Data Collection Technologies – Past, Present, and Future

Sudeshna Sen, Ph.D., NuStats, USA; Stacey Bricka, Ph.D., NuStats, USA

Abstract Over the past few decades, a succession of innovative technologies has revolutionized the surveying industry in the United States. The advent of technological breakthroughs, primarily Computers, GPS, and Internet, have brought about major leaps in the data collection technologies. This paper will document the past, present, and future data collection technologies and their application in the area of travel behavior research. The primary focus of the paper is on current data collection technologies used in survey research in general, including the computer-aided technologies, GPS, and other automated data entry technologies. Some thought is given to emerging technologies such as cellular phone technologies which may influence the gathering of travel behavior details in the future. Recommendations are provided on the methodological design considerations to capture and shape these technologies to meet the needs of travel demand data collection.

1. Introduction

The changing trends in technology have been the driving force in the evolution of surveying methods. With the advent of every new technology, the survey researchers have found avenues to implement them to advance data collection methods. Originally, surveys were conducted using paper-and-pencil methods that evolved into telephone-based surveys. The introduction of innovative technologies have supplemented or replaced conventional data collection methods. In the past three decades, several technologies have made a profound impact on data collection methods in the United States, beginning with Computers in 1970s, Internet in 1990s, and GPS in 1990s.

Computer Assisted Telephone Interviewing (CATI) originated in 1970s and is the oldest computer-aided data collection technology. In the beginning, CATI used computers in a centralized location and was based on a minicomputer system that was eventually replaced by a microcomputer network. In the 1980s, the technolo-
gical change to personal microcomputers facilitated the decentralization of CATI making it possible for interviewers to conduct the surveys from any 'virtual' location. Towards the end of 1980s, the penetration of portable computers, particularly laptops, led to the evolution of Computer-Assisted Personal Interviewing (CAPI). The 1990s witnessed the commercialization of internet that is considered to be the fastest growing technology in the world (Taylor, 2000). The potential of this revolutionary technology was harnessed in Computer-Assisted Self Interviewing (CASI) in the 1990s. Since their inception, CATI, CAPI, and CASI have each undergone various developments that have transformed them into more sophisticated methods. The major changes in these computer-aided data collection methods can be attributed to the development of graphic user interfaces (such as Windows and the World Wide Web) and the subsequent growth in multimedia computing.

Global Positioning System (GPS), a satellite-based positional system, was initiated by the U.S. Military in the 1970s and achieved full operational capacity in 1995. With the rapid improvement in GPS technology, it was quickly adopted by survey researchers as a relatively low-cost, high accuracy solution for meeting various positioning requirements. Beginning in 2007, the GPS technology further improved resulting in a shift from GPS devices with limited portability to wearable GPS devices that could obtain more accurate and detailed location-based information.

The purpose of this paper is to document the status of data collection technology and its application in the area of travel behavior research. Its primary focus is on current data collection technologies, although some thought is given to emerging technologies which may influence the gathering of travel behavior details in the future. The paper is structured as follows. Section 2 discusses the past and present data collection technologies. These technologies are divided into GPS and non-GPS technologies, with example applications from household travel surveys, commercial vehicle surveys, on-board transit surveys and other methods currently in use in the United States. Section 3 discusses the future data collection technologies that can impact travel behavior research. The final section concludes the paper and presents suggestions for areas of further research.

2. Past and Present Data Collection Technologies

The first travel surveys in the United States began in 1950 and were administered by paper and pencil using interviewers who went door to door to survey the households in person. By the 1970s, the surveys were conducted more often by telephone, but still were using paper and pencil techniques. It wasn’t until the 1990s that Computer-Assisted Telephone Interviewing (CATI) technology was introduced. At that point, travel survey methodology started to transition to multi-stage surveys: telephone the household/business to obtain demographic information, mail travel logs to the participants to record travel on an assigned travel day, then telephone to retrieve the details or have the participants mail them back.
The content of the travel surveys became more intricate and the level of quality of the data obtained improved as advances in travel demand modeling placed increasing demands on the data. This increased demand for higher quality/finer resolution data resulted in the introduction of other technologies such as Global Positioning Systems (GPS), web-based surveys, use of bar codes, and scanning. Each of these is discussed in the following sections in context of the surveys that have most benefitted from the technology.

2.1 GPS Technology

Global Positioning Systems (GPS) technology was first introduced into the travel survey arena with the 1996 proof-of-concept Lexington Area GPS survey. Funded by the U.S. Department of Transportation, the purpose of this survey was to test the possibility that trip-based data could be obtained through GPS rather than travel logs. The overall goal was to identify a cost-effective alternative to the trip diaries, which are time-consuming to the respondents and as a result, not always fully completed in terms of all trips traveled and/or the geographic details of the destinations visited.

Since that time, GPS has become a commonplace travel survey technology. In household travel surveys, it has been used to audit trip reporting levels to improve the overall collection of travel behavior details (Bricka and Bhat, 2006), and in some areas replaced the trip diary entirely. In commercial vehicle surveys, GPS helps to identify actual routing of the vehicles as they accomplish their daily deliveries. For on-board travel surveys, GPS is used most often to identify boarding locations (thereby reducing respondent burden and increasing accuracy for a critical data element). More details regarding each of these applications of technology are provided in the following sections.

2.1.1 Household Travel Surveys

The first application of GPS technology in household travel surveys was in the 1996 Lexington Proof-of-Concept study funded by the US Department of Transportation. According to Murakami and Wagner (1999), the objectives of the Lexington study were both to test the equipment and to assess the willingness of the general public to participate. The results found that participation was positive, but several equipment improvements were identified (items that have since been resolved by newer technology and/or changes in methods). In comparing the self-reported data with that obtained from the GPS units, this study found that both trip distances and durations differed greatly between the two data sources.

Despite the positive results from the Lexington study, the use of GPS technology in household travel surveys began with the purpose of auditing trip rates. The hypothesis was that participants did not record all their trips (some overtly, others
due to a lack of understanding about what was to be recorded). In 1997, the Austin (TX) Household Travel Survey was the first full-scale effort in the US to equip vehicles with GPS devices to capture vehicular movement on the assigned travel day. Following that study, almost 20 such efforts have been conducted in the US.

Beginning in 2007, the GPS technology improved to the point of smaller units that collected more data. This resulted in a shift of focus from in-vehicle movement to wearable GPS devices that could obtain travel behavior across all modes of travel. The first wide-scale application of wearable GPS within the context of a household travel survey was in the Chicago Regional Household Travel Survey (NuStats and GeoStats, 2008). Here, all members over 15 years of age of 160 households that reported a tendency towards non-motorized travel were equipped with wearable GPS units for up to a week. Analysis of that data is still ongoing, but preliminary results suggest that it provides rich insights into non-motorized travel and transit usage in the Chicago metropolitan region (NuStats and GeoStats, 2009b).

To date, most of the GPS studies in the U.S. have been ancillary to the main diary-based household travel surveys. However, two studies in the U.S. have started to explore GPS-only collection of travel: a pilot test in Oregon and a full-scale regional effort in Ohio.

In the 2005 Oregon effort, households were randomly assigned to a traditional survey approach, a traditional approach with GPS to audit trip rates, and a GPS-based collection effort (using rudimentary wearable GPS devices). According to Bricka et al. (2009), the objective of this pilot effort was to assess differences in respondent burden, completeness of travel details obtained, and costs. Results showed differences in all three areas as well as variances in non-response bias detected in each effort but confirmed the general thought that GPS was an effective tool for capturing travel behavior details for specific population sub-groups (particularly the elusive young adults). It also showed, however, a technology bias against low-income and less-education population subgroups – a concern echoed in Bricka (2009).

The second GPS-only effort in the United States is underway now in Cincinnati, Ohio. As documented in Giaimo and Anderson (2009), Giaimo et al. (2009), and NuStats/GeoStats (2009b), the study began its year-long data collection effort mid-2009 to collect travel details for 4,000 regional households. In it, all household members over 12 years of age are provided wearable GPS units and asked to carry them for 3 days (children 12 and under are provided simple diaries to track travel). The trip file is created through an imputation process and/or a prompted-recall survey (where the GPS streams are fed back to the respondents for confirmation/clarification).
2.1.2 Commercial Vehicle Surveys

For commercial vehicle travel, GPS technology provides the opportunity to understand travel routes, in addition to the destination and travel time details foundational to travel demand modeling. In addition, accurate vehicle dwell times can be collected, to help with evaluating commercial vehicle-related emissions, cold start times and locations, and idling times. At this time in the US, there are no commercial vehicle studies that have employed GPS technology for data collection. However, studies have been reported in Canada (Sureshan, 2008) and in Australia (Greaves, 2008).

2.1.3 On-Board Travel Surveys

In the United States, most on-board travel surveys are conducted by distributing surveys to respondents as they board a specific bus and asking them to self-complete their survey. Key data elements include the boarding and alighting locations and times, as well as the overall origins and destinations of travel. With on-board travel surveys, the benefit of GPS is in obtaining the boarding location. Surveys are pre-printed with barcodes. Surveyors will scan the first and last barcodes of surveys distributed at each stop, and set a GPS point to link those surveys to the bus stop. This records the place and time of boarding the vehicle. It helps to ensure that a critical data element is accurately obtained and, in addition, reduces respondent burden, thereby increasing overall response rates on the survey. This technology has been in use by NuStats for their on-board surveys since 2007.

2.1.4 Other Uses

As documented in NuStats and GeoStats (2009a), GPS can be used to provide data in support of travel demand modeling in the following ways:

- Objective Measures of Trip Rates, Trip Distances, and Trip Times.
- Route Choice and Activity Space Analysis
- Model Network Development
- Congestion Management Planning
- Emissions Modeling
- Physical Activity/Health Research
- Intelligent Transportation Systems/Operations

2.2 Non-GPS Technologies

Several computer-aided technologies have made a profound impact on transportation survey research beginning with CATI (Computer-Assisted Telephone Inter-
viewing) that originated in 1970s, CAPI (Computer-Assisted Personal Interviewing) in the late 1980s, and CASI (Computer-Assisted Self Interviewing) in 1990s. Computer-aided data collection methods offer lower costs, quicker data turnaround time, and improvements in data quality as compared to traditional paper-and-pencil methods.

Computer Assisted Interviewing (CAI) methods have supplemented or replaced the traditional paper-based methods, with demonstrable gains in data quality and reduction in measurement errors resulting from the following factors. First, use of questionnaires that have been pre-programmed and thoroughly tested reduce human errors such as those associated with incorrect or missing skip patterns, branching, etc. The reduction in errors is more noticeable in surveys with complex questionnaire designs. Second, the interviewing software used for CAI reduce the effects of questionnaire design on data quality by providing tools that allow for more complex and scientifically correct designs. This includes tools for complex branching and skipping patterns that are invisible to the eye of the interviewer. Third, real-time monitoring of data quality facilitated by pre-determined, in-built validity and consistency checks minimize discrepancies in the dataset by helping interviewers recognize and correct data input errors. For instance, in-built logic checks in household travel surveys minimizes a wide range of errors such as trip-sequencing errors associated with invalid or incorrect trip lengths or activity durations at stops made by the members of the household over a 24- or 48-hour period. Finally, CAI methods enable the availability of raw data immediately after data collection for preliminary analysis. This is especially useful for examining the representativeness of the travel survey data in terms of the diversity of key demographic groups and travel behaviors of interest; and monitoring the progress towards survey goals, generally defined by key geographies, socio-demographic, or travel behavior characteristics of respondents.

While there are several advantages of using CAI methods, there are cost implications. As expected, the CAI methods are associated with high front-end costs because they integrate data collection, data entry, coding, and cleaning into a single process. Compared to the front-end costs, the post-interviewing costs associated with processing and cleaning the data are relatively less. For large household travel surveys, the cost savings from reduction in data processing costs and improved data quality outweigh the higher front-end set up and data collection costs, thereby making this a cost-effective method to collect data for large surveys. The following sections provide more details regarding each of these technologies and their application in travel behavior research.

2.2.1 Computer-Assisted Telephone Interviewing
Computer-Assisted Telephone Interviewing (CATI) is the oldest and most common method of computer aided interviewing that has been a key area of development in travel survey research. It is an interactive computer system that enables the interviewers to collect data over telephone and capture the data electronically in real time (Nicholls and Groves, 1986).

Before the introduction of CATI, in-person home interviewing or telephone interviewing was conducted using pencil and paper methods to collect travel behavior data¹. The quality of travel survey data collected using paper and pencil methods was heavily dependent on the interviewer’s ability to interact with people, knowledge of the transportation system, familiarity with the study area, and understanding of the survey questions. As a result, the data underwent extensive processing including data entry, cleaning, coding, validation, and imputation of missing data. Compared to paper-and pencil methods, CATI system offers a more productive, cost-effective and time-efficient solution that enhance and control survey data quality.

In the beginning, CATI would be employed at a centralized location using a computer system with interviewers asking questions that appeared on their computer terminal and typing in the answers provided to them by the respondents. With the advancement of computers from minicomputer system to microcomputer network, increase in the computing power and the expansion of the capabilities of the CATI software over the past several decades, the administration of more complex questionnaire designs has been made possible. Furthermore, the advances in computer technology have also effected the decentralization of CATI system enabling interviewers to plug into the CATI system and interview respondents remotely from their homes.

Though CATI originated in 1970s, it was introduced in travel survey research in 1990s. The earlier CATI programs could handle basic capabilities including simple branching, skips among questions, and real-time validation of data based on simple, in-built logic checks. In comparison, the latest versions of CATI program provide several advanced features that have not only extended their capability to handle complex travel behavior surveys but also enhanced the data quality. These features are being extensively used in recent travel behavior surveys.

1. Ability to construct customized lists of questions for each respondent based on previous responses. This allows for higher degree of personalization of the interviews for each respondent and is useful for Stated Preference (SP) surveys. For instance, SP congestion pricing surveys utilize this feature to generate hypothetical scenarios that vary depending upon the characteristics of the respondents.
2. Capability to construct rosters wherein a series of questions are asked for each member of a list. This feature is extensively used in household tra-

¹ Prior to 1980s, all household travel surveys used home-interviewing with paper and pencil methods for data collection. In 1980s, household travel surveys started using telephone interviewing with paper and pencil methods, as seen in the 1980 Bay Area Travel Survey conducted by Caltrans, and 1981 travel survey conducted by the Bay Area’s Metropolitan Transportation Commission.
vel surveys for collecting detailed information on all persons in the households and all trips/activities for every person.

3. Ability to randomize choices within questions, questions within blocks of questions, and blocks of questions within questionnaires. The randomization eliminates question order or format effects and is used in attitudinal travel surveys wherein information is collected to understand the perceptions and attitudes of travelers.

4. Flexibility to allow for complex checks needed for validating location-based data such as home locations, trip start and end locations; and non-location data such as trip/activity sequence, travel mode sequence etc. The CATI tool facilitates the accurate identification of trips origins, trip destinations; and access and egress modes, using spatial information (such as GIS layers of the transportation system, street network, points of interest etc) to enable subsequent geocoding of trips. This feature is being utilized in recent household travel surveys.

5. Creation and administration of multilingual surveys wherein interviewers can conduct the interview in multiple languages. To our knowledge, most household travel surveys conducted in the last decade have used this feature to administer Spanish versions of the survey to collect information from Spanish-speaking, linguistically-isolated households.

The CATI approach fosters a common cognitive space between the interviewers and the respondents in order to facilitate the exchange of information between interviewers and respondents, and to lessen the respondent burden (Brog, 2000). This contributes towards increasing the productivity of the interviewers that translates to savings in cost per completed survey.

Despite the advantages, CATI surveys are experiencing rapidly falling response rates (Owen, 2002; Curtin et al., 2005). The decline in response rates and the associated increase in the cost of conducting the survey have been attributed to technologies such as answering machines and caller ID. Also, the exponential rise in the ownership and usage of cell phones, and increasing growth of cell-only households present a challenge to the traditional households telephone surveys (Kempf and Remington, 2007; Sen et al., 2009).

From a cost standpoint, studies conclude that CATI surveys are cost-effective only for large scale or regularly repeated surveys (Leeuw et al., 1995) and this again is affected greatly by technological advances. For example, Weeks (1992) indicated that a break-even point while considering CATI was at about thousand interviews. Today, CATI is used for all surveys, even small surveys of 100 households.

2.2.2 Computer-Assisted Personal Interviewing
Computer-Assisted Personal Interviewing (CAPI) is a widely used data collection method that enables interviewers to conduct face-to-face interviews using portable computers such as laptops. Following the interview, the data are sent to a central computing network, either electronically via modem or internet; or by mail (Baker, 1992; Saris, 1991).

CAPI originated in the 1990s with the extension of the use of computers to face-to-face interviewing made possible by the advent of portable laptop computers. CAPI offers all the advantages of computer-assisted interviewing methods over traditional paper-and-pencil techniques. These benefits translate to cost savings in data processing, improved quality of data, and shorter timelines for completion of the survey. Notwithstanding the advantages, there were two major drawbacks in using this method. First, the front-end costs associated with laptops and CAPI software combined with the in-person interviewing costs are higher than for CATI interviewing. Second, the limitations in the portability of laptops and their relatively short battery life can complicate the logistics of in-field data collection. Consequently, CAPI is best used as a complement to the mainstream CATI interviewing.

The next generation of CAPI interviewing is called “Mobile Computer-Assisted Personal Interviewing” (MCAPI). This technology combines the quality aspects of the CAPI with the proliferation of handheld computers such as PDAs to provide a more accessible and cost-effective in-field interviewing technology. Handheld computers are portable, inexpensive, and energy efficient, making them invaluable field instruments as they overcome the limitations of the traditional CAPI laptop technology. While the limited storage capacity and screen size has restricted its application to short surveys such as household screening in large scale surveys (Bosley, Conrad, and Uglow, 1998; Nusser, Thompson, and Delozier, 1996), this technology shows great promise for in-person/in-field interviewing. In addition, the cost issues with MCAPI are more associated with the labor-intensive in-person interviewing rather than the technology itself.

In the area of travel behavior research, MCAPI technology is being increasingly used in stated preference travel surveys (Ampt, et al., 1987; Jones et al., 1989; Jones and Polak, 1992; Polak et al., 1993). Specifically, this technology has been found useful to capture the stated preferences of hard-to-reach population segments such as visitors. In addition, MCAPI technology can be used in non-response follow-up studies that are geared towards collection of socio-demographic and travel behavior information of the households that did not respond to the main survey, wherein the interviewers failed to reach them via telephone, mail, or web.

2.2.3 Computer-Assisted Self Interviewing

Computer-Assisted Self Interviewing (CASI) is a data collection technology where computers take the place of interviewers in guiding the respondents through the questionnaire. Specifically, the surveys are self-administered by respondents
who read the questions on the computer screen and enter the responses. Computer-Assisted Web Interviewing (CAWI) is the most popular type of CASI.

CASI has several benefits over CATI and CAPI (refer Bronner and Kuijlen, 2007 for a detailed review). CASI provides the respondents more privacy and anonymity. With the lack of interviewers, the respondents are more comfortable in reporting sensitive information such as household income (Kreuter et al., 2008). In general, the level and accuracy of reporting sensitive information has been found to increase with CASI (Tourangeau and Yan, 2007). Several studies show that respondents are more likely to report socially unacceptable perceptions, attitudes, and behaviors (Newman et al., 2002). Specifically, self administration eliminates the interviewer’s ability to influence responses or register them according to personal biases. This removes the effect of interviewer bias that is often associated with CATI and CAPI surveys. For instance, Poynter and Comely (2003) found that 64% of the respondents contacted using CASI reported using a mobile phone while driving as compared to 42% contacted using CATI. The study conducted by Taylor et al. (2005) revealed that online respondents were more likely to disclose socially undesirable behaviors as compared to CATI respondents who were more likely to admit to socially desirable behavior. According to the findings, higher number of online respondents reported that they frequently drive over the speed limit. In comparison, higher number of CATI respondents admitted going to church, mosque, or synagogue.

Several studies suggest that data collected using CASI is of higher quality than CATI or CAPI methods due to the flexibility of completion at convenient times and elimination of time pressure giving the respondents more time to think through the questions before answering (Oosterveld and Willems, 2003; Dillman et al. 2001; Kwak and Radler, 2002). These studies suggest that the CASI respondents were found to elicit more in-depth open-ended responses, make less use of extreme categories of response, and answer more number of questions as compared to CATI or CAPI, which makes a considerable impact on the data quality. The ability to use graphical pictures with CASI to add more clarity to the questions is an added advantage and is ideal for transportation research. In addition to improved data quality, CASI is significantly more cost-effective.

However, the data collected using CASI methods is prone to technology bias and self-selection bias – a major drawback of online survey samples. Also, CASI methods are typically associated with low response rates. With the ever-increasing number of people with internet access, studies have shown a growth in response rates (Manfreda et al., 2008, Kaplowitz et al., 2004). Nevertheless, it is important to note that the response rates are closely associated with the survey length. Hence, CASI methods are rarely used in household travel surveys due to the length and complexity of these surveys, but are being increasingly used for stated preferences studies in travel behavior research.

In the past decade, several other technological advancements have led to advancement of CASI methods. In particular, there has been a steady growth in multimedia technology resulting in improvements of audio and video communication. This has made a significant impact on survey data collection using CASI, begin-
ning with the development of audio-CASI followed by video-CASI. With Audio-
CASI, respondents listen to a recorded voice and respond to the questions using
keyboard entry. Video-CASI – an extension of Audio-CASI – also uses video
communication. The variants of CASI have shown demonstrable gains in the qual-
ity of data collected on a variety of sensitive topics (see, for example, Turner et
al., 1998). To our knowledge, CASI variants have not yet been used in travel sur-
vey research. The availability of low bandwidths limits the use of multimedia in
self-administered surveys on the web. As the bandwidth increases, we can expect
to see a wider range of application of CASI variants in travel surveys.

2.2.4 OCR, ICR, and OMR Scanning

Automated data entry technologies are extensively used for processing question-
naires to minimize human errors resulting from traditional data entry methods.
Before these technologies can be used, questionnaires are scanned using an optical
scanner to produce an image file and software is used to process the image file to
differentiate between images, text, checkmarks, etc. Note that these scanners can
be upgraded to enable automatic identification of bar codes – a frequently used
feature to distinguish between surveys.

There are three types of automated data entry technologies: Optical Character
Recognition (OCR), Intelligent Character Recognition (ICR), and Optical Mark
Recognition (OMR). Following the scanning of questionnaires, OCR technology
can recognize a wide variety of fonts in different languages but handwriting is still
problematic. ICR technology can recognize and convert handwritten text and
numbers into machine readable strings and documents. OMR technology can re-
cognize optical marks or checkmarks to response categories of questionnaires. It
decreases the time needed to fill out the questionnaires and improves processing
accuracy because a machine can read the marks quickly and accurately.

These technologies are cost-effective and save time. Consequently, they are be-
ing widely used to electronically enter data. In the area of travel behavior research,
the OMR technology is commonly used to automate the data entry for self-
administered, transit on-board surveys. OCR technology is commonly used for
transcribing electronic license plate captures for cordon-line or external station
surveys.

3.0 Future Data Collection Technologies

In addition to the advances in CASI interviewing mentioned above, dramatic ad-

dances in cellular telephone technology are expected to strongly influence data
collection over the next decade. This includes Smartphone, Text Messaging, and
Android technologies.
Smartphones are on the verge of becoming one of the most powerful tools for data collection. The rapid growth of Smartphones can be attributed to enhanced features and user interfaces, increasing penetration of faster wireless broadband networks, increased screen size, and more attractive pricing. Smartphones leverage the latest cellular technology and wireless internet capabilities, providing a wide range of features such as texting messages, connecting to the Internet, use of third-party applications, and GPS.

In particular, GPS-enabled Smartphones are an emerging data collection technology that holds great promise for the area of travel behavior research. These devices have been known to provide accurate information on location and traffic data such as travel time, speed, acceleration, and direction of travel. Several traffic studies have demonstrated the potential of GPS-enabled Smartphones. These Smartphones provide exhaustive spatial and temporal coverage of the transportation network that help monitoring traffic and computation of reasonable estimates of travel time with GPS-level accuracy (Fontaine and Smith, 2007; Yim and Cayford, 2001; Yim, 2003). However, concerns with the use of this technology such as communication load (as this method requires sending heavy loads of information to a centralized location), handset energy consumption, and privacy (due to knowledge of the exact location of the respondent) need to be addressed for a wider use of this technology (Herrera et al., 2009). The GPS-enabled Smartphones can be used as powerful tool to get accurate location information on the trips made by the respondents.

Text Messaging is a widely used feature of cellular phones, particularly by the younger population, to send short text messages. Though text messaging is currently limited to 160 characters, it is not dependent on making direct two-way contact with the respondent. Thus, it can serve as a powerful survey tool for short, frequent data collection or for inviting respondents to complete surveys at their convenience (Callegaro, 2002). In the context of household travel surveys, text messaging can be used as a cost-effective way to coordinate a time with the potential respondents for the recruitment or retrieval phone calls. Texting may also be used to issue reminders about an upcoming travel day to increase respondent participation in travel surveys.

Android, the open-source mobile operating system championed by Google, can enable developers to build innovative mobile applications or incorporate new cutting edge data collection technologies as they emerge. For example, an application can call upon any of the phone's core functionality such as making calls, sending text messages, or using the camera, allowing developers to create data collection platforms. Android provides free access to a wide range of useful libraries and tools that are already being used world-wide by researchers to build rich applications for inexpensive ways of data collection in developing regions. Seattle's Grameen Foundation Technology Center is using Android to determine if the Ugandan rural health and agriculture text-message tip service is having an effect on the communities it aims to serve (Grameen, 2009; University of Washington, 2009). The University of California, Berkeley's Human Rights Center is using it to record human rights violations in the Central African Republic (University of Washington-
Android holds great potential to turn a cell phone into a versatile data-collection device.

4.0 Conclusions and Future Research

The past decade has been rich in technological advances and many of those have directly improved data collection for travel surveys. From the age-old standards of CATI and CAPI interviewing to the rapid advancements in GPS and cellular phone technologies, the main challenge is in recognizing and harnessing the power of these data collection technologies.

CATI technology will likely continue to provide the foundational interviewing technology for travel behavior surveys over the next few years. However, as web-based survey designs evolve to address self-selection and technology biases, the low-labor cost aspect of that technology will likely put it in the lead over CATI and CAPI options. The challenge here is in designing a web-based survey that can obtain the same level of detail as an interviewer-mediated interaction.

GPS technology has been the industry’s favorite in terms of alternatives to the traditional diary-based travel behavior designs. Its portability, low-respondent burden, and newness makes it exciting, even though to date (in the U.S.), there has not been a full-scale regional model built using data obtained solely from GPS. A future area of research is the integration of GPS with other technologies (such as the Smartphones) so that new applications can be developed to leverage these developments to capture travel behavior data.

Along with these technologies, there are two methodological design issues that much be considered. The first is the representativeness of the surveys obtained through these new technologies. Marred by self-selection and technology bias, transportation professionals need to identify techniques to minimize these biases in order to achieve more representative surveys. The solution may be in deploying a combination or blend of technologies appropriate to the sub-population groups of interest to result in a representative data base.

The second design issue is that of how the data collected through the newer technologies can be used in travel demand modeling. The traditional methods of CATI, aided by CAPI, CASI, and GPS, provide data that is more in-depth and higher quality than ever before. The newer GPS-only studies by design do not obtain the same level of detail as the CATI-focused designs. Is that acceptable? Are there alternatives that provide the same level of detailed data?

Whether its incremental advances related to improvements in existing technologies, or landmark changes in data collection technologies in response to new technologies that enter the market, transportation professionals should be monitoring the developments in order to identify possibilities for improving data collection technologies. Academic researchers can aid in the process by testing the application of these new technologies, while those in the non-academic environment can work to improve the implementation of the technologies. As much as tech-
ology has changed in the past decade, the next five years holds the promise of wide-scale changes and improvements. The task for us is to capture and shape those technologies to meet the needs of travel demand data collection.

References


NuStats and GeoStats (2009b). GPS in Other Studies. Technical Memorandum 2.2.2. Prepared for the Denver Regional Council of Governments as part of the Front Range Travel Counts project, October.


