Chapter 3

Cell Phone Enabled Travel Surveys: The Medium Moves the Message

Jane Gould

Abstract

Purpose — To assess how cell phone technology might impact the collection of travel data in the future.

Design/methodology/approach — Two different types of cell phone enabled studies are considered. First, we examine how the text feature of phones can be used for person-to-person surveys, and second, we explore an aggregate level survey enabled by an anonymous and passive GPS trace.

Findings — This study explores the types of travel information that are likely to be inferred from text surveys and cell phone traces. It recognizes that a passive GPS trace might change the level of measurement and the inferences we make about travel behaviors.

Research limitations/implications — The study is prospective. It anticipates that over the next 10–15 years cell phone tracking technology will improve, as well as the speed and capability of algorithms for post-processing the information.

Practical implications — Cell phone enabled studies may provide a new tool and new level of measurement, as traditional survey response rates decline, and it becomes more difficult and expensive to conduct conventional travel surveys. The capacity of cell phones for travel survey work is improving, but it is not fully realizable today (2012).

Originality/value — This study provides a context to understand how the technology of the cell phone might be integrated with more traditional travel

Transport Survey Methods: Best Practice for Decision Making

ISBN: 978-1-78190-287-5

surveys to streamline data collection, and produce new types of spatial detection, measurement, and tracking.

Keywords: Cell phones; GPS; passive tracking; future surveys; mobile phones; small screens

From the time that cell phones were used chiefly by "early adopters" they have played a significant role in travel and mobility. Initially, cell phones helped people to coordinate their meeting points during travel. In 2003 a Finnish researcher observed that cell phone users communicated their present location (e.g., "I'm at the train station") about 70% of the time, compared to just 5% for landline users (Oulasvirta, 2005). With the addition of Internet service, the cell phone has become an indispensable tool for mobile route planning and navigation. Soon, a next generation of phones, equipped with GPS, will transform how we coordinate with others, make travel payments, and when and how we travel. It will also transform how transportation researchers seek to study and measure travel behaviors.

The process for gathering data is evolving from "intrusive" surveys that take place on the web, phone, in-person, or mail to methods that collect data electronically and passively. Until recently, GPS (with accelerometers) were a stand-alone and separate recording device, but they are increasingly becoming a standard feature of mobile cell phones. The mobile cell phone is a "pattern break" of singular importance to survey designers. The act of recording travel, like any experiment, may interact with and change the underlying behavior. Stand-alone GPS devices, while small, are a reminder that behaviors are being "recorded"; researchers take pains to say that the respondents quickly learn to disregard the presence of the recording devices. Assuming this is so (it is a difficult proposition to test), the separate recording device has an additional drawback: it is difficult to recruit a truly random sample. Stopher (2009a) for example, reported from a 2007 study that in the few instances where a GPS was used as a stand-alone method for travel measurement, the response rates for recruitment were similar to those for conventional surveys.

The timing for passive electronic data collection is opportune, for the foundation of transportation models is built on the representativeness and accuracy of survey data. Stopher (2009a) notes the irony between the precision these models require and the measurement error introduced from factors such as declining response rates, increasing globalization of the population, declining literacy, and respondents' who view surveys as irrelevant. Among those who do participate, survey designers confront respondents who seem to have shorter attention spans, less "investment" to report detailed trips, and more missing trip data overall.

In this chapter, "passive GPS" describes a future data collection using the mobile phone instead of a stand-alone GPS device. When incorporated into a phone, the passive GPS acts as sensor in the background. There is no active decision to use it, other than to turn the telephone on, and carry it. Unlike a stand-alone GPS, people seldom forget to charge it, and seldom leave home without it. Today, stand-alone GPS devices offer superior accuracy but the precision from cell phone traces is improving.[1] An assumption underlying this chapter is that, like Moore's Law of computer chip memory,[2] there will be advances to both the algorithms, and to cell phone operations, like battery life, memory, CPU usage, and "time-to-first fix" (Reddy et al., 2010). Another assumption, the more fundamental one, is that the cell phone will evolve as an even more essential "wearable mobile computer."

We posit that GPS tracking on cell phones will fundamentally and substantively change not only how we collect our data, but also what transportation researchers seek to know. A preview of the change is found in this chapter's literature review, which draws from several fields. "Location-based studies," which use the cell phone for tracking, record data at a system-wide or aggregate level. This is somewhat like beginning with the answers to a survey, but then working back to deduce the questions that were posed. In 2006, Ratti, Pulselli, Williams, and Frenchman noted that aggregate level tracking portended significant change for traffic engineers: cell phone traces could reveal in real time the actual patterns of movement, in lieu of simulating them through models or estimates. In conventional travel surveys there is considerable "noise" in the data, as validity depends on factors like the representativeness of the sample, and then on the integrity and veracity of self-reports. With a passive GPS trace, the "noise" from survey measurement and error are minimized. The trade-off is that as we reduce the "noise" from survey measurement the "signal" or data is modified. We consider how the collection of real-time tracking may change measures like trip frequency, mode-share, and activity coding.

Recognizing that there will still be a need, albeit reduced, for studies that take place at the level of the individual respondent, we also study the applicability of cell phones to send and receive text surveys. Again, the concept of signal to noise ratio is useful. The small screen size and limited attention of the user suggest that cell phone enabled surveys will be shorter and more superficial, hence "noisier." However, there are also new types of content, that is, "signals" created by novel features like picture-taking and real-time transmission.

Organization of this Chapter

1. We briefly review some key studies from the growing literature on mobility and cell phone applications. Much of this work has taken place in computer science and media studies. This will establish the groundwork for the next section.
2. We examine emerging issues for text-based surveys that will be sent and received over cell phones. While the need for these surveys is likely to be diminished, there may be inventive ways of collecting survey data using smart phones.

---

1. As of 2011 most cell phone tracking studies use a less precise method: estimating location by triangulating the distances between cell phone tower locations.
2. The point is that the technologies will advance — exponentially. Moore's Law — linked to digital electronics — observes that the capacity of microchips has doubled every 18 months, and grown in order of magnitude every 5 years (cited in Oulasvirta, 2005).

3. Finally, we assume that the cell phone trace replaces many conventional travel studies. What data elements might be improved, and what information (or signals) is entirely new?

3.1. Literature (Cell Phone Tracking)

New technologies are often packaged in familiar ways: the initial cell phone-trace studies were similar to those with stand-alone GPS devices. In a comprehensive review of GPS, Chorus and Timmermans (2010) attribute the first cell phone study (with a PDA) in 2005 in Japan (see, e.g., Asakura & Hato, 2004), and they cite five additional cell phone trials through 2009. By 2009, there is recognition that the cell phone data has new features and that their ‘‘location-based services’’ may provide a different vantage. In 2005, Asakura, Hato, and Sugino discuss the usefulness of tracking tourists, pedestrians, and cyclists using ‘‘dot data analysis’’ that has both space and time dimensions. Instead of conducting personal surveys of intentions and activities, they propose studying movement at the micro-movement level and then generating thousands of profiles for a simulation model.

A related thread of studies grew out of the computer science and communications field. Eagle and Pentland (2005) proposed that location aware devices, specifically cell phones, could collect data that focused less on the micro-movement of individuals and more on patterning in the collective or aggregate. They observed:

> Surveys are plagued with issues however, such as bias, scarcity of data, and lack of continuity between discrete questionnaires. It is this absence of dense, continuous data that also hinders the machine learning and agent-based modeling communities from constructing more comprehensive predictive models of human dynamics. Over the last two decades these has been a significant amount of research attempting to address these issues by building location-aware devices capable of collecting rich behavioral data.

There were several early demonstrations of this concept particularly by Ratti et al. (2006), Reades, Calbrese, and Ratti (2009), and Reades, Calabrese, Sevstuk, and Ratti (2007) who worked with anonymized records from mobile operators in Italy, and mapped mobility patterns in Milan and then Rome, by capturing the latitude and longitude, and time of day, when calls or texts were engaged. These researchers observed that a more sophisticated procedure would allow the continuous tracing of users' movements at regular intervals, *throughout* the day.

Another early study acknowledged that mobile phone records, which routinely record the location of millions of cell phones, could be mobilized in emergencies, as well as for routine operations by traffic engineers and public works (Madey et al., 2007). This paper demonstrated an algorithm to detect the mode difference between walking and in-vehicle travel.

In 2008 Gonzales, Hidalgo, and Barabasi published a landmark study, based on the analysis of one hundred thousand mobile phone user records, over a six-month period. The research reconstructed travel trajectories, by recording cell phone tower locations whenever those selected in a subsample initiated or received a call or text. They observed a high degree of regularity in the daily travel patterns, captured by a high return probability to a few highly frequented locations (e.g., home, work). These findings were surprising to the researchers, as they studied outside the transportation field.

Over the past year or two, there are several studies that demonstrate that cell phone traces correlate closely with travel statistics, like mode choice and volume of travel, collected by the US Census. There have also been more studies that demonstrate improvement in the post-processing algorithms, and the ability to reproduce a travel route (Bierlaire, Chen, & Newman, 2010; Chen, Newman, & Bierlaire, 2009). A small, but promising study by Reddy et al. (2010) tested a mobile phone equipped with a GPS receiver and an accelerometer. They coded travel mode over 124 hours of data and 16 users, and did not do any user-specific training. They were able to detect with a 93.6% accuracy rate, differences between being stationary, walking, running, biking, and motorized travel.

In the last year, there have also been some novel applications to transportation issues. Becker et al. (2011) analyzed Call Detail Records (CDR) from 35 cell towers (and 300 antennas) within a five-mile radius of downtown Morristown, New Jersey. The team studied CDR patterns to identify travel information not reported in the US Census, like the inflow to the city after commute hours, and the source/destination of tourists and occasional visitors. They suggest this data would "be useful to urban planners for example to reduce traffic congestion by organizing new transit and bike routes or park and ride programs … also (to place) late night shuttle buses to keep inebriated drivers off the road." Privacy issues were addressed by stripping the CDRs of identifiers with the exception of the home zip code.

A similar study was conducted by Isaacman et al. (2011). They also used anonymous CDR (not a continuous GPS trace). They mapped home to work commute trips in both New York City and Los Angeles, and other mobility patterns. They cite a numerical consistency between New York and Los Angeles for the number of "important places" visited, which is similar to the finding in Gonzales et al. (2008). They also try to estimate, from their aggregate level data, the carbon emissions but end up having to assign trips to modes.

A recent study, where the algorithms are constructed from the GPS trace (not CDR) is reported by Van der Hoeven (2010). The Dutch navigation firm, *Tom Tom*, developed an application that tracks travel congestion on major roads in the Netherlands by positioning in real time, all 4 million subscribers to Vodafone's mobile phone services. Van der Hoeven (2010) observes that if all mobile carriers provided data, there would be an accurate and continuous survey of travel demand since nearly everyone in Holland carries a cell phone. At an airport, for example, continual tracking would provide insight into visitors' route choice and mode.

Like the *Tom Tom* service, there is an increasing number of field tests and apps that provide traffic information and transportation-based services that are personalized to the individual users or their mobile device (Bayir, Demirbas, & Cosar, 2010; Manasseh, Aherh, & Sengupta, 2009). The latter use triangulation of distances from cell phone towers to determine the user location, but they note that GPS traces are feasible and "are left as a future work."

3.2. Section 2

A smart phone with GPS "leaps over" many hurdles facing survey research. The AU :3 frequency of travel trips can be collected in the aggregate, without identifying the ID of individual respondents or their message content. This helps address privacy concerns. And, travel information can be collected without attempting to contact, and then incentivize to participate, a random sample. Looking forward, we expect there will be fewer "conventional" household surveys done via the Internet, paper, phone, or in-person. In Section 3.3, we consider the implications of this.

Here, in Section 3.2, we look at more conventional surveys that continue to recruit a sample, and in most cases, use incentives for participation. We anticipate that more of these surveys will take place over smart phones or tablet devices. We focus specifically on self-completed surveys, those where respondents link to a web site or survey form, and text back their responses.

The transformation of data collection to the cell phone, from conventional methods, and even stand-alone GPS devices, is a significant pattern break. Table 3.1 groups some of the key changes. The second column of the table is the baseline, AU :4 summarizing data points, which are collected today. Because much of this data is based on self-report, the so-called signal-to-noise ratio is low. Location-based studies, using the cell phone, reduce this noise and collect "outside" data like time, latitude and longitude, and duration. We discuss (in Section 3.3) how this introduces new ways of measuring travel behavior.

Here, in Section 3.2, we discuss studies that use the cell phone to send/receive text surveys. Survey practitioners are concerned that mobile phones produce more measurement error than landlines because there is less fidelity, a shorter attention span, less privacy, and other distractions (Kennedy & Everett, 2011). However, participants accessing the text feature may have different ways to respond. For example, they can provide immediate on-the-spot feedback and deploy distinctly new features: they can troll the net for outside information, be "pinged" in real time to participate, and be prompted to use the camera feature to share visual content. But, unlike the GPS trace, the attitudinal survey conducted via text or voice on a mobile phone is traceable to the individual respondent. Thus respondents will need to be recruited or "opt-in" and they may expect to be compensated for their opinions or time. In the following discussion we elaborate on this. We also consider how specific types of surveys used in transportation, such as the stated preference survey; the survey of attitudes and opinions; and the level of service study, may be modified and perhaps improved when they are sent and received over the cell phone.

Table 3.1: A comparison of data elements from survey modes.

Approach	Design	Sample or population	Duration of survey	Measurement level	Type of variables							
					Mode	Speed	Person traveling together	Route choice	Time of day	O/D	Activities outside home	Demographics (age, income, educ., household)
Cell phone text[a]	Likely to be panel	Recruited	Variable	Within/ Between samples	Self-report	Self-report	Self-report	Self-report	Objective	Self-report	Self-report	From recruitment
Cell phone trace[b]	Likely to be at aggregate level	Community-wide/ Anonymous	Ongoing	Mass movement behavioral patterns	Objective	Objective	NA	Objective	Objective	Inferred	Inferred	NA…May be inferred

[a]Surveys that are sent/received via text on cell phones are discussed in Section 3.2.
[b]Surveys and studies based on a cell phone passive trace are discussed in Section 3.3.

3.2.1. Cell Phone Surveys, Panels, and Non-Probability Sampling

While large household surveys may be expected to decline in number, researchers may seek to learn individual "motivations" by reaching convenience samples, quota samples, and panel studies. Panels offer a particular advantage for transportation studies, because demographic variables need to be collected only once and the panel can be "pinged" or "prompted" on the cell phone to report characteristics of their trip taking over time. But, it is hard to imagine a cell phone sample being selected through probability-based random sampling.

Panels are likely to be identified by address-based sampling or proprietary lists, and then recruited through an e-mail (not voice) contact. Or, respondents may increasingly self-identify an interest through Facebook, Twitter, or other social media and choose to "opt-in" to a posted survey. Either method depends on non-probability based methods for recruitment. Researchers will then need to make post-survey adjustments and weighting, in order to make their inferences more valid and reproducible.

Recently, a comprehensive review by the American Association for Public Opinion Research (AAPOR, 2010) examined the validity of panels. They raised critical concerns about the overall accuracy of the methodology and the ability to compensate for non-probability sampling. One of the underlying problems is: what criteria motivate respondents to participate in panels? The AAPOR report suggests five typical appeals:

- A contingent incentive or prize
- Self-expression (to register one's opinions)
- Fun (entertainment value)
- Social comparison (to find out what other people think)
- Convenience (easy to join and participate)

These five criteria underscore why social media might be a boon for recruiting subjects, and why Facebook, etc. may be so useful. But, thinking about "why" people agree to participate is important. Researchers cannot assume, even if they have a demographically matched panel, for example, that the participants are fair and unbiased toward the topic. Panels put high demands on their respondents, panel responses are seldom anonymous, and there is often a substantial dropout rate over successive waves. According to the AAPOR report (AAPOR, 2010), results of most panel studies are typically not generalizable to the larger population. Moreover, post-survey adjustments, like post-stratifications and propensity weighting rely on assumptions that may increase overall measurement error.

3.2.2. Cell Phone Surveys "At Large": Adapting to the Small Screen

In order to minimize the disruption and obtrusiveness of a telephone call, researchers are likely to contact their panels using text surveys (or prompts), accessed over the phone.

Very little is known today about compressing a survey into the physical space of a cell phone screen. One can speculate that the reduced screen size requires that fewer items are asked, and that both rating scales and write-in text are shorter. Two other likely impacts, also unknown, are primacy and recency: because of the physical layout on the cell phone "page," there may be a tendency to select the "first" or "easy" answers.

Closely related to this tendency, is the more general category of "Satisficing": satisficing focuses on the cognitive effort that respondents devote to the survey process and some believe that respondents take more "shortcuts" in self-administered surveys (AAPOR, 2010; Kennedy & Everett, 2011). Among the shortcuts are answering all questions uniformly (non-differentiation), answering randomly, answering more quickly, skipping items, or "don't know."

Although there is no research evidence about this yet, surveys completed by cell phone would appear to manifest "satisficing" behaviors. The speed of completion for a questionnaire, the ability to use the phone for other activities like web surfing or navigation, and abbreviated text messages might impact survey quality. The "satisficing" behavior might also be more likely among the type of users, for AU :5 example, young and male, who may be willing to participate in a cell phone survey.

However, there are characteristics of cell phone surveys that could provoke a more reflective stance. The first is "real-time" capability: just as respondents can blog or report news as events unfold, they can also provide impressions and rate items in real time. An important, and still untapped, enhancement is how to integrate users' pictures and video into the survey package.

A second enhancement is that respondents can use online capabilities to see their responses in relation to others. Although reflexive feedback was the key feature of the Rand/Delphi Survey (Rand, 1968) it is unclear how this feature might be integrated into new smart phone surveys.

One of the adjuncts to current GPS studies has been the use of non-probability surveys that use the GPS record to augment self-recall and "prompt" respondents to verify or describe their travel behaviors. The cell phone may turn out to be a compatible medium for additional prompted recall studies; respondents can report their behaviors in situ, and report not only what activity they are engaged in, but why they did *not* choose a different one (i.e., the decision to not cycle or walk). AU :6 However, the prompted recall, particularly since it is a "disruptive-survey" by design, faces the problem of survey recruitment, and whether those who opt-in are systematically different. Stopher (2009b) raises an additional concern that prompted surveys, sent to a person engaged in travel, would flag privacy issues.

3.2.3. Cell Phone Surveys and Stated Preference/Revealed Preference Studies

A customized survey that asks respondents why they did/did not engage in a behavior in real time may give new direction to stated preference studies. An often-levelled criticism of these studies is that they are too abstract for respondents and do not capture real behaviors in real settings. There are several reasons for this.

First, for many transportation choices, it may be too expensive and non-productive for travelers to do extensive search. Second, many transportation behaviors are thought to be habitual (Chorus & Timmermans, 2010). Moreover, the process of participating in a stated preference study is taxing, in terms of respondent recruitment and participation.

Surveys taken on cell phones may eliminate some of the participation burden, and make stated preference studies more realistic and concrete for respondents. The stated preference study, somewhat like a prompted recall, can be conducted in real time, while the participant is engaged in the travel activity.

By way of example, suppose a research team wanted to estimate customer demand for a future train route. Researchers might first identify respondents who travel the proposed route today using a different mode, like bus or airplane. In real time, the researcher could send the traveler a customized stated-preference study. This survey could reflect the traveler's current travel time, cost, and service preferences. The survey software, might then ask respondents, based on the trip they are currently taking, to develop their own weights or to name and weight new attributes. This could improve these studies in a way articulated by Chorus and Timmermans (2010), to be able to change the perception of attributes of alternatives.

3.2.4. Cell Phone Surveys of Attitudes/Opinions

Almost all researchers who write surveys for cell phones will encounter the need to write short "haiku" like questions. This economy of words may be particularly troublesome for those who need to explore multiple dimensions of a topic, or tease out multiple factors, say of "like" and "dislike." Because of the limited screen size, one survey question may have to be asked in lieu of several.

Again, by way of example, consider a survey that asks respondents to rank and rate their satisfaction with a travel mode, say light rail. With traditional Internet or paper surveys, a battery of questions about the experience on the rail might be asked, along with a final one about the quality of the trip overall. In data analysis, a factor analysis would be used to identify the important and critical questionnaire items. Researchers using cell phone studies will often not have the luxury to ask multiple questions without burdening the respondent and risking non-completion.

The brevity of the cell phone survey is likely to work against detailed attitude and opinion studies. Short cell phone surveys may not allow researchers to ask the traditional battery of items needed to reliably measure attitudes and a single question or two may not suffice.

However, this problem is superseded by a different challenge to conventional attitude and opinion measurement. Research that began in the 1980s questions the predictive validity of the attitude survey. There is a branch of investigation in social psychology cautioning that attitudes reported in surveys are superficial and not reflective of genuine viewpoints, which are reflected in more subconscious reasoning and reflections (Beatty, 2010). If this research stream is viewed seriously, it upends the validity of doing attitudinal studies with a battery of like/dislike probes.

3.2.5. Cell Phone Surveys and Level of Service Questionnaires

Levels of service (LOS) questionnaires are frequently used in many areas of transportation and they probe conditions, such as whether the service was reliable, timely, clean, and comfortable. Unlike the attitude and opinion survey, they report relatively objective phenomena and might be fairly easy for respondents to report/complete on a small cell phone screen.

However, with a passive GPS and the opportunity for remote sensing, there is likely to be less dependence on these surveys. Sensors are able to update on arrival and departure times; cameras can record conditions, and identify, in the case of transit, load factors.

While the need for traditional LOS surveys may decline, there may be totally novel ways for travelers to report their requirements. For example, a user-based LOS will empower cell phones users to geotag locations where they would like additional public transit service or identify where bicycle lanes are inadequate or failing.

3.2.6. Cell Phone Surveys and the Search for External Records/Demographics

As survey researchers design for the small screen and short format of the cell phone survey, there is likely to be pressure to either recruit more panel studies or shorten the number and type of demographic variables that are asked. Collecting demographic data, like income, age, gender, and household size, is vital to researchers, but is also likely to become more difficult when surveys are ''downsized'' to fit on a smart phone. In order to retain the spontaneity and speed of a mobile survey, researchers may face a trade-off, asking fewer personal and in-depth questions. Moreover, collecting these demographic profiles then creates privacy risks for both cell phone users and cell phone companies.

Opportunely, there is some evidence from transit research that demographic information might be inferred from the GPS trace, and then linked to census data. In addition, in the special case of transit, smart cards can be linked, in theory, to the point of sale, such as a credit card, and then to other external data sources.

Apart from their links to demographic information, smart cards leave a ''breadcrumb trail'' that can also be used for O/D studies, to analyze transfer activity, and to study day of the week patterns. Some transit researchers have suggested that passenger surveys, are no longer necessary because this information can be inferred from the location (Chapleau, Trepanier, & Chu, 2008).

Transit studies provide a hint that as cell phone enabled surveys increase, researchers will seek external sources of data. An initial location-based study by Lee-Gosselin and Harvey (2006) identified unique external data sources such as license plates and electronic tolls. Demographic information is likely to be inferred from geolocations. It should not be overlooked that an external source for geolocation data is often the U.S. Census. While it is deemed to have a high degree of accuracy, census data is a mix of self-report and interviewer data, it is centrally cleaned and standardized, and there are time lags before it is published. The final irony is that census information is itself based on surveys of individuals and households.

3.3. Section 3

Surveys conducted over the phone, on paper, etc. are fundamentally ''noisy'' because their accuracy and reliability depends first on recruiting a representative sample, and second, on the quality and veracity of individual reports. Over the next 10–15 years, what if transportation data was collected primarily by passive GPS? There are many original and new inferences to be made from the passive GPS trace. For example, the time of travel becomes readily available and might provide new insights for models of congestion pricing and environmental impacts. Another novel dimension might be the objective speed of travel, not perceived/reported travel time. In the following discussion, we consider the practice of employing an aggregate level GPS trace in lieu of conventional survey data to infer travel mode, trip counts and trip activities, route choice, and dynamic route change.

3.3.1. Passive GPS and Inferences from Travel Mode: Walking/Bicycle/Vehicle

3.3.1.1. Walking trips Ground-truth counts consistently show that survey respondents underreport their walking trips. The reasons are multiple: survey respondents ''forget,'' they view the walking trip as too brief, or, in the context of a survey, it burdens the time to completion. However, a more profound source of bias is that walking trips are not well captured as an ''activity.'' Walking is both the mode (e.g., to get to the gym) and the activity (recreation). In the future, reliable data from a GPS/passive may provide more complete information than surveys.

The passive GPS may facilitate an understanding of ''where'' and ''why'' walking takes place, alongside the role of the built environment (or a reinterpretation of existing studies). Using aggregate level data, the GPS enumerates walking distance with much more precision than a self-reported survey. And, when walking routes are viewed alongside maps of the built environment, GPS data may provide insight about factors that facilitate or impede pedestrian activity, like dark or windy corridors, the absence of sidewalks, or topography.

The GPS/passive trace has an immediate application and extension from walking to transit research. Conventional surveys oversimplify transit taking because respondents do not focus on transit's multimodal aspect. Survey respondents also overestimate the time spent waiting for transit. The passive GPS trace can provide more complete information on the characteristics of walking to/from a bus stop and door-to-door travel time (Li & Shalaby, 2008). A particularly interesting ''in-site'' experiment is measurement of the walking premium associated with rapid-bus service.

3.3.1.2. Bicycles/Cycling A GPS trace offers many of the same advantages for learning about bicycle travel, as it does for walking. Bicycle trips, like walking, may be undercounted by traditional surveys because they are less likely to take place on a regular, reoccurring basis. An additional problem is that many bicycle riders are young and male, and as Murakami (2008) notes they shun survey taking. The GPS trace may be quite attractive to this demographic group, because cyclists can share with each other and with planners their information on route choice. It will also provide more accurate data on the multiday patterns of bicycle commuting and mixed mode choice. Finally, the GPS trace will better count and delineate the differences in route choice and time of day — between recreational bikers and commuters. This new volume and detail on cycle usage may help communities reexamine the accuracy and usefulness of their bike-master plans.

3.3.1.3. Parking Our understanding of vehicle use may be enhanced by the passive GPS trace. Today, few travel surveys ask about the dynamics of parking and many transportation models still do not explicitly include the time and cost to park a vehicle. But, in large urban areas, the availability and price of parking will influence whether a trip takes place at all and then, the choice of travel mode too (e.g., taxi or train).

"Location-based" software that run on the cell phone are likely to help count parking events, providing that new algorithms are written. Today, GPS helps drivers identify "open" parking spaces. An interesting experiment, already taking place, records the difference in search time and curb-behaviors for GPS users and nonusers (Rodier & Shaheen, 2010). Using a measure of duration, we might also learn whether parking itself generates "new" trips and cold-starts, as drivers move their vehicles to new locations when their "metered time" expires.

Parking rates, recorded in the ITE manual, might be recalculated, with new aggregated GPS results. These counts can provide a more accurate and reliable means of estimating trip generation by activity centers with different characteristics. For example, parking requirements associated with say an urban hospital with good public transit service might be compared to a hospital, in a similar urban setting, lacking transit service.

3.3.2. Passive GPS and Inferences about Trip Counts and Activities

Counting the number and frequency of travel trips with a passive GPS trace is likely to disrupt current and known data keeping. First, GPS devices are known to enumerate more trips than self-reported survey: the undercounting of trips by self-report occurs because respondents (a) avoid survey fatigue, (b) cannot recall their complete travel, or (c) do not cognitively grasp "trip chains." Second, data keeping from self-reported surveys typically reports behaviors over a short duration, usually one day, on occasion, a week, and seldom for a mix of both weekdays and weekend.

There is ample evidence that travel behavior is more dynamic, and less predictable over longer time periods.

Yet another limitation of self-reported survey data is that activities are not as discrete as their measurement. For example, in a travel diary a respondent might code their ‘‘evening out’’ as recreation. In real time, using text and IM on their cell phone that respondent engages in multiple activities: they order take-out on the phone, window shop for a future purchase, network with friends to find a meeting point, and catch up on work related e-mails while traveling to meet. Researchers may find they can no longer validly code trips into discrete categories.

However, the passive GPS may afford new opportunities: it may, at the aggregate level, capture the previously unmeasured influence of ‘‘connected’’ behaviors. Writing from the marketing field, Tancer (2009) describes how Internet behavior can be analyzed collectively: the volume and location of searches reveal larger patterns, which are never detected individually. Transportation planners know that there are network level impacts that are associated with weather, holidays, large sporting events, and the like. The FHWA, for example, funded a project to study the patterns and regularity of crowd movements during large planned special events (cited by Calabrese, Pereira, Lorenzo, Liu, & Ratti, 2010). The passive GPS trace may link travel patterns with previously unexplored social and temporal phenomena.

3.3.3. Passive GPS and Inferences about Route Choice and Dynamic Travel Information

A fundamental assumption of the activity diary/travel survey is that underlying ‘‘activities,’’ which are then completed at ‘‘destinations,’’ motivate trip taking. While fill-in surveys can facilely present a list of ‘‘in-home/out-of-home’’ activities, they are not a good medium for probing what motivated the activity or how the travel-trip was chosen. Doherty (2006), Lee-Gosselin, Doherty, and Papinksi (2006), and others have used prompted surveys and custom designed web studies to fill-in this data gap.

Paradoxically, while passive GPS traces measure macro-level movement, they may have a dual role, helping to mine the confluence of individual choice. Increasingly cell phones are being used to provide travelers with real-time dynamic traffic information (*Technology Review*, 2011). It is becoming feasible to correlate route and mode changes with the distribution of traffic updates and alternative choices.

As an example, consider a transit rider consulting a real-time, crowd-sourced bus-location app on their cell phone (e.g., ‘‘Tiramisu,’’ see Steinfeld, Zimmerman, Tomasic, Yoo, & Aziz, 2011). Depending on the predicted wait time, the potential rider may, on the spot, select a different route or a different mode (walking), or they may choose to complete their activities (say shopping and recreation) in a different temporal sequence. The original mode, destination, and activity have each changed with updated information. Potentially, the passive GPS will record travel that is less predictable than traditional surveys and will provide insights into how people respond to dynamic information, like traffic congestion, in situ.

Passive GPS will also provide useful information to anticipate natural experiments. How do travelers react when there are disruptions to the transportation network, preannounced road closings, detours, or bridge closings? Or, more mundanely, how might variable congestion pricing impact the network, in terms of both temporal and spatial shifts? Given baseline data, the GPS trace can then compare the changes in traffic volumes, route-taking, and temporal distributions. With a bank of experiential data, self-reported surveys may no longer need to ask travelers what they "might do."

3.4. Conclusions and Future Issues for Study

The tools we use to collect travel data are intricately linked to our data processing capabilities (Kitamura, Chen, Pendyala, & Narayanan, 2000). In the 1950s and 1960s computing power and data storage were limited; the four-step model was not able to incorporate predictive data elements, like time of day and multimodal trip taking. In the 1970s, data processing capabilities expanded and models then incorporated more behavioral and household level observations. Today, the speed of computer processing and new data mining tools make it feasible to envisage continuous mobility traces that process quantum amounts of aggregated and anonymous data points.

The cell phone GPS trace will soon collect travel information silently and continuously and reveal actual patterns of movement, rather than those estimated through simulation or models. Fewer survey studies may mean less information about the demographic and psychological motivations for travel, which are important predictors in current modeling. The trade-off is that as we reduce the "noise" from survey measurement the "signal" will be modified, and record far more system-wide and baseline travel data. In many ways this is a stronger signal because it will be produced in real time, across different transportation modes.

More than five years ago, Doherty (2006) wrote a paper with the provocative title "Should We Abandon Activity Type Analysis." He suggested then that transportation researchers look again at the use of traditional activity classifications like work and shopping. His weeklong scheduling study found that behavior was far more dynamic than what was captured by the conventional survey measures. This work suggests that survey results were "noisier" than originally imagined because the studies classified behavior through the "lens" of researchers, not of respondents. As we move forward, we are less likely to want to actively recruit samples that report on their daily travel behavior if we fundamentally ask new questions.

It is recognized that we will still encounter segments of the population who do not wish to carry the phone or have earlier software. However, the market shows tremendous growth. In six countries more than 50% of the population uses smart phones: these are Australia, the United Kingdom (sic), Sweden, Norway, Saudi Arabia, and the United Arab Emirates. An additional seven countries — the United States, New Zealand, Denmark, Ireland, Netherlands, Spain, and Switzerland — now have more than 40% smart phone penetration (Think with Google, 2012).

The source notes that, "a global movement is happening as smart phone adoption moves mainstream. Mobile devices have become indispensable to people's lives and are driving massive changes in consumer behavior."

For reaching emergency services, GPS is generally regarded as a useful feature. In the future, acceptance of its value may tip public sentiment to support development of a vast "data commons," particularly for traffic information and congestion mitigation. Recruitment for additional, specialized studies, say for attitudinal data, short stated preference studies, or demographic information, might facilely take place, still using the telephone, if users are compensated with either free cell phone plans or air time.

In conventional travel surveys there is considerable "noise" in the data, as validity depends on factors like the representativeness of the sample, and then on the integrity and veracity of self-reports. With a passive GPS trace, the "noise" from survey measurement and error are minimized. The trade-off is that as we reduce the "noise" from survey measurement the "signal" or data is modified. We consider how the collection of real-time tracking may change measures like trip frequency, mode-share, and activity coding.

For some segments, like bicycle users, there is immediate value in visualizing daily travel, and making it into more of a participative venture. Hence, for some groups the passive GPS trace may evolve into a more interactive and personalized visual travel logger. It is also possible that "collective," "crowd-sourced" travel behaviors might emerge, say for daily rail and bus riders. It is a useful speculation whether a GPS trace can be used, in the words of Kitamura, Fujii, and Pas (1997), to make people more aware of their travel patterns, and help them individually and collectively, assess the impact of transportation on their quality of life. At the collective level, real-time traffic information is already being used to redirect vehicle trips and reduce air pollution and fuel consumption. Perhaps the GPS data could be used, at a macro level, to analyze natural experiments; for example, when drivers encounter less traffic on a regular commute, do they make additional new trips on the same tour, or new trip later in the day or evening?

It is necessary and inevitable as this new technology is scoped, to address whether the GPS traces are confidential and anonymous. This was anticipated by Ratti et al. (2006), who viewed privacy issues as a concern when the data was provided to a third party, other than the mobile phone operator. Working with anonymous, aggregated data, individual movements cannot be tracked and individual privacy is a non-issue (Ratti et al., 2006, citing Fisher & Dobson, 2003; see also Wigan & Clarke, 2009). In April 2011 (parenthetically, while this chapter was being written) cell phone traces became a popular news story. A hacker revealed that Apple I-phones, and to a lesser extent Android phones, were surreptitiously recording tracking locational data. Once installed, the operating system upgraded logged location data, including time and date, and stored it in a hidden file. The software used triangulation from cell phone towers and Wi-Fi. It was not as accurate as it would be using GPS data (Homeland Security Newswire, 2011). Ironically, Quercia, DiLorenzo, Calabrese, and Ratti (2011) had developed a cell phone application to measure outdoor advertising exposure that scrambles the name/identity of the recipient and reports erroneous locations. Ingeniously, the randomized response algorithm still makes it possible to accurately collect and process aggregated location data.

In this discussion of cell phone privacy issues, we should recall that the demand for location-based services began with an European Universal Service Directive (Directive 2002/22/EC) which required fixed and mobile network operators to transmit the location of people calling ‘‘112’’ emergency lines, in the best possible way based on their national emergency standards and the technological possibilities of the networks (GIS news, cited by Ratti et al., 2006). A later implementation in the United States was mandated by the Federal Communications Commission and launched emergency e-911 service. As cell phones with smart applications grow, they will increasingly be used to help people navigate, select routes, and save fuel or travel time. They will also be used to purchase goods, and carry transit fares. In many ways, they will become the indispensable ‘‘computer that you wear.’’ Like e-911, the privacy issues raised by their breadcrumb trail are likely to be offset by new applications that are seen to ‘‘empower’’ consumers rather than ‘‘surveille’’ them (Lee-Gosselin, Doherty, & Shalaby, 2010). AU :7

So, as data becomes ubiquitous and continuous, the challenge will be not to conceal it — but to exploit it usefully. Developing ‘‘good’’ data algorithms has some similarity to developing ‘‘good’’ survey questions because in both cases the researcher influences what is measured: for example, inferences about speed and stops are used to deduce the mode of travel, and inferences from GIS overlays and duration are used to code activities. Srinivasan, Bricka, and Bhat (2010) observe, ‘‘The use of passive GPS technology in travel surveys shifts considerable burden from the respondent to the analyst.’’ There is no standardization so different algorithms will make different observations about the same ‘‘objective’’ travel event (Stopher, 2009b). An analogy from self-reported surveys is that there are multiple ways to get at the same information and elicit different answers (e.g., ‘‘What is your age’’; When is your birthday’’; ‘‘How old are you?’’).

The final observation, and a challenge for the future, is that although cell phone traces are an entirely new way of measuring mobility patterns, researchers will indubitably choose to interpret them in familiar ways. Resistance is due, in part, to professional norms of the research community and in part to the extensive investments made in existing simulations and models. But, in the literature review we noted how computer scientists and media specialists were beginning to probe mobility data for spatial/temporal patterns. Not unlike the fleetness of the horseless carriage, the passive GPS brings a previously unimagined volume and speed to travel data. The continuing push to investigate this data from other disciplines, and the opportunity for both new revenue and new data streams will likely refocus transportation efforts.

Acknowledgments

I would like to acknowledge the assistance and offer of mobile survey data from Geoff Palmer at Survey-on-the-Spot, and the assistance of the marketing department at the MBTA (Massachusetts Bay Transit Authority) which allowed me to internally test mobile surveys on their Blackberry phones. I would also like to acknowledge the help of Ryan Chin, who helped me navigate the MIT Media Lab, in a different context.

References

American Association for Public Opinion Research (AAPOR). (2010). Research synthesis on online panels. *Public Opinion Quarterly*, *74*(4), 711–781.

Asakura, Y., & Hato, E. (2004). Tracking survey for individual travel behaviour using mobile phones: Recent technological development. *Transportation Research C*, 207–233. AU :8

Asakura, Y., Hato, E., & Sugino, K. (2005). Simulating travel behaviour using location positioning data collected with a mobile phone system. *Springer Operations Research/Computer Science Interfaces Series*, *31*, 183–203.

Bayir, M., Demirbas, M., & Cosar, A. (2010). A web- based personalized mobility service for smartphone applications. *The Computer Journal*, *54*(5), 800–814.

Beatty, G. (2010). *Why aren't we saving the planet? A psychologist's perspective*. London: Routledge.

Becker, R., Cáceres, R., Hanson, K., Loh, J. M., Urbanek, S., Varshavsky, A., & Volinsky, C. (2011, June). Clustering anonymized mobile call detail records to find usage groups. 1st Workshop on Pervasive Urban Applications (PURBA). Retrieved from http://www.kiskeya.com. Accessed on June 23, 2011.

Bierlaire, M., Chen, J., & Newman, J. (2010). *Modeling route choice behavior from smartphone GPS data*. Transport and Mobility Laboratory Report No. 101016, Ecole Polytechnique Federale de Lausanne, Switzerland.

Calabrese, F., Pereira, F., Lorenzo, G., Liu, L., & Ratti, C. (2010). The geography of taste: Analyzing cell-phone mobility and social events. *Pervasive Computing*, *6030*, 22–37. doi:10.10070978-3-642-12654-3_2

Chapleau, R., Trepanier, M., & Chu, K. K. A. (2008). The ultimate survey for transit planning: Complete information with smart card data and GIS. Presented at the 8th international conference on transport survey method, Annecy, France.

Chen, J., Newman, J., & Bierlaire, M. (2009). *Modeling route choice behavior from smart-phone GPS data*. Retrieved from http://transpor2.epfl.ch/proceedings/CHEN09_IATBR.pdf. Accessed on June 2, 2011.

Chorus, C., & Timmermans, H. (2010). Ubiquitous travel environments and travel control strategies: Prospects and challenges. In M. Wachowicz (Ed.), *Movement aware applications for sustainable mobility: Technology and applications* (pp. 30–51). AU :9

Doherty, S. (2006). Should we abandon activity type analysis? Redefining activities by their salient attributes. *Transportation*, *33*(6), 517–536. doi:10.10070811116-006-0001-9

Eagle, N., & Pentland, A. (2005). *Eigenbehaviors: Identifying structure in routine*. Retrieved from http://vismod.media.mit.edu/tech-reports/TR-601.pdf

Fisher, P. F., & Dobson, J. E. (2003). Who knows where you are, and who should, in the era of mobile geography? *Geography*, *88*(4), 331–337.

Gonzales, M., Hidalgo, C., & Barabasi, A. (2008). Understanding individual human mobility patterns. *Nature*, *453*, 479–482.

Homeland Security Newswire. (2011, April 22). Cell phone privacy. . Retrieved from http://www.homelandsecuritynewswire.com/cell-phone-privacy. Accessed on May 24, 2011.

Isaacman, S., Becker, R., Caceres, R., Kabourov, S., Martonosi, M., Rowland, J., & Varshavsky, A. (2011). Identifying important places in people's lives from cellular network data. Conference Paper. Retrieved from http://www.cs.arizona.edu

Kennedy, C., & Everett, S. (2011). Use of cognitive shortcuts in landline and cell phone surveys. *Public Opinion Quarterly*, *75*(2), 336–348. doi:10.1093/poq/nfr007

Kitamura, R., Chen, C., Pendyala, R., & Narayanan, R. (2000). Micro-simulation of daily activity-travel patterns for travel demand forecasting. *Transportation*, *27*, 25–51. doi:10.1023/A:1005259324588

Kitamura, R., Fujii, S., & Pas, E. (1997). Time-use data, analysis and modeling: Toward the next generation of transportation planning methodologies. *Transport Policy*, *4*(4), 225–235. doi:10.1016/S0967-070X(97)00018-8

Lee-Gosselin, M., Doherty, S., & Papinksi, D. (2006). Internet-based prompted recall diary with automated GPS activity-trip detection: System design. *Proceedings of the 85th annual meeting of the Transportation Research Board*, Paper No. 06-1934, Washington, DC.

Lee-Gosselin, M., Doherty, S. T., & Shalaby, A. (2010). Data collection on personal movement using mobile ICTs: Old wine in new bottles? In M. Wachowicz (Ed.), *Movement-aware applications for sustainable mobility: Technologies and approaches* (pp. 1–14). Hershey, PA: Information Science Reference.

Lee-Gosselin & Harvey, A. (2006). Non web technologies. In P. Stopher (Ed.), *Travel survey methods: Quality and future directions* (pp. 561–568). Elsevier. AU :10

Li, Z., & Shalaby, A. (2008). Web based GIS system for prompted recall of GPS assisted personal travel surveys: System development and experimental study. *87th Transportation Research Board meeting*, Washington, DC.

Madey, G., Barabasi, A., Chawla, N., Gonzalez, M., Hachen, D., Lantz, B., Pawling, A., … Yan, P. (2007). Enhanced situational awareness: Application of DDDAS concepts to emergency and disaster management, Computational Sciences — ICCS. *Lecture Notes in Computer Science*, *4487*, 1090–1097.

Manasseh, C., Aherh, K., & Sengupta, R. (2009). The connected traveler: Using location and personalization on mobile devices to improve transportation. ACM Digital Library. Proceedings of the 2nd International Workshop on Location and the Web. AU :11

Murakami, E. (2008, January). Hard to reach populations. Presentation to N.Y. Metropolitan Transportation Council and Region 2 UTRC meeting. In *Resources on Household Travel Survey Methods*, New York, NY.

Oulasvirta, A. (2005). Grounding the innovation of future technologies. *Human Technology*, *1*(1), 58–75.

Quercia, D., DiLorenzo, G., Calabrese, F., & Ratti, C. (2011). Mobile phones and outdoor advertising: Measurable advertising. *IEEE Pervasive Computing*, *10*(2), 28–36. doi:10.1109/MPRV.2011.15

RAND. (1968). Retrieved from http://www.rand.org/pubs/memorandaRM588.html. Accessed on September 1, 2011.

Ratti, C., Pulselli, R. M., Williams, S., & Frenchman, D. (2006). Mobile landscapes: Using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design*, *33*(5), 727–748. doi:10.1068/b32047

Reades, J., Calbrese, F., & Ratti, C. (2009). Eigenplaces: Analyzing cities using the space-time structure of the mobile phone network. *Environment and Planning B*, *36*, 824–836. doi:10.1068/b34133t

Reades, J., Calabrese, F., Sevstuk, A., & Ratti, C. (2007). Cellular census: Explorations in urban data collection. *IEEE Pervasive Computing*, *6*(3), 30–38.

Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., & Srivastava, M. (2010). Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks*, *6*(2) article 13.

Rodier, C., & Shaheen, S. (2010). Transit based smart parking: An evaluation of the San Francisco bay area field test. *Transportation Research, Part C*, 225–233.

Srinivasan, S., Bricka, S., & Bhat, C. (2010). Methodology for converting GPS navigational streams to the travel-diary data format. Transportation Research Board annual meeting paper.

Steinfeld, A., Zimmerman, J., Tomasic, A., Yoo, D., & Aziz, R. (2011). Mobile transit rider information via universal design and crowd-sourcing. Transportation Research Board annual meeting paper.

Stopher, P. (2009a). The travel survey toolkit: Where to from here. In P. Bonnel, M. Lee-Gosselin, J. Zmud & J. Madres (Eds.), *Transport survey methods* (pp. 15–46). Bingley, UK: Emerald Group Publishing.

Stopher, P. (2009b). Collecting and processing data from mobile technologies. In P. Bonnel, M. Lee-Gosselin, J. Zmud & J. Madres (Eds.), *Transport survey methods* (pp. 361–391). Bingley, UK: Emerald Group Publishing.

Tancer, B. (2009). *Click*. London: Harper Collins.

Technology Review. (2011). Social surveillance yields smarter directions by Tom Simonite. *Technology Review*, February 2.

Think with Google. (2012). Retrieved from http://www.thinkwithgoogle.com/mobileplanet. Accessed on July 12 2012.

Van der Hoeven, F. (2010). Setting the stage for the integration of demand responsive transport and location-based services. In M. Wachowicz (Ed.), *Movement aware applications for sustainable mobility: Technology and applications*. IGI Global Publishing. doi:10.4018/978-1-61520-769-5.ch004

Wigan, M. R., & Clarke, R. (2009). Transport and surveillance aspects of location-based services. *Transportation Research Board*, *2105*, 92–99. doi: 10.3141/2105-12

AUTHOR QUERY FORM

	Book: Johanna-3610686 Chapter: 3	Please e-mail or fax your responses and any corrections to: E-mail: Fax:

Dear Author,

During the preparation of your manuscript for typesetting, some questions may have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list.

Disk use

Sometimes we are unable to process the electronic file of your article and/or artwork. If this is the case, we have proceeded by:

☐ Scanning (parts of) your article ☐ Rekeying (parts of) your article

☐ Scanning the artwork

Bibliography

If discrepancies were noted between the literature list and the text references, the following may apply:

☐ **The references listed below were noted in the text but appear to be missing from your literature list. Please complete the list or remove the references from the text.**

☐ ***UNCITED REFERENCES:*** **This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or delete it. Any reference not dealt with will be retained in this section.**

Queries and/or remarks

Location in Article	Query / remark	Response
AU:1	Please check "paper" has been changed to "chapter" while referring to the current chapter.	
AU:2	Please check the edits made in the following sentence: "There were several early…"	
AU:3	As per style "Section 2" and "Section 3" cannot be given as heading. Therefore Please provide a heading for Sections 3.2 and 3.3.	
AU:4	In the following sentence: "The second column of…" please specify the period for "today."	

AU:5	Please check the edit made in the following sentence: "The "satisficing behavior…""	
AU:6	As bold text is not preferred as per style, it has been changed to italics. Please check.	
AU:7	Please check whether this should be "surveil."	
AU:8	Please provide the volume no. in the following refs. Asakura & Hato (2004); Rodier & Shaheen (2010).	
AU:9	Please provide the name and location of the publisher in the ref. Chorus and Timmermans (2010).	
AU:10	Please provide the location of the publisher in the refs. Lee-Gosselin & Harvey (2006); Van der Hoeven (2010).	
AU:11	Please provide the venue of the Proceedings/meeting in the following refs. Manasseh et al. (2009); Srinivasan et al. (2010); Steinfeld et al. (2011).	