Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 32 (2018) 659–665

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Steering Committee for Transport Survey Conferences

TTS2.0: A research and development (R&D) project on passenger travel survey methods

Khandker Nurul Habib*[a], Eric J. Miller[a], Siva Srikukenthrian[a], Martin Lee-Gosselin[b], Catherine Morency[c], Matthew J. Roorda[a], Amer Shalaby[a]

[a]*Civil Engineering, University of Toronto, Toronto*
[b]*Universtiy of Laval, Laval, Quebec*
[c]*Ecole Polytechnique de Montreal*

Abstract

This paper describes a multi-year R&D project on developing the next generation of passenger travel survey methods for operational use in the Greater Golden Horseshoe (GGH) that was currently underway at the time of writing and reports on results obtained to date. While the current household travel survey method used in the GGH for passenger travel surveys – the Transportation Tomorrow Survey (TTS) – has been used for efficiently gathering standardized, high-quality travel information for the Greater Golden Horseshoe (GGH) region over the past 25 years, changes in communications technologies and other methodological issues are presenting increasing challenges to the current TTS method. In many ways, these issues are germane to any household travel survey. Passenger travel survey data are critical to our collective, on-going transportation planning, policy analysis, and decision-making. So, it is essential to develop and test new data collection methods that will be available to supplement or replace the current practices of household travel surveys. The TTS2.0 is such an R&D program that is investigating and testing a broad range of alternative designs, leading to a recommended new, comprehensive passenger travel survey program for the GGH. It has been carefully designed to address this challenge.

Keywords: Household travel surveys; core-satellite design; multi-instrument travel data; web-based surveys; smartphone apps; data fusion

* Corresponding author. Tel.:+1-416-946-8027;
E-mail address: khandker.nurulhabib@utoronto.ca

2352-1465 © 2018 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/)
Peer-review under responsibility of the International Steering Committee for Transport Survey Conferences (ISCTSC).
10.1016/j.trpro.2018.10.004

1. Introduction

TTS 2.0 is a three-year R&D project intended to develop the next generation of passenger travel survey methods for operational use in the Toronto-centred Greater Golden Horseshoe (GGH) region in southern Ontario, Canada. While the current household travel survey method used in the GGH for passenger travel surveys – the Transportation Tomorrow Survey (TTS) – has been used for efficiently gathering standardized, high-quality travel information for the GGH region over the past 30 years, changes in communications technologies and other methodological issues are presenting increasing challenges to the current TTS method. At the time of this paper's preparation, it is approximately two years through the planned three-year research project. The purpose of this paper is to provide an overview of the overall project, some of its key data collection design concepts, and a high-level report on key findings to date,

TTS 2.0 builds directly on a major report prepared in 2012 by most of this paper's authors for the Transportation Association of Canada (Miller, et al., 2012). In the 2012 study, the state of art and practice in Canadian person-based urban travel data collection methods was reviewed in detail. The report argues strongly that the current methods employed in most major Canadian cities require significant updating to reflect changing planning and modelling needs, emerging new methods for data collection, and growing technical problems with conventional data collection methods. Responding to this report, and recognizing growing challenges facing the TTS data collection program, the Transportation Information Steering Committee (TISC) – the consortium of the Ontario Ministry of Transportation and regional and local municipalities in the GGH that is responsible for the TTS data collection program – commissioned the University of Toronto Transportation Research Institute's (UTTRI) Data Management Group (DMG), through this paper's authors, to undertake a major redesign of the region's data collection program: TTS 2.0.

A key design principle underlying TTS 2.0 thinking is that of a core-satellite (multi-instrument) approach to the design of a comprehensive travel behaviour data collection program (Goulias, et al., 2011; Miller et al., 2012). The next section of the paper briefly describes this key concept. A second key assumption in the design of the TTS 2.0 research project is that we needed to "start from scratch", with "everything being on the table", in terms of considering as many different approaches to travel data collection as possible. We viewed this as particularly critical given both the complexity of modern travel behaviour in large urban areas and the rapidly changing nature of emerging data collection methods. The third section of the paper presents an overview of the research approach adopted in TTS 2.0 to address these needs for design comprehensiveness and flexibility. The paper's fourth section briefly reports on selected key findings that have been obtained to date within the project, while the final section offers a few very brief comments concerning next steps and expected outcomes of the project.

2. The Core-Satellite Design Paradigm

A key organizing principle in approaching the TTS 2.0 design is the adoption of a core-satellite paradigm. Detailed specification and testing of this design as a practical approach to comprehensive and flexible travel data collection within a large urban region lie at the heart of developing TTS 2.0.

The primary objective of a core survey is to maintain a stable and representative sample of travel demand, through a large-sample survey of the population that collects a limited (but key / "core") population and travel attributes. Satellite surveys are small-scale surveys that are opportunistically implemented with the intent to complement/supplement the core survey data with additional information and/or sample of the population. Implementation of a core-satellite paradigm for urban passenger travel demand requires sound and practical understanding of all aspect of travel surveys, including issues related to sample frames, recruitment procedures, mode of survey, etc. Deep understanding of many of these issues requires practical experiments, which are often expensive and time-consuming to undertake. As discussed further in the next section, the TTS2.0 project involves an elaborate strategy for detailed investigation of this broad range of issues and methods.

In particular, it is clear that the field is evolving very rapidly in its use of both web-based and smartphone collection methods, as well as in the use of passive data in a variety of ways. At the same time, it is not clear that one clear "winner" among these various approaches exists, at least for the foreseeable future. It appears that "mixed methods" will be the likely way forward, in which the relative strengths of each method are exploited (and the relative weaknesses are avoided) as much as possible (Lo et al., 2017; Srikukenthiran et al., 2017). As briefly described in Section 4, each of these major approaches is being investigated in depth within the TTS 2.0 project.

Further, this core-satellite, "mixed methods" approach to a comprehensive data collection program inevitably will require the use of various data fusion methods for combining data from different sources in a statistically optimal way. A key element of TTS 2.0 will be the explicit incorporation of "fusion by design" in designing the various components of the data collection program. Work on this issue to date within TTS 2.0 is also briefly discussed in Section 4.

Designing a data collection program according to the core-satellite survey paradigm offers the opportunity to collect more detailed data than would be collected through a traditional household travel survey. Consisting of three components, the core survey, satellite surveys, and complementary datasets, the design of each component should be tailored to suit the characteristics of the target population, facilitate the enabling of the desired capability, and serve the goals of the survey. Within the GGH, there is strong interest in ensuring that survey data are compatible with the TTS, based either on the spatial, temporal, and/ or semantic contexts. This provides the opportunity to utilize data collected through a large-sample travel survey that utilizes a survey method that has already been validated.

3. TTS 2.0 Research Approach

The TTS2.0 study consists of three major, interconnected components:

1. ***Methodological research*** to clarify the current and emerging state of the art in urban passenger travel data collection methods (and data sources) and to address a wide range of technical issues that need to be resolved as best as possible prior to the design of a program for field testing alternative data collection methods.
2. ***Field testing*** selected promising data collection methods for testing to determine their efficacy and practicality in the GGH context. These tests range from preliminary, small-sample pilot tests to large-sample field tests of methods that are directly scalable for full regional implementation.
3. ***The design*** of the recommended TTS 2.0 data collection and management program for the GGH for full implementation post-2018.

The classic survey design issues of sampling frame, sampling units, and sample recruitment lie at the heart of any data collection effort and, to a large extent, are the crux of concerns with the current TTS. Shifting to web- or smartphone-based survey instruments will not eliminate this problem but only redefine it. A particularly important element of this problem is the choice of sampling unit. In the current TTS, the household is the basic sampling unit, with travel by all individuals (11 years of age or older) within the sampled households being recorded. This is both an efficient approach to gathering trip-making behaviour of the population as a whole, and it facilitates "household-based" travel modelling in which household contextual factors (competition for household cars, joint household activities, etc.) can be accounted for.

Building upon these general considerations, the broad range of methodological research tasks have been staged in three groups that are being sequentially executed over the three-year project time period to feed into the pilot/field tests being executed over this same time period in an optimal way. The three groups are:

- April-October 2015: Research into land-line, the web, and smartphone-based survey methods for input into the design of the 2016 pilot and field tests (Harding et al 2017a, 2017b and 2017c). Continuous survey methods were also investigated during this first year, given the potential for this approach to significantly impact the overall data collection process design (El-Assi et al. 2017).
- April-October 2016: Research into passive data sources and satellite design options for input in the design of the 2017 pilot and field tests that may well incorporate these elements in various ways, in addition to revised designs arising out of the 2016 tests.
- April-October 2017: Research into data fusion methods for possible input in the 2017 field tests and of the TTC 2.0 design.

Two rounds of pilot and field tests were executed in 2016 and 2017. Each followed the same general pattern of:

- Pilot test design (January-March).
- Pilot test execution (April-August)
- Field test design (primarily June-August).
- Field test execution (September-December, overlapping with 2016 TTS).
- Analysis and evaluation of pilot and field test results (August-March).

The pilot tests and field test design needed to overlap, as well as the analysis of the pilot test results with the start of the field tests, in order to keep the overall 3-year timetable feasible and to be able to run the field tests in parallel to the 2016 TTS. A longer time period was allocated to the first round of pilot tests (October 2015 – March 2016) in order to allow for the large number of options that we expected to need to evaluate for this first round of tests. It was expected that the shorter January – March 2017-time period for the design of the 2017 pilot tests should be sufficient.

4. Selected Preliminary Results

4.1 Web-Based Surveys

In terms of survey mode for potential core-satellite structure, the use of web-based technologies is deemed to be critical because of their potential to decrease respondent burden, improve data quality, and lower costs. Many such surveys are custom-built since commercial survey builder platforms (e.g. Survey Monkey, Qualtrics, etc.) are not usually tailored for activity-travel data collection. These platforms lack features that allow efficient data collection of specific travel information, such as location piping between questions and the use of interactive maps for geospatial data collection. As an alternative approach to travel survey development, a custom web platform (TRAISI) was developed (Chung et al. 2017). Objectives in the development of TRAISI included:

- Addressing known weaknesses in the current TTS CATI data collection software concerning location accuracy and household proxy bias.[†]
- Developing an easy-to-use user interface that provides high-quality information concerning all trips, regardless of mode(s) chosen and complexity of the trip.
- Minimizing respondent burden.
- Providing a flexible and extensible software platform for the rapid deployment of a variety of travel surveys.

Results to date of TTS2.0 field and pilot studies with the newly designed software indicate that the use of web-surveys compared to conventional CATI methods can significantly reduce the proportion of proxy responses in a household travel survey. I.e., given the advantages of web surveys, it is possible to survey more than one member of the household without significantly increasing the response burden.

4.2 Smartphone Apps

With the increasing popularity of cellular telephones, smartphone-based methods for travel data collection provide a possible attractive alternative for collecting travel behaviour information. As a part of the TTS2.0 project, a rigorous assessment of the strengths and weaknesses of smartphones for addressing regional transportation data needs was undertaken. This assessment evaluated smartphones apps as well as trace processing algorithms for extracting useful information from data collected using smartphones.

The first major experiment, conducted in the summer of 2016, used multiple apps and compared their trace data

---

[†] In the current TTS, only one respondent per household provides travel information for all household members.

with rigorously observed ground truth to assess how well the apps performed (Harding et al 2017a, 2017b and 2017c). Smartphone traces were collected by a project research assistant using 17 different apps (Android and iOS combined) installed on 21 phones carried together at all times. Random identifiers were assigned to each app to allow for a high level of detail to be presented without having this report be seen as an endorsement of any one product. A total of 314 trips were made, consisting of 553 legs (including "main mode" access and egress components). Over the course of roughly 130 hours of travel time, 3,655 km were travelled and 1,063 battery recordings were taken. Results of the investigation include:

- Higher recording frequency and location accuracy lead to improved trip end and mode inference accuracies. The relationship between frequency, accuracy and overall performance, however, is non-linear. Likewise, there is a link between higher battery drain and more accurately detected travel. In both cases, there is a point beyond which more frequent logging at high accuracy does not improve performance but merely leads to higher respondent burden in the form of battery drain. This suggests that there is potential for a high-performing app to be designed and tailored for regional travel survey needs that would perform well without causing excessive battery drain.
- Accuracy in terms of accurately detected trip ends is high in most cases.
- Mode inference performance, however, varied significantly across the apps and travel contexts. This indicates that until significant improvements can be made to improve mode inference algorithms, validation data must be sought from survey participants. This is especially true in a setting where the mix of travel modes is diverse, as well as where multi-modal trips account for a significant share of overall travel.

Building on the results of the 2016 tests, a second major smartphone field test was conducted in the fall of 2017. This involved the widespread use of one of two different apps (one iOS-based and one Android-based)[‡] by a convenience sample of GGH trip-makers. Two primary methods were used to recruit volunteers. The first was recruitment from a list of people who had completed the Fall 2016 TTS and had agreed to be re-contacted for follow-up surveys. The second involved widespread use of social media and even an interview on a local public affairs television show to promote participation in the survey. A total of 1550 respondents were recruited who recorded at least 1 trip. Results of this field test were still being analysed at the time of this paper's preparation.

4.3 Passive Datasets

The core-satellite paradigm allows harnessing the potential of passively collected data for travel behaviour investigation. In addition to the data collected from custom smartphone apps discussed in the previous section, typical passive datasets often considered for use in tracking trip-making include:

- Call data records (CDR), which trace cellular phone communications with cell towers.
- Transit smartcard transaction data.
- Onboard vehicle tracking records.
- Credit card transaction data.
- Other third-party location tracking app data, collected/aggregated from a variety of commercial vendors.

Such datasets offer the potential of massive amounts of data (so-called "Big Data") more or less continuously collected over extended time periods tracking very large samples of the trip-making public. As discussed further in the next sub-section, the individuals generating these spatial-temporal traces are generally anonymous and so the data are missing valuable social-economic information. Other potential challenges in using such data include:

- Spatial precision can vary considerably from one source to another, as well as even within a given trace, depending on local conditions.
- These data are generally not collected for travel behaviour analysis purposes and so are not always as well designed or stored as one would like. In particular, travel mode and trip purpose are generally not recorded

[‡] The two apps were bundled/branded under the name of *City Logger* for marketing and recruitment purposes.

and so must be imputed. As noted in the discussion of smartphone traces, this can be a non-trivial problem to resolve.

- These data are often collected by private firms and so the availability and cost of the data can be a significant issue.

In its work, the TTS 2.0 project has reviewed the literature on travel demand data sources but has focussed its analysis to date on transit smartcard transaction data obtained through Metrolinx's[§] recently implemented Presto card. As is typical of most smartcards, which inevitably are primarily designed for fare collection rather than information collection, the data examined to date present a "mixed bag" in terms of usability for travel analysis and modelling purposes. At the time of writing of this paper, work is continuing on investigating how these data might best be utilized within an overall data collection program for the GGH region.

4.4 Data Fusion

One of the critical technical issues related to any core-satellite survey program is the proper understanding of the possible fusion techniques that can integrate the information gathered from multiple instruments/data collection methods to create a comprehensive, joint database. For example, passively collected data such as GPS, smartphone, and transit smart-cards provide useful and detailed information on travel behaviour. These sources often offer continuous streams of travel information of a large population at a relatively inexpensive cost. These passive data sources, however, generally are partial in terms of the range of attributes observed. In particular, they very often lack any information concerning the socio-economic attributes of the trip-maker. They also typically track individuals, with no information concerning the characteristics of the households within which these individuals reside.

Given this, data fusion is generally essential to combine the various partial datasets available into a more comprehensive, detailed joint dataset which combines travel and socio-economic attributes of trip-makers together, as well as, ideally, placing observed individuals within household contexts. While data fusion has the potential to enrich travel survey data, the actual implementation can be very challenging. The largest barrier is the incompatibility of datasets, as most datasets are not designed for integration. Methods to resolve data incompatibility are difficult to generalize as each dataset is unique and needs to be treated on a case-by-case basis. Another barrier is a lack of validity measures – current methods offer insights on variable distributions and correlations, but finer levels of measurements do not yet exist due to mathematical constraints. As for the fusion techniques, many are rigorously discussed at the theoretical level, but few are put into actual use. Partially caused by the complexity and incompatibility of real-life datasets, there is no specific statement on which method should be recommended under certain situations. Nevertheless, the TTS 2.0 project has been reviewing available data fusion methods, and an integral component of the final TTS 2.0 data collection will include recommendations concerning the data fusion methods to be employed within the overall core-satellite design. Recommendations will also be given on how surveys should be designed so as to maximize the ability to effectively fuse the data obtained from the various methods and data sources used.

5. Summary & Next Steps

As briefly summarized in this paper, the TTS 2.0 project has investigated a wide range of survey design issues and methods as it works towards a set of recommendations for a new, comprehensive personal travel data collection program for the Toronto-centred Greater Golden Horseshoe region within southern Ontario. Key contributions of the work to date include: development of new, powerful software for implementing web-based surveys; extensive testing of smartphone apps for collecting trip information; assessment of the strengths and weaknesses of other passive sources of travel-related behaviour (notably transit smartcard transaction data); and investigation of data fusion for use in multi-instrument data collection designs.

[§] Metrolinx is the regional transportation planning authority for the Toronto region.

The emphasis throughout this work has been on developing a core-satellite design, in which the on-going collection of a comprehensive dataset describing travel behaviour within the GGH is accomplished through an optimized set of surveys (or other data collection methods). These consist of a main or "core" survey which is supplemented by one or more "satellite" surveys. The core survey will be a large-sample, relatively simple survey designed to generate high response rates (low response burdens) and to gather key / "core" information about travel and trip-makers in the region. The satellite surveys then target specific sub-markets (transit or active transportation users; elderly trip-makers; etc.) to complement/supplement the core information. Use of advanced data fusion methods will be essential to meld these various datasets together effectively and efficiently.

Final recommendations concerning the TTS 2.0 data collection program are expected to be made by fall 2018. These will be presented in a final project report (which will also document all the activities and findings of the three-year project), as well as in a series of focussed project reports and journal papers.

References

Chung, B., S. Srikukenthiran, E.J. Miller, K.M.N. Habib. 2017. *Development of a Web Survey Builder Platform for Household Travel Surveys*. The paper presented at the 11th International Conference on Transport Survey Methods, Quebec, September 24-29, 2017

El-Assi, W., C. Morency, E.J. Miller, K.M.N Habib, 2017. "*Investigating the Capacity of Continuous Household Travel Survey in Capturing the Temporal Rhythm of Travel Demands*". The paper presented at the 11th International Conference on Transport Survey Methods, Quebec, September 24-29, 2017.

Goulias, K., R. Pendyala and C. Bhat, 2011. Total design data needs for the new generation large scale activity microsimulation models, *Proceedings of the 9th International Conference on Survey Methods in Transport, Scoping the Future While Staying on Track*, Termas de Puyehue, Chile.

Harding, C., S. Srikukenthiran, E.J. Miller, K.M.N. Habib, K.M.N. 2017. *Evaluation of cost-effectiveness and feasibility of in-person surveys as an augment to the regional travel survey: a case study in the Toronto area*. The paper presented at the 11th International Conference on Transport Survey Methods, Quebec, September 24-29, 2017.

Harding, C., Z.T. Zhang, S. Srikukenthiran, E.J. Miller, K.M.N. Habib. 2017. *On the User Experience and Performance of Smartphone Apps as Personalized Travel Survey Instruments: Results from an Experiment in Toronto*. The paper presented at the 11th International Conference on Transport Survey Methods, Quebec, September 24-29, 2017.

Harding, C., Z.T. Zhang, S. Srikukenthiran, E.J. Miller, K.M.N. Habib, 2017. *An Experimental Study for Evaluating Smartphone Apps as Personalized Passenger Travel Survey Tools in Toronto*. The paper presented at the 11th International Conference on Transport Survey Methods, Quebec, September 24-29, 2017.

Miller, E.J., K. Nurul M.N. Habib, M. Lee-Gosselin, C. Morency, M.J. Roorda and A.S. Shalaby, 2012. *Changing practices in data collection on the movement of people,* final report to the Transportation Association of Canada, Île d'Orléans, Québec: Lee-Gosselin Associates Ltd.

Lo, A., Srikukenthiran, S., E.J. Miller, K.M.N. Habib. 2017a. *Impact of a Multiple Survey Frames on Data Quality of Household Travel Surveys: The case of the 2016 Transportation Tomorrow Survey*. The paper presented at the 11th International Conference on Transport Survey Methods, Quebec, September 24-29, 2017.

Srikukenthiran, S., T. Lin, E.J. Miller, K.M.N. Habib. 2017b. *Inverted Sampling Frames to overcome Under-Coverage of specific Population Cohorts: Examining the viability of recruiting households via employers and institutions*. The paper presented at the 11th International Conference on Transport Survey Methods, Quebec, September 24-29, 2017.