse, for instance). The higher-than-TTS vehicle ownership rate would also corroborate this (Table 6).

TABLE 9 - TRIP RATES AND DISTANCE BY PURPOSE, MAIN RESPONDENT-ONLY

Source	CAPI2017	TTS2011
Nb of trips*	2477	1514
Av. Time for Out-of-Home activities (hrs)	6.9	7.0
Trips Home	37%	39%
Distance Home	39%	39%
Trips Work	15%	20%
Distance Work	18%	26%
Activity duration Work	43%	68%
Trips Other	24%	20%
Distance Other	25%	20%
Activity duration Other	36%	22%
Trips Shopping	18%	12%
Distance Shopping	11%	9%
Activity duration Shopping	17%	7%
Trips Pickup/Dropoff	5%	9%
Distance Pickup/Dropoff	5%	6%
Activity duration Pickup/Dropoff	4%	2%

Trips School	1%	1%
Distance School	0%	1%
Activity duration School	1%	2%
**Unknown purpose trips removed* *Only main respondent travel*		

Looking at Table 9, which presents information on only the main respondent within a household in both the TTS and CAPI surveys, the differences found are both intuitive and encouraging. To begin, as expected for a survey carried out during the summer and where a greater emphasis was placed on prompting to collect complete travel diaries, the share of trips for work decreased markedly (20% to 15%). On the reverse, the share of trips made for Shopping and Other purposes increased (18% from 12% and 24% from 20%, respectively). This is both consistent with expectations, as well as indicative of a higher share of discretionary travel being recorded.

In addition to emphasis being placed on collecting information on discretionary trip making when face-to-face, a more detailed breakdown of the potential trip purposes for which information was sought (see Table 10) may also have encouraged reporting trips that otherwise could be forgotten.

TABLE 10 - DETAILED BREAKDOWN OF TRIP RATES BY PURPOSE, MAIN RESPONDENTS ONLY

TTS	%	%	CAPI
Home	39.3%	38.9%	Home
Work	20.0%	10.4%	Work (habitual location)
		3.9%	Work-related
Other	19.5%	9.1%	Recreation
		4.8%	Visiting friends, family
		4.4%	Restaurant, bar
		2.5%	Services
		2.0%	Health / Personal care
		0.9%	Other
		0.3%	Worship, religion
Shop	12.2%	15.1%	Shopping and Errands
		2.5%	Take-out

Facilitate Passenger	8.4%	2.6%	Drop someone off
		2.4%	Pick someone up
School	0.6%	0.4%	School / Education

5.5 Differences between 'in-person' and 'on their own' responses

Table 11 shows the differences in reported trip rates and other attributes for individuals interviewed in person or who chose to use the survey code left at their home.

TABLE 11 - DEMOGRAPHICS AND REPORTED TRAVEL OF IN-PERSON AND ON-THEIR-OWN CAPI RESPONDENTS – NOT ALL HOUSEHOLD MEMBERS, BUT ONLY THE SURVEY TAKER

Response	In-person	On their own	Total
Reported income	50%	77%	60%
High income	67%	57%	62%
Home owner	90%	90%	90%
Occupation Professional/Management/Technical	88%	87%	87%
Number of vehicles	1.32	1.22	1.28
Phone owner	67%	60%	64%
Age respondent	51	53	52
Female	48%	52%	50%
Full Time Employed	63%	53%	59%
Full time student	4%	4%	4%
Retired	22%	28%	24%
Transit pass owner	11%	11%	11%
Bikeshare member	2%	3%	2%
No travel	12%	13%	12%
Trips per day	2.83	3.25	2.98
Carshare member	9%	11%	10%
Driver's license holder	90%	93%	91%
Commute to work on foot	14%	10%	13%
Commute to work by bike	10%	16%	12%
Commute to work by local transit	31%	30%	31%
Commute to work by driving alone	38%	35%	37%
Free parking at work	27%	31%	29%
Commute to work by ridesharing	3%	5%	4%
Commute to work by regional transit	1%	0%	1%
Commute to work by other modes	4%	4%	4%
Frequency walk*	3.2	3.1	3.2
Frequency cycle	1.1	1.2	1.2
Frequency local transit	2.2	2.3	2.3
Frequency drive alone	2.2	2.4	2.3
Frequency rideshare	2.2	1.8	2.0
Frequency regional transit	0.3	0.3	0.3
			Frequency responses were 0 = Very Rarely; 1 = Rarely; 2 = On Occasion; 3 = Frequently; 4 = Very frequently

As can be read in

Table 11, summer 2017 respondents who answered the survey on their own after receiving the invitation and survey code in the mail were more likely to report their income. This is consistent with prior findings on the provision of sensitive information. They were also less likely to be high-income earners, but this result must be interpreted in the context given the higher share of retired respondents (28 to 22%).

The overall share of respondents with a landline is around 64%, which is higher than anticipated, but which also makes sense given the prevalence of larger households in the sample. There are no major differences in the share of homeowners, nor in the share of respondents reporting to work in professional, managerial and technical professions. This is unsurprising given the near ubiquity of home ownership in the tracts visited.

Moving along to employment status, a larger share of persons who responded face-to-face are employed full-time (63 to 53%). This could be interpreted as an indication that CAPI methods are more adept at collecting information from non-retired individuals or residents who might otherwise ignore a mailed survey invitation. The rate of transit pass ownership is consistent for both groups at 11%. This is slightly lower than the 13.3% of TTS 2011 non-proxy respondents in the same tracts reporting having a metropass, but further research would need to be conducted to see if the decrease from 13.3% to 11% is consistent with decreases in transit pass ownership over the summer months for non-students. Seasonality may once again explain the difference, as might the greater share of larger households.

The percent of the sample (either responding in person or online) with a Toronto BikeShare membership (2-3%) or who are carshare members (10%) are higher than anticipated and would need to be looked into further. It would be interesting to look into whether CAPI respondents were particularly progressive in their mobility tool ownership or are merely reflective of a changing and more multi-modal Toronto. Data are not available at this time to perform comparisons with locally relevant membership data.

The number of vehicles per household are rather consistent, as are age and gender balance. The share of respondents reporting no travel on the survey day is very similar (12 or 13%), with a slight increase in the number of trips per day (3.2 to 2.8) for web-survey respondents. This difference in reported trip rates can be explained in part by the fact that retired persons (who are more likely to have responded to survey using the code online) have more time on their hands and patience to respond to surveys. One would have, however, expected the trip rates to be higher when the surveys were filled out with interviewers. This may be a sign that more hands-on training and monitoring might be required for future efforts.

Finally, one of the most important differences between in-person and online surveys are completion rates. When face-to-face, only 2 or 3 interviews were begun without being completed – less than 0.5%. This stands in sharp contrast with the 79.6% completion rate for respondents who used the survey code on their own, whether after speaking to one of our interviewers (72.3%) or simply after finding the survey code in their mailbox (82.6%). See Table 12.

TABLE 12 - COMPLETION RATES FOR SURVEYS ONLINE

Spoke to resident	Surveys completed online	Survey completion rate
No	227	82.6%
Yes	86	72.3%
Total	313	79.6%

Low completion rates are a major concern with travel surveys, as the demographics and travel of households do *not* complete the survey can not be demonstrated to be similar to the demographics and travel of households who do complete the survey.

5.6 TTS versus optional additional household diaries

As mentioned in section 3.5, the decision had initially been made to collect travel diary information on one respondent per household, later *allowing for* other household members to provide their travel information using a diary-only version and unique household code. The motivation for interviewers in promoting this was that every additional household member who filled out a travel diary would get them $2.50, while the motivation on the prospective respondent side was that each diary would qualify as an entry into the monthly raffle. This approach was a failure and should not be replicated again. The follow-through was 62 additional diaries collected, 32 collected in-person with the interviewers and 30 online.

This is not to say all household members should be asked to provide diaries, but rather that collecting information on the travel of household members should be done systematically. This can take the form of collecting diaries from every household member above the age of 11 (as is currently done) or - in an effort to reduce response burden and proxy bias while increasing completion rates- can take the form of randomly selecting 1 or more persons within the household and asking that their travel be reported. The former approach is common in North America, while the latter is more in line with survey practices in other parts of the world, notably in Switzerland (Office Fédéral de la Statistique, 2012). However, that discussion, important as it may be, is peripheral to this report.

6 CONCLUSIONS

The work carried out during the summer of 2017 demonstrates that CAPI surveys can be carried out cost-effectively in specific circumstances. Response rates are on par with the 2016 TTS (16 vs 17%), with very minimal follow-up with prospective respondents. Analysis of the demographics of respondents indicates that it is possible to collect information from younger individuals using such an approach, with the limitation that the method has only been tested in low-density residential environments in the city. Finally, analysis of the trip information collected also indicates that the method leads to recorded travel of equivalent quality to the TTS.

What all this would indicate is that in-person surveys should be considered as a way to augment the core TTS data collection methods in tracts where the properties of the urban form and dwelling-type composition make this economically and logistically feasible.

The way in which the experiment was carried out may be one of the most important takeaways: if cluster sampling can be applied, the ability to make use of interviewers for both survey data collection and mail delivery can lead to data collection costs on par with telephone and mail surveys, all the while producing an equivalent quality of data. It may also be possible to optimize shift scheduling and pick-up and drop-off of interviewers, such that the approach becomes viable in less central and transit adjacent areas.

7 SCALING UP

Were in-person surveys to be scaled up to thousands of respondents and brought out of the realm of academic field tests, estimations have been made as to the likely costs involved – see In the scaled-up version, certain costs are included which were essentially absorbed by the university during the summer field test - printing and mailing costs, notably. Otherwise, the most significant item added is that of management and coordination, included as a full-time position for 18 months. The manager brought on would be in charge of hiring and managing the team of interviewers, ensuring the survey is operational at all times, as well as be in charge of finding the locations to be visited, preparing shift maps, ensuring interviewer performances are being properly monitored and preparation of a conduct report. Team-leader bonuses are also included for each team, and with post-survey processing estimated based on TTS2011 costs.

Table 13. The scaling up is not meant to be used to replace the TTS in its entirety, but rather a realistic scenario is devised whereby one project manager could oversee 4 teams of 3 interviewers, sharing 8 tablets and collecting data in two consecutive spring seasons, with a bit of data collection in fall to better differentiate seasonal from methodological biases.

While the targets set may appear too small to contribute to the TTS, it should be stated that the same way a 10% sample of household travel survey respondents are on certain projects given GPS loggers to better understand trip underreporting, a 12,000 household sample of respondents providing high-quality data can act in a similar manner, checking for differences in reporting behaviour that might be related to fatigue, as well as allowing another means by which to reach certain demographic groups. Coordination would have to be well managed to ensure there is no sample frame overlap with the main TTS, but otherwise should not be problematic.

Also, while there are good reasons why a 5% sample can be desirable, it should be noted that in a context where real-world municipal roll-outs can range from 2,500 to 10,000 households (Stopher & Greaves, 2007), a 12,000 sample is not a negligible source of data.

In the scaled-up version, certain costs are included which were essentially absorbed by the university during the summer field test - printing and mailing costs, notably. Otherwise, the most significant item added is that of management and coordination, included as a full-time position for 18 months. The manager brought on would be in charge of hiring and managing the team of interviewers, ensuring the survey is operational at all times, as well as be in charge of finding the locations to be visited, preparing shift maps, ensuring interviewer performances are being properly monitored and preparation of a conduct report. Team-leader bonuses are also included for each team, and with post-survey processing estimated based on TTS2011 costs.

TABLE 13 – APPROXIMATE COST ESTIMATES FOR A SCALED-UP CAPI SURVEY

Item/Survey	TTS2011	CAPI2017	CAPI2018
Software development and testing	$115,000	$0	$1,600
Interviewer staff, training and transportation	$1,600,000	$11,315	$135,775
Coding staff and training	$235,000	$0	$0
Hardware and software	$68,000	$0	$0
Phones / Tablets	$160,000	$540	$2,160
Printing and mailing	$225,000	$728	$16,490
Office space and furniture	$400,000	$0	$0
Sample	$35,000	$0	$0
Office expenses and supplies	$17,000	$0	$120
Management and coordination	$482,000	$11,352	$105,500
Computer support	$293,000	$0	$0
Post-survey processing	$20,000	$0	$1,480
Raffle	$0	$300	$600
Total	$3,668,000	$24,235	$263,725
Completed	159,600	984	11,808
Addresses visited		6,100	73,200
Cost per completed survey	$22.98	$24.63	$22.33
2017 dollars	$25.02	$24.63	$22.33

8 PROBLEMS FACED, ERRORS MADE AND OPTIONS TO EXPLORE FOR FUTURE DATA COLLECTION WORK

This section, very much an addendum to the body of the report, will include descriptions of survey design, field and technical issues that arose during the course of the data collection effort, along with proposed solutions.

8.1 Design

8.1.1 Response alternative issues

While the choice of response alternatives to present to respondents is beyond the scope of this report, we would like to highlight a few cases which respondents reported as being problematic or were often misunderstood.

'Seldom' versus 'never'

One type of question included in the CAPI surveys is: "How frequently do you use the following modes of transportation:", followed by a list of modes - Figure 3.

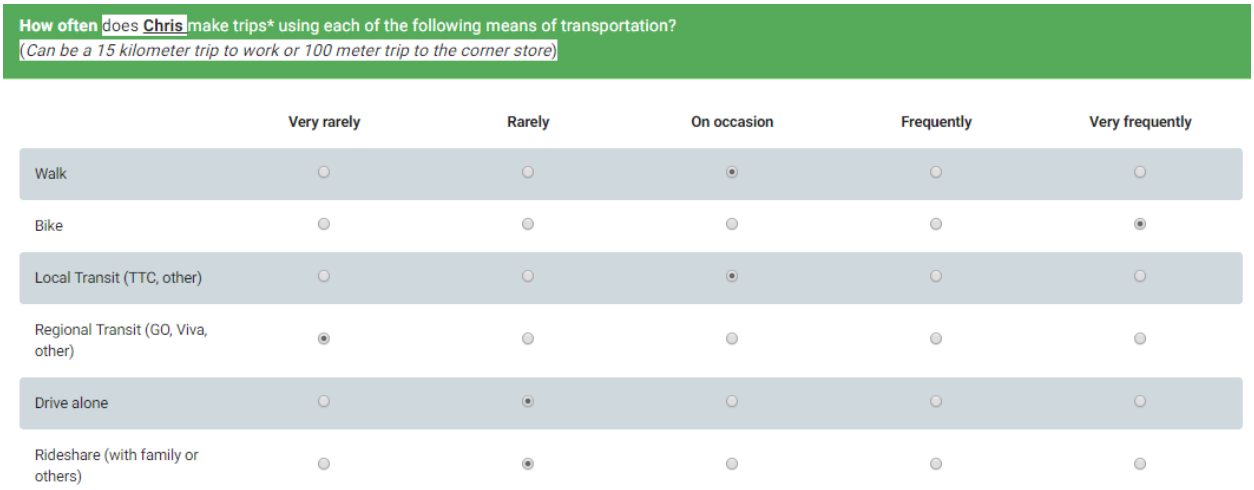

FIGURE 3 - MODE FREQUENCY MATRIX

In the matrix, a 'Never/Less than once per year' option should be included to distinguish 'Very Rarely' from 'Never'. Many people talked about this as being an annoyance. While the idea from the start was to avoid having to define what 'Never' entailed, this was one of the points brought up most frequently in terms of survey dissatisfaction.

Gender

A topic of current relevance is the issue of asking about a gender. The decision made for the in-person work was to include an 'Other' response alternative, but our interviewers were told this was 'othering'. In future, an 'Other/Non-binary identifying' or similar type of response may want to be added. The wording on such a question should be verified before any future significant rollout.

Employment status

A recent trend is that of persons working from home, as well as working multiple part-time jobs or performing contract work. These individuals may or may not have multiple habitual work locations, as well as may find it difficult to know how to respond to a question about whether they are full-time employed or part-time employed.

From a travel demand modeling perspective, the most important thing to understand is the frequency and scheduling of out of home activities, so a person working 'part-time' at 2 jobs out of the home, but totaling 45 hours a week may well be making as many trips as a person working 'full time'. Likewise, a person working part-time 'from home' and part-time 'out of home' may feel it's most appropriate to answer that they are a full-time worker with a habitual place of work, but if we record this incorrectly, our models will, in turn, incorrectly generate out-of-home travel episodes.

A proposition made by one of the researchers is to look into the implications of asking about the number of hours worked for pay at home, as well as out of the home, or to ask for the number of days per week that work is carried out out-of-home. This would allow for much more clear modeling work to be performed and trip rates to be estimated for these different groups – where workers then exist along a spectrum instead of being categorized as part or full time, which may be becoming less fruitful definitions of employment status.

Along the same line, self-employed, as a characterization of employment status, is less helpful from a travel demand modeling perspective, as it remains unclear what that entails. A person can be self-employed and work from home, or be self-employed and be on the road every day of the week. As the point is to better understand travel behaviour, not employment trends, it may be worth revising the question.

Mobility-related impairment

Our interviewers were told multiple times over the summer that there was no room in the survey to indicate mobility was impaired if no travel episodes were reported. As such, the question "Do you have a mobility impairment?" was added. The appropriate groups should be consulted to find out the best way to phrase the question, as well as the most apt response alternatives, but of those persons exposed to the question, 4.1% responded that they had some form of mobility impairment.

8.1.2 Loops

Not specified or explicitly addressed in the survey, but leisure walks, jogs, walking the dog and other activities not departing from a location and heading to another were recorded with our interviewers as recreational activities when reported by respondents, but are not taken into account in the TTS.

With the increased use of TTS-type data for epidemiological research, there is no right or wrong answer in terms of whether or not these trips *should* be recorded, but there should clearly be a discussion of how to process these trips and whether or not to include them when reporting on distances traveled and trip rates.

8.1.3 Completion time

This issue is rather straightforward, but at the launch of the survey, there was no estimated response time provided when starting the survey. Household travel surveys like the TTS do vary considerably in

length depending on the number of individuals in the household and the number of trips they report, but not providing an estimated length was highlighted by our web respondents as an annoyance. If a survey lasts 5 minutes or less, it may be possible to not present such information, but as soon as the survey lasts longer than this, it becomes problematic. Beyond irritation, respondents may have chosen to start the survey on a different device or at a different time if they had known the length would be so significant. This may then lead to a decreased completion rate.

8.1.4 Website issues

A mistake made early on that should be addressed in any future project is to ensure an SSL certificate is purchased and applied before any traffic is routed to a site. When the TTS2.CA website launched and prospective respondents were directed to it, 'https' had not yet been secured, so some prospective respondents were reticent to answer the survey. This is not because a majority of residents are tech savvy enough to know what a security certificate means, but rather that the browsers they employ send warnings when navigating to unsecured sites.

Anything that can potentially hurt the credibility of the project while aiming to collect detailed demographic and travel information is a problem.

8.2 Field

8.2.1 Apartment and condo residents

The summer 2017 data collection effort explicitly sought to collect information on residents of single-family homes in tracts where this was the predominant form of dwelling and where previous efforts to collect travel data led to considerable under-representation of certain demographic groups. This experiment having been carried out rather successfully, further efforts should be carried out to measure the potential effectiveness of the method of collecting information on apartment and condo tower residents.

8.2.2 Survey code

The way the initial and second contact letters were written, the invitation served only to introduce the project, while the follow-up contained the survey code. This was a conscious decision made to avoid people feeling they had to do anything before the interviewers would come by, and replicated the concept of a mailed pre-notification letter before a telephone or in-person contact.

While it did make it such that a majority of interviews were carried out in-person with our interviewers, which has some benefit, it also increased the labour costs involved and may have caused some households who would have been willing to answer the survey, not to follow-through.

For any future effort, we believe it is worth considering printing a survey code for the household on both the initial and follow-up letters. The difference in cost is nil and there is a potential to increase overall response rates and reduce labour costs if residents answer the survey before interviewers show up at the door.

8.2.3 Address list problems

The address list used to carry out the project was downloaded from the City's Open Data portal (City of Toronto, 2017). It is riddled with problems, however, as there are:

- addresses labeled as *unknown* instead of low-density residential, as well as;

- addresses labeled as low-density residential despite having been subdivided or being apartment buildings or commercial spaces;

- addresses that exist in duplicate.

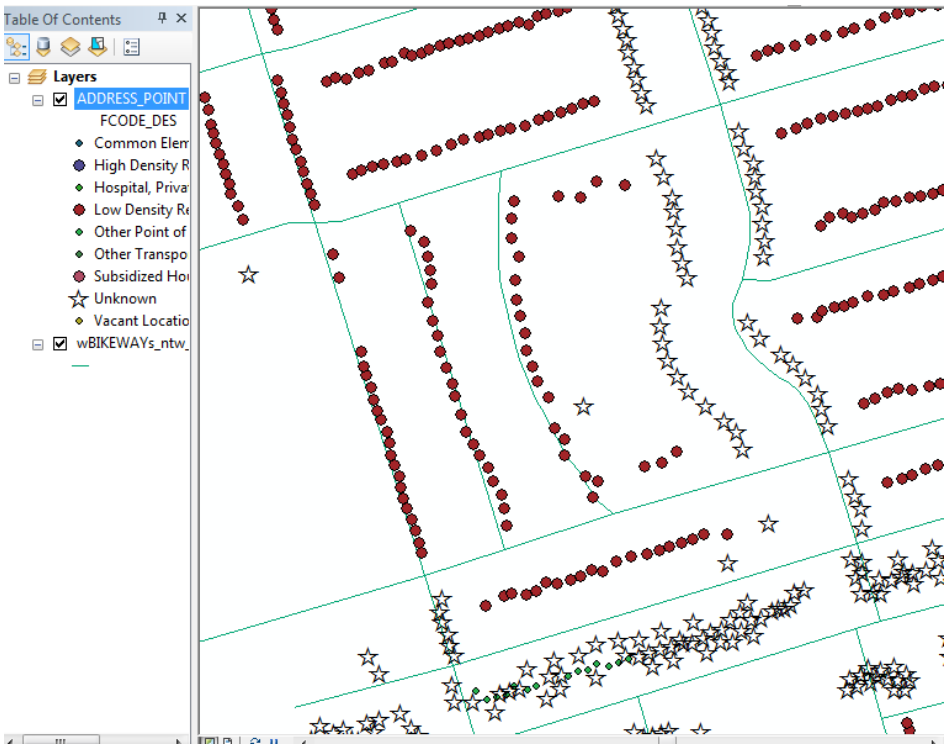

FIGURE 4 - MAP SHOWING THE PREVALENCE OF 'UNKNOWN' DWELLING TYPE ADDRESSES (WHITE STARS) IN THE CITY'S OPEN DATA

These issues both decrease the likelihood that the best tracts will be selected for in-person survey work, as well as cause problems while in the field. As all invitation and follow-up letters carry an address and a unique survey code, printing invitations without an address, but with a survey code is possible, but far from ideal. It might be possible to deliver the initial invitations with a clipboard or other device at the ready to mark down new addresses, as well as mark which survey code is associated with the new invitation dropped. This list could subsequently be used to update the project list and have everything in order by the time interviewers are dispatched.

While the approach described *can* be used to address issues where basement and other apartments have their own mailbox but don't have their own entry in the Open Data address list, there are also entire city blocks of single-family homes that are missing.

One idea that emerged during the summer might be to combine resources and help update the Open Data address list in the field while carrying out the work. As such, the costs incurred in spending the additional time at each address would not only help improve the representativeness of the data collected for the project but would also have longer-term benefits for the City or other users of the address lists.

Finally, certain types of problems encountered in the field were kept track of, but not in a complete and systematic manner. Examples of issues marked down were homes under construction and where it was clear no person was currently living, addresses where no resident spoke English, as well as addresses where the person who answered the door indicated the home was being rented out on a short term basis – e.g. Airbnb. These issues, in future, should be recorded more systematically such that their effect on completion rates can be better understood.

8.2.4 Sunlight, Safety, and Scheduling

The interviewers were asked at project end what they thought could be changed to make them feel safer, the work involving going door to door with little in terms of safety apparatus. Safety considerations have been reported as reasons why in-person interviews ceased in certain locations, and so should be taken very seriously (Bonnel, 2001) (Stopher & Greaves, 2007).

Sunlight was brought up when we were wrapping data collection in August and the sun was setting earlier and earlier. One thing that is clear both in terms of productivity, as well as safety, is that interviews can only be carried out until the end of nautical twilight, which is when the sun is between 6 and 12 degrees below the horizon. Beyond this point, it becomes too dark to see, prospective respondents are reticent to open the door to strangers and interviewers should no longer be walking around given the low visibility.

Because of daylight saving time, this means surveys can't be carried out in-person on weekdays in the fall, but starting March 12, nautical twilight is 8:22 PM, extending to past 10PM from May 22 (timeanddate.com, n.d.). With data collection beginning at 4:45 PM and running until 8:30 PM early season, and shifting back a half hour by end of the season, collecting data from the end of March to end of June would allow for data collection during the evening while ensuring a safe work environment for interviewers.

As per Stopher and Greaves (2007), March to June and September to November are both acceptable, common periods where surveys can be carried out. Some data collection c*ould* happen in September if desired, where collecting during both fall (TTS and CAPI) and spring (CAPI-only) periods can provide an idea of seasonal effects controlling for survey instrument, allowing for Spring-Fall correction factors to be derived.

While one might expect that active transportation figures might be higher in the spring than fall, biasing mode shares, temperatures in Toronto are actually lower in the spring than in the fall, with an average high of 4 degrees in March and 22 degrees in September. The fact that the sun sets later in the day may offset the colder temperatures and make it such that the two effects cancel each other out, but further research would need to be carried out looking at active transportation mode shares in a similar climate to confirm this.

8.2.5 Quality assurance

One of the concerns throughout this process was that the interviewers, who were monetarily incentivized to complete a greater number of interviews each shift, could either make interviews up or rush through interviews to maximize the amount completed in a given hour. To assess whether the latter of these two issues was a valid concern, the head RA monitored trip rates periodically, as well as tagged along on certain interviews. This helped ensure the quality of the responses received was high, as clarifications could be periodically offered and training carried out on the fly. For times when

the head RA was not along for interviews, however, there was always the possibility that interviewers were not performing their duties exactly as desired.

While the head RA knew the location of interviewers, as the interviewers and RA would meet at the beginning and end of each shift at the location where data were being collected, as well as could be monitored by querying the 'Find my device'* feature of the tablets (see Figure 5), the content of the conversations was not monitored in the same way as a call center.

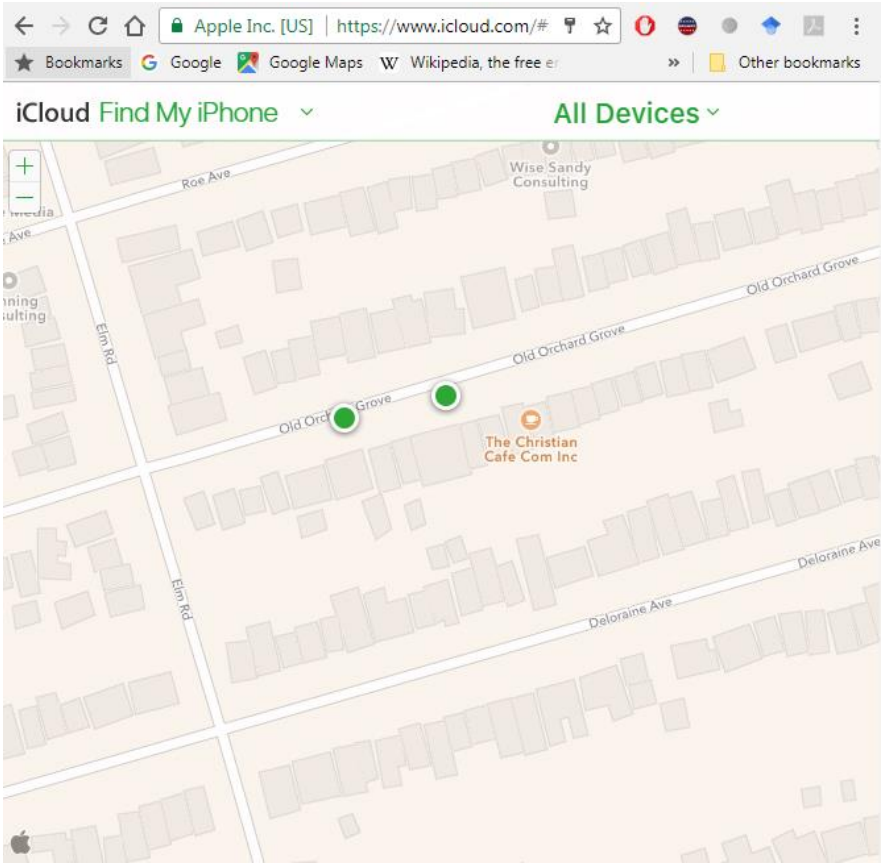

FIGURE 5 – MONITORING INTERVIEWER LOCATION USING FIND MY PHONE

**Note, as the interviewers had not been advised that Find my Phone could be used, this was only done for the purposes of generating the screenshot, but was not actually employed during workdays in the summer.*

There is no indication that travel was misreported or that the interviewers misled the survey team in any manner, but for the same reason market research firms and call centers record the conversations their agents have, it may be desirable for larger-scale efforts to look into recording the audio for interviewers and interviewees. This could be carried out by installing an application on the iPads to record audio when an interview begins or could be handled otherwise.

Such a decision involves a deeper conversation about the legal and ethical implications of the recording, but for the benefit of ensuring the highest quality data be collected, should be considered for future in-person data collection efforts.

8.2.6 Mailers

A very straightforward way of reducing costs is to ensure the envelopes used allow for 8.5”X11”sheets to be folded in 2 rather than 3. Folding in three was chosen because relatively inexpensive 2-window envelopes were found which could be used to both show an institution seal and show an address, but a 2-fold version would save considerable labour time.

8.2.7 Proximity to transit

For reasons outlined in section 3.2, the tracts chosen for further analysis were adjacent to subway stations or proximate to high capacity transit infrastructure. If the effort were to be broadened, it may be worth investing in some means by which to ferry interviewers from a more central location to the locations where they are to carry out their work. Velomobiles like Organic Transit’s ELF could be acquired for ongoing projects and made use of for delivery of materials, as well as interviewer drop-off. This would be a sustainable way to carry out travel for the project and the vehicles could be reallocated after project completion.

9 REFERENCES

Abt Associates Inc., 2012. *Family and Medical Leave in 2012: Methodology Report,* Cambridge: Abt Associates Inc..

Albaum, G. S. & Smith, S. M., 2010. *An Introduction to Marketing Research,* New Mexico: Qualtrics Survey University.

Alsnih, R., 2006. Characteristics of Web-based surveys and applications in travel research. *Travel Survey Methods, Quality and future directions,* pp. 569-592.

Ashley, D., Richardson, T. & Young, D., 2009. *Recent Information on the Under-Reporting of Trips in Household Travel Surveys,* Melbourne, Australia: Australasian Transport Research Forum.

Badoe, D. & Steuart, G., 2002. Impact of interviewing by proxy in travel survey conducted by telephone. *Journal of Advanced Transportation.*

Baker, R. P. & Couper, M. P., 2007. The Impact of Screen Size and Background Color on Response in Web Surveys. *General Online Research Conference (GOR).*

Barr, M. L., van Ritten, J. J., Steel, D. G. & Thackway, S. V., 2012. Inclusion of mobile phone numbers into an ongoing population health survey in New South Wales, Australia: design, methods, call outcomes, costs and sample representativeness. *BMC Medical Research Methodology,* p. 12:177.

Bayart, C. & Bonnel, P., 2012. Combining web and face-to-face in travel surveys: comparability challenges. *Transportation,* Volume 39, p. 1147–1171.

Bell, D. S., Mangione, C. M. & Kahn, C. E., 2001. Randomized testing of alternative survey formats using anonymous volunteers on the world wide web. *Journal of the American Medical Informatics Association,* pp. 616-620.

Bergstrom, J. C. R., Erdman, C. & Lakhe, S., 2016. *Navigation Buttons in Web-Based Surveys: Respondents' Preferences Revisited in the Laboratory*. [Online]
Available at: http://surveypractice.org/index.php/SurveyPractice/article/view/303/html_51

Bernard, M., Chaparro, B., Mills, M. & Halcomb, C., 2003. Comparing the effects of text size and format on the readability of computer-displayed Times New Roman and Arial text. *International Journal Human-Computer Studies,* pp. 823-835.

Bernard, M. L., Chaparro, B. S., Mills, M. M. & Halcomb, C. G., 2003. Comparing the effects of text size and format on the readibility of computer-displayed Times New Roman and Arial text. *International Journal of Human-Computer Studies,* pp. 823-835.

Boehm, L. E., 1989. *Reliability of Proxy Responses in the Current Population Survey,* s.l.: U.S. Bureau of Labour Statistics.

Bonnel, P., 2001. *Postal, telephone and face-to-face surveys: How comparable are they?,* s.l.: International Conference on Transport Survey Quality and Innovation, Kruger Park, South-Africa.

Bonnel, P. & Le Nir, M., 1998. The quality of survey data: Telephone versus face-to-face interviews. *Transportation,* Volume 25, p. 147–167.

Bose, J. & Giesbrecht, L., 2004. Patterns of proxy usage in the 2001 National Household Travel Survey. *ASA: Survey Research Methods, Bureau of Transportation Statistics.*

Bourbonnais, P.-L. & Morency, C., 2013. Web-Based Travel Survey: A Demo. In: *Transport Survey Methods: Best Pracice for Decision Making*. s.l.:Emerald Group, pp. 207-223.

Bullock, P., Stopher, P. & Horst, F., 2003. *Conducting a GPS Survey with Time-Use Diary,* Washington, Dc: Transportation Research Board.

Cerdá, A., 2014. *Innover dans la collecte de données transport pour des données actuelles et de meilleure qualité*. Montreal, Transportation Association of Canada.

Chapleau, R., 2003. Measuring internal quality of a CATI travel survey. *International Conference on Transport Survey Quality and Innovation,* p. 23.

Chen, M., Srikukenthiran, S., Habib, K. & Miller, E., 2016. *Sampling frame for household travel surveys: under-coverage issue,* s.l.: University of Toronto Transportation Research Institute.

Chiao, K. A. et al., 2011. Continuous Improvement in Regional Household Travel Surveys: New York Metropolitan Transportation Council Experience. *Transportation Research Record: Journal of the Transportation Research Board,* pp. 74-82.

Childers, T. L. & Jass, J., 2002. All Dressed Up With Something to Say: Effects of Typeface Semanics Associations on Brand Perceptions and Comuser Memory. *Journal of Consumer Psychology,* pp. 93-106.

City of Toronto, 2017. *Address Points (Municipal) - Toronto One Address Repository,* Toronto: s.n.

Clarke, T. & Costall, A., 2007. The emotional connotations of color: a qualitative investigation. *Color Research and Application,* pp. 406-410.

Cobb, C. & Krosnick, J. A., 2009. *Experimental Test of the Accuracy of Proxy Reports Compared to Target Report with Third-Party Validity.*. Hollywood, Florida, s.n.

Colour Blindess Awareness, 2016. *Colour Blindness*. [Online]
Available at: http://www.colourblindawareness.org/colour-blindness/

Conrad, F. G., Peytchev, A., Couper, M. P. & Tourangeau, R., 2010. Increasing Respondents' Use of Defintions in Web Surveys. *J Off Stat,* pp. 633-650.

Conrad, F. G., Schober, M. F. & Coiner, T., 2007. Bringing features of human dialogue to web surveys. *Applied Cognitive Psychology,* pp. 165-187.

Couper, M. P., Baker, R. & Mechling, J., 2011. *Placement and Design of Navigation Buttons in Web Surveys*. [Online]
Available at: http://surveypractice.org/index.php/SurveyPractice/article/view/93/html

Couper, M. P., Kennedy, C., G., C. F. & Tourangeau, R., 2011. Designing Input Fields for Non-Narrative Open-Ended Responses in Web Surveys. *Journal of Official Statistics,* pp. 65-85.

Couper, M. P. & Tourangeau, R., 2004. *Picture This! Exploring Visual Effects in Web Surveys.* [Online] Available at: http://poq.oxfordjournals.org/content/68/2/255.extract

Couper, M. P., Tourangeau, R., Conrad, F. & Crawford, S., 2004. What they see is what they get: Response options for web surveys. *Social Science Computer Review,* pp. 111-127.

Couper, M. P., Tourangeau, R., Conrad, F. G. & Singer, E., 2006. Evaluating the Effectiveness of Visual Analog Scales. *Social Science Computer Review,* pp. 227-245.

Couper, M. P., Traugott, M. W. & Lamias, M. J., 2001. Web Survey Design and Administration. *Public Opinion Quarterly,* pp. 230-253.

Couper, M. P., Traugott, M. W. & Lamias, M. J., 2001. Web Survey Design and Administration. *Public Opinion Quaterly,* pp. 230-253.

Crawford, S., Couper, M. & Lamias, M., 2001. Web surveys: Perceptions of burden. *Social Science Computer Review,* pp. 146-162.

Data Management Group, 2014. *TTS 2011, Design and conduct of the survey,* Toronto: s.n.

Dillman, D., Smyth, J. & Christian, L., 2009. *Internet, mail, and mixed-mode surveys: the tailored design method (3rd ed.)*. Hoboken: John Wiley and Sons Inc..

DMG, 2014. *Transportation Tomorrow Survey 2011: Design and Conduct of Survey,* Toronto: Data Management Group.

DMG, 2015. *TTS 2.0: An R&D Program to Develop the Next Generation Passenger Travel Survey Method for the Greater Golden Horseshoe,* Toronto: University of Toronto Transportation Research Institute.

DMG, 2016. *Transportation Tomorrow.* [Online]
Available at: http://www.transportationtomorrow.on.ca/publications.html

Doyle, J. K., 2005. Face-to-Face Surveys. In: *Encyclopedia of Statistics in Behavioral Science.* s.l.:s.n.

Duffy, B. & Smith, K., 2005. Comparing data from online and face-to-face surveys. *International Journal of Market Research, 47*(6), pp. 615-639.

Dumont, J., 2009. *Trip Reporting and GPS-Based Prompted Recall: Survey Design and Preliminary Analysis of Results,* Toronto: University of Toronto.

Dyson, M. C. & Haselgrove, M., 2001. The influence of reading speed and the line length on the effectiveness of reading from a screen. *International Journal of Human-Computer Studies,* pp. 585-612.

Dyson, M. C. & Kipping, G. J., 1998. The effects of line length and method of movement on patterns of reading from screen. *Visible Language,* pp. 150-181.

Edmonton, 2016. *Making Tracks*. [Online]
Available at: https://www.edmonton.ca/transportation/traffic_reports/travel-surveys.aspx

eyequant, 2013. *108 Million Web Users Are Color Blind: How Do They See Your Website?*. [Online]
Available at: http://blog.eyequant.com/blog/2013/07/02/108-million-web-users-are-color-blind-how-do-they-see-your-website

Faulkner, C., 1998. *The essence of human-computer interaction*. London: Prentice Hall.

Fraser, T. & Banks, A., 2004. *Designer's Color Manual: The Complete Guide to Color Theory and Application*. s.l.:Chronicle Books.

Fuchs, M., 2009. The Reliability of Children's Survey Responses: The Impact of Cognitive Functioning on Respondent Behaviour. *Statistics Canada's International Symposium Series*.

Galesic, M., Tourangeau, R., Couper, M. P. & Conrad, F. G., 2009. Eye-Tracking Data: New Insights on Response Order Effects and Other Cognitive Shortcuts in Survey Responding. *Public Opinion Quarterly,* pp. 892-913.

Galitz, W. O., 2002. *The Essential Guide to User Interface Design*. [Online]
Available at: http://ps.fragnel.edu.in/~dipalis/prgdwnl/eguid.pdf

Gingrich, P., 2004. *Association between Variables,* Regina: Gingrich, Paul.

Guéguen, N. & Jacob, C., 2004. Solicitation by E-Mail and Solicitor's Status: A Field Study of Social Influence on the Web. *CyberPsychology & Behavior,* pp. 377-383.

Guner, H. & Inal, Y., 2015. The Effect of Banner Location on Banner Recognition in a Turkish Government Website: An Eye Tracking Study. *Human-Comperter Interation: Users and Contexts,* pp. 65-72.

Habib, K. N. & El-Assi, W., 2015. *Sample Size Requirements for Regional Household Travel Surveys,* Toronto: UTTRI.

Harding, C., Nasterska, M., Dianat, L. & Miller, E. J., accepted. Effect of land use and survey design on trip underreporting in Montreal and Toronto's regional surveys. *European Journal of Transport and Infrastructure Research (EJTIR)*.

Harding, C., Zhang, Y. & Miller, E., 2016. *Waterfront Toronto Smartphone Data Collection Project, Technical report,* s.l.: University of Toronto Transportation Research Institute.

Hassounah, M. I., Cheah, L.-S. & Steuart, G. N., 1993. *Under-Reporting of Trips in Telephone Interview Travel Surveys,* Toronto: Join Program in Transportation, University of Toronto.

Healey, B., 2007. Drop downs and scroll mice: The effect of response option format and input mechanism employed on data quality in web surveys. *Social Science Computer Review,* pp. 111-128.

Heerwegh, D., 2004. Using progress indicators in web surveys. *59th AAPOR Conference*.

Heerwegh, D. & Loosveldt, G., 2002. Online persuasion strategies. *International Workshop on Household Survey Nonresponse.*

Heerwegh, D., Vanhove, T., Matthijs, K. & Loosveldt, G., 2005. The effect of personalization on response rates and data quality in web surveys. *International Journal of Social Research Methodology,* pp. 85-99.

Holbrook, A. L., Green, M. C. & Krosnick, J. A., 2003. Telephone versus face-to-face interviewing of national probability samples with long questionnaires; Comparisons of respondentsatisficing and social desirability response bias. *Public Opinion Quarterly,* Volume 67, p. 79–125.

Hoogendoorn-Lanser, S., Schaap, N. T. W. & OldeKalter, M.-J., 2015. The Netherlands Mobility Panel: An innovative design approach for web-based longitudinal travel data collection. *Transportation Research Procedia,* Volume 11, p. 311 – 329.

Hu, S., Balluz, L., Battaglia, M. P. & Frankel, M. R., 2010. Improving Public Health Surveillance Using a Dual-Frame Survey of Landline and Cell Phone Numbers. *American Journal of Epidemiology,* pp. 703-711.

Jackson, A. C. et al., 2013. Improving Gambling Survey Research Using Dual-Frame Sampling of Landline and Mobile Phone Numbers. *Journal of Gambling Study,* pp. 291-307.

Joinson, A. N., Reips, U.-D., Buchanan, T. & Schofield, C. B. P., 2010. Privacy, Trust, and Self-Disclosure Online. *Human-Computer Interaction,* pp. 1-24.

Joinson, A. N., Woodley, A. & Reips, U.-D., 2007. Personalization, authentication and self dis-closure in self-administered Internet surveys. *Computers in Human Behaviour,* pp. 275-285.

Jones, R. & Pitt, N., 1999. Health surveys in the workplace: Comparison of postal, e-mail and World Wide Web methods. *Occupational Medicine,* pp. 556-558.

Kaya, N. & Epps, H. H., 2004. Color-emotion association: past experience and personal preference. *AIC 2004 Colour and Paints, Interime Meeting of the International Color Association,* pp. 31-34.

Krug, S., 2006. *Don't Make Me Think: A Common Sense Approach to Web Usability*. Berkeley: Neew Riders.

Labrecque, L. I. & Milne, G. R., 2011. *Exciting red and competent blue: the importance of color in marketing*. [Online]
Available at: http://link.springer.com/article/10.1007/s11747-010-0245-y

Labrecque, L. & Milne, G., 2011. Exciting red and competent blue: The importance of color in marketing. *Journal of the Academy of Marketing Science ,* pp. 40-45.

Leavitt, M. O. & Sneiderman, B., 2004. *Research-Based Web Design & Usability Guidelines*. [Online]
Available at: https://www.usability.gov/sites/default/files/documents/guidelines_book.pdf

Lebrasseur, D., Morin, J. P., Rodrigue, J. F. & Taylor, J., 2010. Evaluation of the innovations implemented in the 2009 Canadian Census Test. *In Proceedings of the American Statistical Association Survey Research Methods Section,* pp. 4089-4097.

Lee-Gosselin Associates Limited, 2012. *Changing Practices in Data Collection on the Movement of People,* Sainte-Pétronille: Lee-Gosselin Associates Limited.

Lee, S., Brick, M., Brown, R. & Grant, D., 2010. Growing Cell-Phone Population and Noncoverage Bias in Traditional Rnadom DIgit Dial Telephone Health Surveys. *Health Research and Educational Trust,* pp. 1121-1139.

Ling, J. & Schaik, P. v., 2006. The influence of font type and line length on visual search and information retrieval in web pages. *International Journal of Human-Computer Studies,* pp. 395-404.

Livingston, M. et al., 2013. Surveying Alcohol and other drug use through Telephone Sampling: a Comparison of Landline and Mobile Phone Samples. *BMC Medical Research Methodology,* p. 13:41.

Loa, P., Srikukenthiran, S., Habib, K. N. & Miller, E. J., 2015. *Current State of Web-Based Survey Methods,* Toronto: University of Toronto Transportation Research Institute.

Lu, B. et al., 2013. Design and Analysis of Dual-Frame Telephone Surveys for Health Policy Research. *World Medical and Health Policy,* pp. 217-232.

Lynch, P. J. & Horton, S., 2001. *Web style guide: Basic design principles for creating Web sites*. New Haven: Yale University Press.

Mahnke, F. H., 1996. *Color, Environment, and Human Response: An Interdisciplinary Understanding of Color and Its Use as a Beneficial Element in the Design of the Architectural Environment*. s.l.:Wiley.

Malatest, 2017. *Survey Status by Municipality, Region and PD,* s.l.: s.n.

Malatest, 2017. *Transportation Tomorrow Survey: TTS 2016 Challenges and Lessons Learned (Draft),* Toronto: malatest.

Manfreda, K. L., Batagelj, Z. & Vehovar, V., 2002. Design of Web Survey Questionnaires: Three Basic Experiments. *Journal of Computer-Mediated Communication.*

Mavletova, A. & Couper, M. P., 2014. Mobile Web Survey Design: Scrolling Versus Paging, SMS Versus E-mail Invitations. *Journal of Survey Statistics and Methodology,* pp. 498-518.

McGeeney, K. & Marlar, J., 2013. Mobile Browser Web Surveys: Testing Response Rates, Data Quality and Best Practices. *The American Association for Public Opinion Research (AAPOR) 68th Annual Conference.*

McHugh, M. L., 2013. The Chi-square test of independence. *Biochemia Medica,* pp. 143-149.

Meier, B. P. & Robinson, M. D., 2004. Why the sunny side is up: Association between affect and vertical position. *Psychological Science,* pp. 243-247.

Micheal S. McCarthy, D. L. M., 2002. Effects of typographic facturs in advertising-based persuasion: A general model and inital emperical tests. *Psycholinguistics & Marketing,* pp. 663-691.

Muehlenhaus, I., 2013. *Web Cartography: Map Design for Interactive and Mobile Devices*. Boca Raton, FL: CRC Press.

Murray, D. C. & Deabler, H. L., 1957. Colors and mood-tones. *Journal of Applied Psychology,* pp. 179-283.

National Cooperative Highway Research Program, 2016. *Standardized Procedures for Personal Travel Surveys,* Washington: NCHRP.

Nielsen, J., 2004. *Guidelines for Visualizing Links.* [Online]
Available at: https://www.nngroup.com/articles/guidelines-for-visualizing-links/

Nielsen, J., 2007. *Banner Blindness: Old and New Findings.* [Online]
Available at: https://www.nngroup.com/articles/banner-blindness-old-and-new-findings/

Nielson, J., 2006. *F-Shaped Pattern for Reading Web Content.* [Online]
Available at: https://www.nngroup.com/articles/f-shaped-pattern-reading-web-content/

Novemsky, N., Dhar, R., Schwarz, N. & Simonson, I., 2007. Preference Fluency in Choice. *Journal of Marketing Research,* pp. 347-356.

NuStats, 2002. *2000 - 2001 California Statewide Household Travel Survey,* Austin: NuStats.

NuStats, 2007. *Chicago Regional Household Travel Inventory,* Austin: NuStats.

NuStats, 2013. *2010-2012 California Household Travel Survey Final Report Appendix,* Austin: NuStats.

Nygren, E., 1996. From paper to computer screen: human information-processing and user interface design. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology,* p. 188.

Odbert, H. S., Karwoski, T. F. & Eckerson, A. B., 1942. Studies in synesthetic thinking: Musical and verbal associations of color and mood. *Journal of General Psychology,* pp. 153-173.

Office Fédéral de la Statistique, 2012. *La Mobilité en Suisse,* s.l.: s.n.

Ontario, 2017. *How to make website accessible.* [Online]
Available at: https://www.ontario.ca/page/how-make-websites-accessible#section-2

Patterson, Z., Fitzsimmons, K., Widener, M. & Reid, J., 2018. *Recruitment, Burden, Incentives and Participation in Smartphone Travel Surveys.* Washington, s.n.

Pearson, J. & Levine, R. A., 2003. Salutations and Response Rates to Online Surveys. *The Impact of Technology on the Survey Process,* pp. 351-362.

Peytchev, A., Couper, M. P., McCabe, S. E. & Crawford, S. D., 2006. Web survey design: Paging versus scrolling. *Public Opinion Quarterly,* p. 596–607.

Peytchev, A. & Hill, C., 2010. Experiments in mobile web survey design: Similarities to other modes and unique considerations. *Social Science Computer Review,* pp. 319-335.

Pierce, B., Casas, J. & Giaimo, G., 2002. *Estimating Trip Rate Under-Reporting: Preliminary Results from the Ohio Household Travel Survey,* Washington: Transportation Research Board.

Pope, D. & Baker, R., 2005. Experiments in color for Web-based surveys. *FedCASIC response*.

Resource Systems Group Inc., 2013. *Utah Travel Study,* Utah: Utah Department of Transportation (UDOT).

Richardson, A., 2005. *Proxy Responses in Self-Completion Travel Diary Surveys,* Washington: The Urban Transport Institute for Reliable Urban Transportation Information.

Schmidt, K. E., Liu, Y. & Sridharan, S., 2009. Webpage aesthetics, performance and usability: Design variables and their effects. *Ergonomics,* pp. 631-643.

Schwarz, N. et al., 1991. Rating scales: Numeric values may change the meaning of the scale labels. *Public Opinion Quarterly,* pp. 618-630.

Shrestha, S. & Lenz, K., 2007. *Eye Gaze Patterns while Searching vs Browsing a Website*. [Online] Available at: http://usabilitynews.org/eye-gaze-patterns-while-searching-vs-browsing-a-website/

Siemiatycki, J., 1979. A Comparison of Mail, Telephone, and Home Interview; Strategies for Household Health Surveys. *American Journal of Public Health,* Volume 69, pp. 238-245.

Sills, S. J. & Song, C., 2002. Innovations in survey research an application of web-based surveys. *Social science computer review,* pp. 22-30.

Stapleton, C. E., 2013. The Smartphone Way to Collect Survey Data. *www.surveypractice.org,* pp. Vol. 6, no 2.

State Health Access Data Assistance Center, 2013. *Technical Report for the 2011 Minnesota Healtrh Access Survey: Survey Methodology, Weighting and Data Editing,* Minnesota: SHADAC.

Statistics Canada, 2014. *Residential Telephone Service Survey, 2013*. [Online] Available at: http://www.statcan.gc.ca/daily-quotidien/140623/dq140623a-eng.htm [Accessed 30 8 2017].

Stopher, P. & Greaves, S. P., 2007. Household travel surveys: Where are we going?. *Transportation Research Part A,* Volume 41, p. 367–381.

Stopher, P. R., Wilmot, C. G., Stecher, C. & Alsnih, R., 2003. Standards for household travel survey - some proposed ideas. *10th Triennial Conference of the International Association for Travel Behaviour Research,* p. 24.

Student Move TO, 2016. *Student Move TO*. [Online] Available at: http://www.studentmoveto.ca/

SurveyMonkey, 2016. *SurveyMonkey Blog*. [Online] Available at: https://www.surveymonkey.com/blog/2016/05/17/progress-bars/

Swain, L., Drew, J., Lafrance, B. & Lance, K., 1992. The creation of a residential address register for coverage improvement in the 1991 canadian census. *Survey Methodology,* 18(1), pp. 127-141.

Tal, G., Favetti, M. & Nicholas, M., 2015. *Exploring the Use of a Web Based Map Tool for Travel Behavior Data Collection: Lessons from Plug-in Vehicle Owner Surveys,* Davis, CA: TRB 2016 Annual Meeting.

timeanddate.com, n.d. *Sunrise, Sunset, and Daylength.* [Online]
Available at: https://www.timeanddate.com/sun/canada/toronto
[Accessed 24 11 2017].

Toepoel, V., Das, M. & Soest, A. V., 2008. Effects of Design in Web Surveys: Comparing Trained and Fresh Respondents. *Public Opinion Quarterly,* pp. 985-1007.

Tourangeau, R. et al., 2006. Everyday Concepts and Classification Errors: Judgments of Disability and Residence. *Journal of Official Statistics,* pp. 385-418.

Tourangeau, R., Conrad, F. G. & Couper, M. P., 2013. *The Science of Web Surveys.* s.l.:Oxford Scholarship Online.

Tourangeau, R., Conrad, F. G., Couper, M. P. & Ye, C., 2014. *The Effects of Providing Examples in Survey Questions.* [Online]
Available at: http://poq.oxfordjournals.org/content/early/2014/02/17/poq.nft083

Tourangeau, R., Couper, M. P. & Conrad, F., 2004. Spacing, Position, and Order - Interpretive Heuristics for Visual Features of Survey Questions. *Public Opinion Quarterly,* pp. 368-393.

Tourangeau, R., Couper, M. P. & Galešic, M., 2005. Use of eye-tracking for studying survey response processes. *ESF Workshop.*

Tourangeau, R., Groves, R. M., Kennedy, C. & Yan, T., 2009. The Presentation of a Web Survey, Nonresponse and Measurement Error among Members of Web Panel. *Journal of Official Statistics,* pp. 299-321.

TRB Travel Survey Methods Committee, n.d. *Chapter 25 Costs and trip rates of recent household travel surveys.* [Online]
Available at: http://www.travelsurveymanual.org/Chapter-25-1.html
[Accessed 20 11 2017].

TRB Travel Survey Methods Committee, n.d. *Chapter 3 Options for Travel Surveys.* [Online]
Available at: http://www.travelsurveymanual.org/Chapter-3-1.html
[Accessed 20 11 2017].

Trouteaud, A. R., 2004. How you ask counts: A test of internet-related components of response rates to a Web-based survey. *Social Science Computer Review,* pp. 385-392.

USDOT, 2016. *National Household Travel Survey,* s.l.: Westat.

UWSC, 2010. *Survey Fundamentals: A Guide to Designing and Implementing Surveys.* [Online]
Available at: https://oqi.wisc.edu/resourcelibrary/uploads/resources/Survey_Guide.pdf

Verreault, H. & Morency, C., 2015. *What about Proxy Respondent bias Over Time?,* Montreal, Quebec: CIRRELT: Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation.

Verreault, H. & Morency, C., 2016. Integration of a phone-based household travel survey and a web-based student travel survey. *Transportation,* p. New York.

Vischeck, 2015. *Vischeck simulates colorblind vision.* [Online]
Available at: http://www.vischeck.com/

W3C, 2016. *About* W3C. [Online]
Available at: https://www.w3.org/Consortium/

Walters, J., Apter, M. J. & Svebak, S., 1982. Color preference, arousal, and the theory of psychological reversal. *Motivation and Emotion,* pp. 193-215.

Wargelin, L. & Kostyniuk, L. P., 2014. *Self and Proxy Respondents in Household Travel Surveys: A Comparison,* s.l.: Research Gate.

WCAG, 2016. *WAI: Strategies, guidelines, resources to make the Web accessible to people with disabilities.* [Online]
Available at: https://www.w3.org/WAI/WCAG20/quickref/#content-structure-separation

Westat, 2016. *2015-2016 National Household Travel Survey: What's Different and New.* [Online]
Available at: http://nhts.ornl.gov/trb/2016/Workshop-Westat.pdf

Wexner, L. B., 1954. The degree to which colors (hues) are associated with mood-tones. *Journal of Applied Psychology,* pp. 432-435.

Wilkinson, S. & Payne, S., 2006. Eye tracking to identify strategies used by readers seeking informaiton from on-line texts. *Proceedings of the 13th European Conference on Cognitive Ergonomics: Trust and Control in Complex Socio-Technical Systems,* pp. 115-116.

Wright, W. D., 1988. Talking about color. *Colour Research and Application,* pp. 135-203.

Wroblewski, L., 2008. *Web form design: Filling in the blanks.* Brooklyn: Rosenfeld Media.

Yan, T., Conrad, F. G., Tourangeau, R. & Couper, M. P., 2011. Should I Stay or Should I go: The Effects of Progress Feedback, Promised Task Duration, and Length of Questionnaire on Completing Web Surveys. *International Journal of Public Opinion Research,* pp. 131-147.