Survey Research Methods (2015)
Vol. 9, No. 1, pp. 1-13

ISSN 1864-3361
http://www.surveymethods.org

A feasibility test of using smartphones to collect GPS information in face-to-face surveys

Kristen Olson
University of Nebraska-Lincoln

James Wagner
University of Michigan

Survey organizations rely on interviewers to make informed and efficient decisions about their efforts in the field, including which housing units they approach to knock on doors, seeking to make appointments, and obtain interviews. This paper presents initial findings from a feasibility test on a face-to-face survey, the US National Survey of Family Growth, in which the interviewers were equipped with GPS-enabled smartphones. This test included an experiment to determine if the use of the GPS-enabled smartphone altered interviewer behavior. We evaluate interviewer compliance with the GPS request, the effects of using the GPS device on interviewer behavior, and the quality of the recorded GPS points as related to interviewer behaviors. This test also included two surveys of interviewers that were completed after using the smartphone. We report results from these surveys and link one set of survey results to interviewer GPS compliance rates. Implications for future use of GPS devices to monitor and understand interviewer travel behavior are discussed.

Keywords: interviewers; GPS; mobile devices; face to face surveys

1 Introduction

Interviewer-administered face-to-face surveys are used to collect data in surveys such as the US National Health Interview Survey, the US National Crime Victimization Survey, the European Social Survey, and the UK's Understanding Society. Interviewers complete a wide variety of tasks while in the field for a survey, from knocking on the door of a housing unit to see if anyone is at home to completing interviews (e.g. Campanelli & Purdon, 1997; Morton-Williams, 1993). Understanding how interviewers make decisions about these tasks (e.g., which housing units to approach, when to do so, and in what order) is important. Previous research has shown that interviewers in field surveys vary substantially in outcomes such as contact rates and response rates (e.g. Durrant & Steele, 2009; O'Muircheartaigh & Campanelli, 1999; Pickery & Loosveldt, 2002; Purdon, Campanelli, & Sturgis, 1999). Why interviewers vary in these field outcomes is not well understood.

One possible reason for this variation is differences in the decisions that interviewers make about which housing units to visit and how they travel between these housing units. These travel decisions also have important cost implications which to date have received attention only through simulation (Bienias, Sweet, & Alexander, 1990; Chen, 2012) or in aggregate (Judkins, Waksberg, & Northrup, 1990; Kalsbeek, Botman, Massey, & Liu, 1994). These travel patterns may also influence or reflect field performance. For example, interviewers may observe that someone is home at a sampled unit and move past other sampled housing units. This kind of behavior may increase contact rates while also increasing total miles travelled. Alternatively, interviewers may increase their travel as a result of making unproductive calls in one neighborhood, which leads to them move on to a different sampled neighborhood.

One reason that interviewer travel decisions have received limited empirical attention is the lack of good quality data. Each interviewer task is supposed to be recorded by the interviewers in administrative systems designed to monitor the survey process. These process data or paradata (Couper, 1998; Couper & Lyberg, 2005) include call record data, timesheet data, expense reports, and interviewer observations. Given that interviewers are reporting on their own behavior, it is likely that some tasks are incorrectly reported or underreported. For example, Biemer, Chen, and Wang (2013) conducted a survey of interviewers that revealed systematic underreporting of call attempts. Studies of inconsistencies in paradata by the US Census Bureau and Statistics Canada also call into question the quality of call record data (Bates, Dahlhamer, Phipps, Safir, & Tan, 2010; Laflamme & Karaganis, 2010).

An alternative, objective measurement of interviewer travel in the field can come from equipping interviewers with global positioning system (GPS) devices (Nusser, 2007). GPS devices can measure location (latitude and longitude),

Kristen Olson, Department of Sociology, University of Nebraska-Lincoln, 703 Oldfather Hall, Lincoln, NE 68588-0324 (kolson5@unl.edu)

elevation, speed, direction, date, and time, among other measurements, at regular time intervals.

Survey organizations in the US such as RTI, NORC and the US Census Bureau have used GPS devices, most often for field listing. In the simplest incarnation, field listers are equipped with a GPS device or smartphone with GPS capabilities to record the latitude and longitude of each housing unit on a sample frame (Cecchi & Marquette, 2012; Dekker, English, Winfrey, & Seeger, 2013; Levinsohn et al., 2010; Morton et al., 2007; Seeger, 2011). Other surveys used GPS devices to actively or passively capture the location of an interview to help detect the position of the interviewer while field listing or interviewing, sometimes in an effort to reduce "curbstoning" (Cecchi & Marquette, 2012; Ellis, Sikes, Sage, Eyerman, & Burke, 2011; Haddaway, 2013; Keating, Loftis, McMichael, & Ridenhour, 2014; Sikes, 2009; U.S. Census Bureau, 2014), or to help interviewers plan travel routes in field studies (Nusser & Fox, 2002). In addition, real-time collection of trips taken by respondents has been implemented in the US and Canada in travel surveys (Chung & Shalaby, 2005; Wolf, Guensler, & Bachman, 2001), in the Netherlands in time use studies (Sonck & Fernee, 2013), and in the US, Europe, and Australia to measure physical activity (Krenn, Titze, Oja, Jones, & Ogilvie, 2011).

The previous uses of GPS devices for interviewer evaluation have been linked to a static location, such as a sampled housing unit, and for basic monitoring of field data collection (Kurkowski, 2013). To our knowledge, no previously published studies have used GPS devices to study interviewers while they are moving in the field between housing units. As a result, the effect of collecting GPS data about the movement of field interviewers in a production survey on interviewers and field outcomes and the quality of that data is unknown.

One concern with using a new device is that the resultant data may be incomplete or contain errors (Couper, 2013). If collecting GPS data requires action from the interviewer, such as turning on a GPS device, then there can be missing data when the interviewer forgets to do this action. Although interviewers may be instructed to follow certain actions in training, previous research has shown that interviewers often deviate from desired behavior. For example, interviewers fail to read questions exactly as written – a fundamental part of their job - between 3% and 73% of the time (Ongena & Dijkstra, 2006). Interviewers fail to record all of their visits to sampled housing units in call records (Biemer et al., 2013). People also forget to carry devices. For example, when *respondents* are asked to carry activated GPS devices missing data rates for GPS data range from 2.5% to 92%, with one of the strongest predictors being the length of the study period (a long study took four or more days) (Krenn et al., 2011).

Missing data can also result when a file is corrupted in transmission, when there is GPS signal noise, or when GPS devices fail to record data for technical reasons (Lemmens, 2011). GPS failure rates of over 10% have been reported for high quality GPS-only wearable devices (e.g. Chung and Shalaby, 2005, 21.5% missing) and iPhones (e.g. Zandbergen, 2009, up to 12.3% missing). Measuring GPS signals in "urban canyons" is particularly challenging. Satellite signals bounce off of tall buildings, making detecting GPS signals difficult (Gong, Chen, Bialostozky, & Lawson, 2012). Other problems with loss of GPS data can occur in tunnels or in certain types of cars (Stopher & Greaves, 2007). Missing data can also occur during trips when a GPS device has to obtain a connection after not having one (Kerr, Duncan, & Schipperjin, 2011; Krenn et al., 2011; Stopher & Greaves, 2007). In a set of surveys in developing countries, up to 30% of data passively collected from GPS-enabled mobile devices carried by interviewers was missing due to unreliability of the GPS chips (Haddaway, 2013). Thus, there are factors that are not related to the interviewer that may lead to missing data.

In addition to missing data, it is well-established that GPS data have measurement error (Chung & Shalaby, 2005). Measurement error in GPS data results from slow connectivity (e.g., device turning on, being in a fast moving car, entering a building), physical structures (e.g., high buildings, dense tree cover, being indoors), inadequate satellite signals, errors in transmitting from the satellites to the ground, or the need to use the less accurate method of triangulation using cell phone towers (Goodchild, 2007; Kerr et al., 2011; Lemmens, 2011; Zandbergen, 2009). In general, dedicated GPS devices are more accurate and have better signals than cell phones (Wu et al., 2010), and GPS devices are more precise when they are stationary than when they are moving (Rainham, Krewski, McDowell, Sawada, & Liekens, 2008).

The interviewer may also forget to turn the GPS device off, resulting in abnormally long files or recordings during illogical times (Lemmens, 2011). This is less of an issue than missing data in terms of identifying travel patterns, but leads to more difficulty in identifying when an interviewer was actually working. How often any of these errors occur when interviewers use GPS devices is unknown.

Using a monitoring device such as a GPS device may also change interviewer behavior. Previous research has shown that audio recording interviews leads to fewer interviewer-related errors during questionnaire administration (Billiet & Loosveldt, 1988; Fowler & Mangione, 1990). It is plausible that active monitoring of field behaviors will lead interviewers to change their field effort. If interviewers are aware of the monitoring of their travel movements, they may be more productive by making more call attempts and visiting more sampled cases.

GPS devices have been successfully used to examine travel patterns of *respondents* in real time in general population surveys. The applicability of these findings to monitoring interviewer field travel in real time on a national scale is limited. Almost all of the GPS-based travel surveys in the US have used GPS devices installed in a car, are almost exclusively confined to a limited geographic range such as a city or a state, and are conducted on a small sample for a small number of days (Bricka, 2008). Studies of physical activities have used high quality portable devices for a short period of time with small samples in limited geographic areas (Krenn et al., 2011). Thus, the effects of collecting GPS data on *interviewers* from a nationally representative geographic area, including both urban and rural areas, over multiple months using smartphones are unknown.

Data from GPS devices may be useful for answering critical questions about interviewer variability in costs and outcomes. As part of a feasibility test, GPS data were collected by interviewers working on the National Survey of Family Growth Continuous 2011-2019. We address the following questions – to what extent do interviewers comply with the request to use a GPS logging application? Does this request affect field behavior? What do interviewers think about the smartphone and GPS app? Do the GPS data provide information covering an interviewer's workday? We evaluate results from an experimental assignment of interviewers to use a GPS logging application, or "app," on a smartphone. We also provide an initial examination of the quality of the data collected via the smartphone app and results from interviewer surveys about the smartphones, as well as some initial insights about interviewer travel gleaned from these data. Given the novelty of the use of GPS devices in sample surveys to study interviewer travel, it is unknown whether interviewers will reliably comply with the instruction to use a GPS app on a smartphone to track their movements, the quality of the GPS data in a large scale national survey, or how the GPS request will affect field behavior. This paper aims to fill this gap. We limit the scope of this paper to whether interviewers consistently use the GPS app, the quality of the data from this app as related to interviewer decisions, and interviewers' attitudes towards the device. Understanding these factors are critical first steps to establishing feasibility of using a GPS logging app for gathering cost and effort data in the field.

2 Data and Methods

The data come from a national face-to-face survey, the US National Survey of Family Growth Continuous 2011-2019 (NSFG) conducted by the Survey Research Center at the University of Michigan. The NSFG uses a nationally representative area probability sample design, with a continuous sample design that rotates PSUs annually. A new, independent sample of housing units is released four times a year (quarterly) and is in the field for 12 weeks. For any given year, the NSFG has 40 to 45 female interviewers on staff. Interviewers are employed by the University of Michigan to work at least 30 hours each week on the NSFG. Interviewers are paid by the hour; mileage and other transportation costs are reimbursed. The interviewers work in 35 primary sampling units with over 450 unique area segments. There are approximately 20,000 housing units sampled each year across these area segments. The NSFG interview is completed in two stages. The first stage is a "screening" interview to determine eligibility. The second stage is a "main" interview conducted with sampled eligible persons. Cumulatively, these sampled units receive more than 100,000 call attempts. This paper focuses on data collected between Fall 2011 (Q1) and Summer 2013 (Q8) (AAPOR RR4 between 71% and 76% over these quarters).

The NSFG has deployed a web-enabled smartphone with its field interviewers. GPS data have been collected using the off-the-shelf GPS Logger app on Android-platform smartphones (Motorola Atrix 2) since September 2011. The data are saved to a file that is transmitted to a secure server daily. In addition to latitude and longitude, these telephones capture time, date, elevation, speed, direction of travel, and indications of the quality of measurement (not reported here) at one minute intervals.

We considered three options for collecting GPS data. First, interviewers could "snap" GPS locations at the time of the interview; this would not yield data collected about non-interviews or about travel itself. Second, the smartphone could be set to collect GPS data without any interviewer intervention. As interviewers may carry their work smartphone with them even when they are not working, this option raises issues related to privacy (collection of personal movements in addition to work-related movements) and file length and complexity (irrelevant travel with work-related travel). Third, interviewers could turn on and off the GPS device itself during their work hours. This provides information about travel in the field to both non-interviews and interviews, minimizes the recording of irrelevant non-work travel, and limits privacy concerns. We anticipated that NSFG interviewers would carry their work smartphone with them even when they were not working (e.g., running errands, dropping children off at school), and wanted to respect their privacy in these non-work related tasks. This also reduces the length of the GPS data file, easing merging of the GPS data with call records (described below). Allowing the interviewer to turn the device on and off also provides them the right to withdraw from the GPS measurement, but raises issues then about compliance with the request. This is the option used in this study, and we empirically evaluate compliance rates as an important outcome.

The interviewers received in-person training on how to use the app as part of the study-specific training. During training, we explained that the data generated by the app were for research purposes only. Thus, the effect of the GPS device on behavior may be somewhat attenuated compared to using it as a management device. All NSFG interviewers were instructed to take their smartphones with them, turn the app on when they left their home for fieldwork, and turn the app off when they returned from their daily interviewing work. To start the app, the interviewers simply had to open it and press a "begin logging" button. Interviewers were also trained that they were permitted to use the mapping app on the phone, but that they were not allowed to enter sampled addresses directly into the app for confidentiality reasons. Instead, they needed to enter nearby intersections for use of the mapping app. Work-related phone calls could be made, but the phone was not for personal use.

Interviewer compliance rates were monitored on a weekly basis, with focused attention during Q1-Q3 and Q5-Q6, when the PSU rotation led to bringing in a new set of interviewers. Supervisors discussed these rates with interviewers and sought improved compliance. This also led to the discovery of technical difficulties which were addressed by central office staff. Because we were concerned that the use of the GPS Logger app would affect field outcomes, we randomly assigned interviewers to use the GPS app in either the first quarter or the second quarter of data collection. In particular, we hypothesized that the use of the app would make interviewers aware that their field behavior was being recorded, and thus change calling behaviors, similar to the effect of audio recording interviews on interviewer behavior while administering a questionnaire (Billiet & Loosveldt, 1988; Fowler & Mangione, 1990). We hypothesized that recording of movement will make interviewers more likely to change their calling behavior to look "active" in the field, and thus make more call attempts and visit more sampled cases. In the first quarter (Q1), a random sample of half of the interviewers were instructed to submit GPS files (n = 22 interviewers); this group was instructed to not submit GPS files in the second quarter (Q2). In Q2, the complementary random sample of interviewers were instructed to submit files (n = 24 interviewers), after being instructed not to submit files in Q1. Because the interviewers were asked to turn the app on and off themselves, they were aware of their participation in the experiment and in the use of the app. Although a secondary research question (and unusual implementation plan), this design allows us to evaluate whether there are differences in compliance and other field behaviors when there is a 12 week lag between training and instructions to use the GPS device. It also allows us to evaluate whether there are "carryover" effects – that is, whether the initial use or lack of use of the GPS device changes field behaviors during subsequent quarters.

After Q2, we abandoned the random assignment and asked *all* interviewers, including new hires, to submit the files (n = 69 interviewers total over the eight quarters examined here). There is a GPS file for each day that the interviewer turned the app on and off.

We also conducted two web surveys with NSFG interviewers about their travel behavior and the smartphone. One survey was conducted during summer 2012 (n = 29, AAPOR RR1 62%), and the second was during summer 2013 (n = 25, AAPOR RR1 71%).The first survey was an anonymous survey of the NSFG interviewers and, as such, cannot be linked to field outcomes. The second survey has been linked to field outcomes. Table 1 summarizes the timing of the GPS data collection and interviewer surveys over the eight quarters.

To calculate compliance rates, we needed to know the total number of days for which a GPS record could be feasibly collected. We aggregated data from the call records to an interviewer-day level for each day with an entry made in the call records, and combined them with the GPS data. Days during the first two quarters in which the interviewer was not randomly assigned to submit GPS files are excluded. In total, we have 12,187 interviewer-days (i.e., a record for each interviewer for each day of the study period for which we have a call record) over the first eight quarters of data collection.

We examine four categories of outcomes. First, we examine results of the experimental random assignment of the instruction to submit GPS files on interviewer compliance and other field outcomes. Second, we examine variability in compliance across interviewers and differences in field outcomes for days with and without GPS data recordings. Third, we review results from the web surveys of interviewers about the smartphone, and examine these as correlates of compliance with the GPS request. Finally, we evaluate the quality of the GPS data using the time at which the measurements occurred and the length of the file.

3 Findings

3.1 Effects of the Experimental Assignment to use the GPS App

We start by examining how well interviewers complied with the instruction to record their movements using the GPS logging application. To do this, we examine the total number of interviewer-days in which a GPS file was received out of the total number of eligible days worked as indicated in the call records. Out of the 12,187 interviewer-days across the first eight quarters of the NSFG, 7,168 (58.82%) had an associated GPS file. This rate varied across the quarters. Q1 immediately followed training, and thus the interviewers randomly assigned to use the GPS logging app during this quarter had immediate implementation of their training on this app. The interviewers randomly assigned to start app usage in Q2 had a 12 week delay in their initial use of the application. Not surprisingly, Q1 had a much higher compliance rate than Q2 – 72.5% versus 58.8%. However, when accounting for the clustering within interviewers, this difference was not statistically significant (Rao-Scott χ^2 = 2.01, p = 0.16). In the subsequent quarters, compliance rates ranged from a low of 42.7% (Q4) to a high of 69.0% (Q5) (Figure 1). The

Table 1
GPS and Interviewer Survey Implementation Procedures, NSFG 2011-2019, Q1-Q8

Date	Survey Process	Who Uses App?
Q1, Fall 2011	Interviewers trained	Random half
Q2, Winter 2011		Complementary random half
Q3, Spring 2012		All interviewers
Q4, Summer 2012	Q4 Interviewer Survey conducted	All interviewers
Q5, Fall 2012	New interviewers hired and trained	All interviewers
Q6, Winter 2013		All interviewers
Q7, Spring 2013		All interviewers
Q8, Summer 2013	Q8 interviewer survey conducted	All interviewers

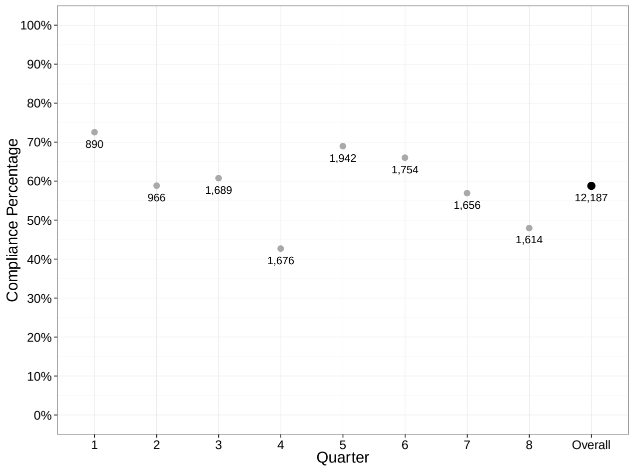

Figure 1. GPS App Compliance Rate by Quarter and Overall, Q1-Q8, NSFG. The points on the graph are labeled with the number of interviewer-days in each quarter.

variation in rates across quarters can be attributed in part to differential levels of monitoring at the central office. Fewer staffing resources were available for monitoring during Q4 and Q8 than the other quarters.

Compliance rates declined steadily over weeks of the field period overall and in most of the quarters of data collection (Figure 2). This difference in compliance rates over the weeks of the field period is statistically significant ($F = 6.65$, $p < 0.0001$). We see the most modest drops in GPS app compliance over the course of the field period in Q1 and Q5 when monitoring was the most extensive, and the steepest drops in Q4 and Q8 when monitoring was less prevalent.

The GPS logging app was intended to record information about interviewer travel behaviors unobtrusively. The interviewers were aware of the app because they had to turn it on and off for each trip. As a result, calling patterns or other field effort outcomes may have been affected. Table 2 shows a variety of field outcomes for interviewers who were randomly assigned to use the app in Quarter 1 (Assigned GPS App Q1) and those who were randomly assigned to use the app in Quarter 2 (Assigned GPS App Q2), whether or not they actually complied with the app instruction (an intent-to-treat analysis). In the first column (Assigned GPS App Q1), Q1 (in bold) is the quarter to which the interviewers are randomized to use the app, and Q2 is the quarter in which they are instructed not to use the app. In the second column (Assigned GPS App Q2), Q1 is the quarter during which the interviewers do not use the application, and Q2 (in bold) is the quarter in which they are randomized to use the application. Thus, we have both a between subjects design (Assigned GPS App Q1 vs. Assigned GPS App Q2) and a within subjects design (Q1 vs. Q2) for this experiment.

The first field outcome we evaluate is the mean number of call attempts made to a sampled case. We display two types of call attempts – the first set of call attempts includes all calls, whether they were made face-to-face or via telephone. The second set of call attempts subsets the calls to only those made via in person attempts (interviewers assigned the modes in the call records; we would not expect effects of the GPS on telephone calls, but errors may exist in the mode assignment). As can be seen in Table 2, the interviewers who were randomly assigned to use the GPS app in Q1 had more calls overall (6.77 in Q1 and 6.53 in Q2) than those randomly assigned to use the GPS app in Q2 (5.76 in Q1 and 5.74 in Q2), but this almost one call difference (6.77 vs. 5.76) was not statistically significant at the $p < 0.05$ level ($p = 0.07$ for assigned to Q1 vs. assigned to Q2 in Q1). Also, interviewers did not have significantly different numbers of call attempts during the quarter to which they were assigned to use the GPS logging application. There is no significant difference between the groups in the proportion of calls with contact out of all call attempts (ranging from 0.34 to 0.39 across both quarters and groups) or the proportion of completed interviews or screeners out of the number of calls with interview or refusal (ranging from 0.79 to 0.84 across both quarters and groups). Interviewers who were assigned to use the GPS logging app in Q1 visited statistically significantly more segments in Q1 than in Q2 ($t = 2.09$, $p = 0.04$); there was no difference in calling in Q1 versus Q2 for interviewers who were assigned to use the GPS logging app in Q2. Over-

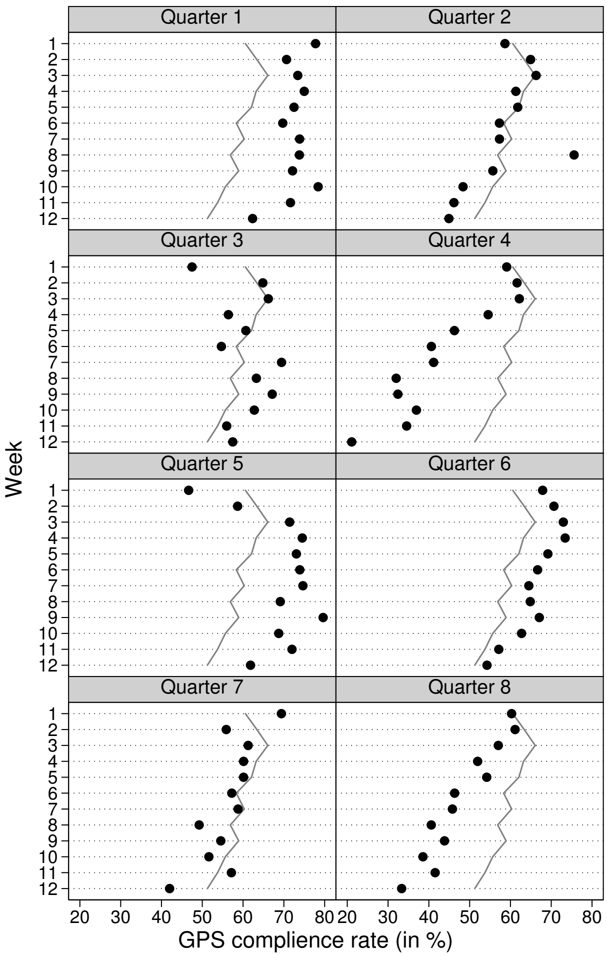

Figure 2. Interviewer compliance rates with GPS logging app, by week in the field period and quarter of data collection with line representing all quarters, NSFG Q1-Q8

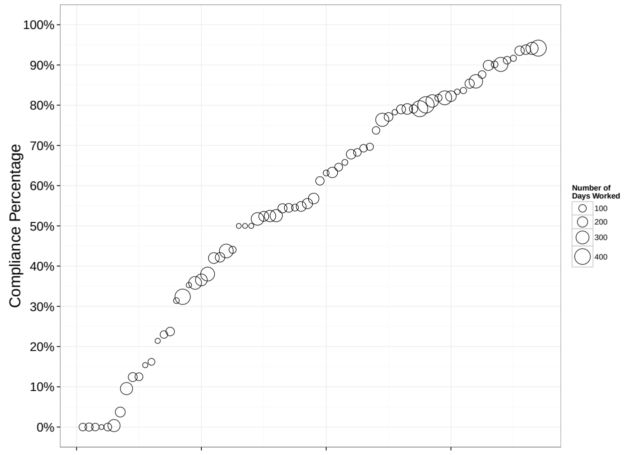

Figure 3. Interviewer GPS Compliance Rates, NSFG Continuous 2011-2019, Q1-Q8

all, interviewers who were assigned to use the GPS logging application had statistically similar calling behaviors than interviewers who were not assigned to use this application.

3.2 Compliance with the GPS App Instruction

Next, we evaluate variability in compliance with the GPS instruction across interviewers. Interviewers varied greatly in the rate with which they complied with the GPS logging application instruction (Figure 3). In this context, compliance involves turning the application on and off during any day that any face-to-face call attempts were made and submitting the generated files. We requested that these files be submitted daily, but did not count late submissions as non-compliant. Across the eight quarters and 74 interviewers, five (6.8%) interviewers failed to use the GPS recording device on any days, and none used the application on all of the days for which face-to-face call attempts were recorded. Of the 74 interviewers, 46 (66.7%) submitted files more than half of the days on which they recorded face-to-face call attempts. The average rate of submission across all interviewers is 55.2% - that is, the average interviewer submitted a file for just over half of the days on which they made any face-to-face call attempts, and there are an average of 103.88 days ($SD = 90.96$) of GPS data among interviewers who complied with the GPS instruction. There is no difference in compliance rates for experienced versus inexperienced interviewers (Rao-Scott $\chi^2 = 0.20$, 1 df, $p = 0.65$) or for interviewers in urban versus other areas (Rao-Scott $\chi^2 = 0.03$, 1 df, $p = 0.87$).

Because interviewer compliance with the instruction to use the GPS logging app varied between and within interviewers, an important question is whether the call outcomes on the interviewer-days with GPS recordings are meaningfully different from those without GPS recordings. We examine outcomes of calls during the screening and main interview components of the NSFG fieldwork – the number of call attempts, cases attempted, completes, contacts, and appointments.

As shown in Table 3, for virtually all of these outcomes, in both the screener and main interview stage of recruitment, the interviewer-days on which the GPS measurements are available are larger in number than those for which there is no GPS measurement available. That is, on interviewer-days with a GPS recording, interviewers attempted significantly more cases for screener interviews ($diff = 1.57$, $t = 2.04$, $p = 0.05$), had significantly more completed screener (diff=0.60, $t = 5.01$, $p < 0.0001$ and main ($diff = 0.20$, $t = 3.66$, $p = 0.0005$) interviews, and had significantly more cases with contact for both screener ($diff = 0.93$, $t = 6.35$, $p < 0.0001$) and main interviews ($diff = 0.58$, $t = 3.55$, $p = 0.0007$). This could be because the use of the GPS

Table 2
Mean number of calls, contact rate, cooperation rate and mean number of segments per trip

	Assigned GPS App Q1			Assigned GPS App Q2			Assigned Q1 vs. Assigned Q2	
	Q1[a]	Q2	F	Q1	Q2[a]	F	F	F
Calls (FtF + Tel)	6.77	6.53	1.22	5.76	5.74	0.00	3.38	0.91
Calls (FtF only)	6.08	5.77	0.39	5.31	5.03	0.27	4.12	0.39
Contact Rate	0.37	0.34		0.39	0.39		3.78[b]	
Cooperation Rate	0.80	0.79		0.80	0.84		4.11[b]	
Segments Per Day	1.57	1.46	4.40*	1.46	1.42	1.05	5.18*	0.00

Contact rate = $\frac{\text{\# calls with contact}}{\text{\# calls}}$; Cooperation rate = $\frac{\text{\# completed interviews or screeners}}{\text{\# calls with screener, interview or refusal}}$
[a] Quarter in which the group of interviewers was assigned to use the GPS logging app. NSFG interviewers could use the telephone to contact screened eligible sampled persons to schedule appointments.
[b] Rao-Scott chi-square test is for overall association between experimental group×quarter and dependent variable.
* $p < .05$ ** $p < .01$ *** $p < .001$

device increased interviewers' awareness of their calling behaviors and thus they did more calling on these days. Alternatively, interviewers may have been more likely to remember to turn on the GPS app when they were planning on working longer shifts or anticipating greater productivity, rather than simply going to a single housing unit (i.e., to fulfill an appointment time). We do not have data available to disentangle these two mechanisms. Interviewers attempted approximately 1.5 more cases for screener interviews, on average, on days with GPS logging information available than on days without it available (11.3 vs. 9.8 cases, respectively). There was no difference in contact rates for days when these files were available compared to when they were not collected.

3.3 Interviewer Survey

We conducted a web survey of the NSFG interviewers about the smartphone and GPS logging app and travel-related behaviors during Quarters 4 and 8 after interviewers had used them for about one year. The Quarter 4 survey was conducted anonymously because of a technical error. Because of this anonymity, we cannot link the sample file to the field data. To evaluate the quality of the Q8 data, we start by examining whether the subset of interviewers who participated in the Q8 survey differ from those who did not. There are few statistically significant differences in field outcomes and productivity measures between interviewers who participated in the Q8 survey and those who did not (Table 4), with the only statistically significant difference being on the number of hours worked ($t = -2.74$, $p = 0.009$) –those who responded worked more hours than those who did not. Interviewers who completed the Q8 interviewer survey had somewhat, but not significantly higher GPS compliance rates (50% for survey nonrespondents, 62.5% for survey respondents, $t = -1.41$, $p = 0.16$).

We now turn to evaluating the survey data itself. Satisfaction levels were similar in Q4 and Q8. About forty percent (Q4; 40% Q8) reported being either "satisfied" or "very satisfied" with the smartphone in both quarters, and roughly one-third (31% Q4; 36% Q8) of the interviewers reported using the voice and calling capabilities of the smartphone "always" or "most of the time." Use of the GPS and mapping abilities changed over the two surveys. In Q4, these capabilities were used "always" or "most of the time" by 21% of interviewers, increasing to 36% in Q8. In comparison, 72% (Q4; 84% Q8) of the interviewers reported using the GPS logger application "always" or "most of the time," with only 10% (Q4, 8% Q8) reporting using it "hardly ever" or "never."

Overall, interviewers who used the phone's voice and GPS capabilities were more satisfied with it. In Quarter 4, over half of interviewers who reported being satisfied with the phone used the phone's voice/calling or GPS/mapping capabilities "always" or "most of the time." In contrast, at least half of the interviewers who were dissatisfied with the phone reported that they "hardly ever" or "never" used these features (voice: 50%, GPS/mapping: 62.5%). These associations are significant (Fisher's exact test $p < 0.02$ for both analyses), and hold in Q8. We have anecdotal evidence that this lack of satisfaction could be due to lack of knowledge in how to use the calling or mapping features on the phone and not previously having a smartphone, although we cannot evaluate these hypotheses with these data.

Failure to use the GPS logging app in the smartphone resulted largely from technical problems, not from discomfort with having movements tracked via the GPS device. Forgetting to turn the phone on (73% Q4, 67% Q8) and the battery dying (54% Q4; 59% Q8) were the most common reasons reported by interviewers for not using the smartphone at some point during the field period; only 10% in each quarter

Table 3
Field Outcomes by Availability of GPS file

	No GPS file available	GBS file available	t-value
Screener			
# of calls	9.75	11.33	-2.04*
# cases attempted	8.33	9.89	-2.07*
# completes	1.74	2.34	-5.01***
# cases with contact	2.51	3.44	-6.35***
# appointments	0.01	0.01	-0.36
Main Interview			
# of calls	4.26	4.72	-1.48
# cases attempted	3.49	3.94	-1.65
# completes	0.64	0.84	-3.66***
# cases with contact	2.28	2.86	-3.55***
# appointments	0.43	0.58	-3.44***
Contact Rate (in %)	42.5	44.9	-1.08
N	5019	7168	

Tests account for clustering of field outcomes within interviewer
* $p < .05$ ** $p < .01$ *** $p < .001$

Table 4
Productivity measures and field outcomes for respondents and nonrespondents to interviewer web survey, NSFG, Quarter 8

	Respondents	Nonrespondents	t-test	p-value
Field Outcomes:				
Screener Response Rate (in %)	92	93	0.80	0.43
Main Interview Rate (in %)	78	80	0.76	0.45
Number of Main Interviews	199.3	162.1	−1.09	0.28
Number of hours	2154.9	1474.4	−2.74	0.01
Number of miles	21887.2	15265.9	−1.84	0.07
Hours per interview	12.2	10.5	−1.57	0.12
Miles per interview	128.5	106.5	−1.14	0.26
GPS Compliance Rate (in %)	62.5	50.0	−1.41	0.16
Number of interviewers	23	29		

Interviewers with fewer than 10 interviews excluded from this table.

reported not knowing how to turn the app itself on. Interestingly, no interviewer reported that they did not want their movements recorded in either survey.

We can look at whether there are associations between responses to the interviewer survey and GPS compliance rates for a subset of interviewers ($n = 25$). The only variable that is significantly associated with GPS compliance rates is self-reported app use. Interviewers who report using the GPS app more often have higher GPS compliance rates ($F = 7.13$, $p = 0.0017$). There is a general trend for higher compliance rates for interviewers who report forgetting the phone, forgetting to turn it on, and report the battery dying, likely reflecting that these interviewers are more conscientious about the task or more likely to be aware that they have forgotten the phone occasionally (Appendix Table 1).

3.4 GPS Data Quality

Using GPS to follow interviewer movements in the field is a worthwhile method only if the data collected are of high enough quality to be used efficiently. There are three aspects of data quality – whether the amount of erroneous data is limited to minimize the number of post-processing hours required for using these data, whether there is adequate coverage of the movements of interviewers when the GPS device is turned on, and whether the quality of the GPS points are such that we can detect meaningful travel patterns through sampled neighborhoods. We now briefly evaluate the quality of the data recorded by the GPS app specifically related to the interviewers' actions (the first two aspects), rather than the technical limitations of the smartphone or the GPS Logging app (the last aspect).

We start with a simple method to detect erroneous data points. We would not expect interviewers to work during late night or early morning hours. The time that the points were recorded is available for 1,943,764 GPS points. Of these approximately 1.9 million points, 11.3% occurred between 9:00 PM and 6:00 AM, times that interviewers are unlikely to be working. During qualitative interviews with interviewers, we learned that many interviewers simply forgot to turn off the GPS device. During more traditional work hours, 33.4% of the points were recorded between 7:00 AM and 12:00 PM, 43.0% between 1:00 PM and 5:00 PM, and 12.3% were recorded between 6:00 PM and 8:00 PM.

Another measure of data quality is the length of time that the GPS file was recording interviewer movements. This is one aspect of adequate coverage of interviewer movements. We calculated the total number of minutes in each GPS log file. The total times recorded range from 0 minutes to about 8320 minutes (over 138 hours or nearly 6 days). The mean number of minutes in each file is 408 minutes (6.8 hours) and the median is 349 minutes (5.8 hours). We compare all of these GPS files to the number of hours reported in all of the timesheet data, which include days for which interviewers did not travel. We expect that the time recorded in the GPS files will be *shorter* than the timesheet data because the GPS files include only days the interviewer travelled to area segments, whereas the timesheet data include both travel and non-travel days (i.e., days with telephone meetings or administrative tasks rather than field visits). On average, interviewers report working for 4.6 hours across all of the days with hours submitted, including both travel and non-travel days. Thus, the GPS files are on average *longer* than the hours reported in the timesheet data, opposite our hypothesis. This could be due to the GPS files collecting data during non-working hours. When we exclude the GPS files that were collected during non-working hours (between 9:00 PM and 6:00 AM), the length of the GPS files drops to an average of 342 minutes (5.7 hours), with a median of about 326 minutes (5.4 hours), still longer than the timesheet data. This could also be due to the tendency for the GPS data to be collected on days when more calls were made.

When we instead compare the length of the hours worked to the length of the GPS files on days for which we have *both* GPS data and timesheet data, we see a slightly different picture. On the days where we have both sources of data, the average length of time worked is virtually identical in both files – 7.04 hours in the timesheets compared to 6.86 hours in the GPS files, an average difference of only 0.18 hours ($SD = 5.26$), or about 11 minutes. The range of the differences between the two files is quite large – from the GPS file being over 50 hours longer than the time recorded in the timesheets to the timesheets having almost 16 hours more time indicated than the GPS files.

A final comparison for whether the GPS files adequately cover interviewer movements is the number of miles traveled. The GPS files indicate that the interviewers travel an average of 149.5 miles per logging file with a median distance of 51.98 miles. In contrast, the average distance reported in the timesheets (that is, the mileage submitted for reimbursement) is 77.80 miles, with a median of 66.00 miles. Thus, the average GPS file has distances roughly twice that of the timesheet data, but the GPS median distance is slightly lower than the timesheet median distance.

Overall, it appears that, when the GPS device is turned on, the files collected cover the hours that are being worked and the miles traveled, but that at least 10% of the points recorded are erroneously collected during non-work hours.

3.5 What do the GPS data look like?

The GPS data collected are a rich source of information about interviewers' travel patterns. Figure 4 displays a simulated example of the GPS data that were collected with the GPS Logger App. The GPS data are the small numbered circles. Numbers indicate the temporal order in which the GPS point was recorded. Housing units are shown with stars, and sampled housing units are circled. The black line that begins at point 42 and ends at point 55 indicates a route that was created using GIS software that could have been followed from points 42 and ending at point 55.

A few qualitative observations can be made from our initial examination of the GPS data. First, interviewers do not call on sampled housing units in the order that they appear in the segment. Instead, they travel throughout the segment, skipping over sampled housing units, and retracing their path. Second, as shown in the cluster of dots at the top of Figure 4, the GPS data jump around even when the interviewer is stationary. Third, we cannot tell what side of the street the interviewer is on from the GPS data, potentially complicating linking issues when sampled housing units are on both sides of a street.

Cleaning the GPS data for analysis is a highly time consuming task. Chung and Shalaby (2005) report seven preprocessing steps to link raw GPS data to GIS coordinates, and an additional 11 steps to identify the mode of transportation. An additional challenge for survey research is that the GPS data are not automatically linked to the call records, and errors in both the call records and GPS data mean that there is not a one-to-one relationship (Wagner, Olson, & Edgar, 2013). This linkage and future research will be critical to understanding how interviewers move in the field, how to control travel costs, and what kinds of travel decisions are related to increased efficiency and better field outcomes.

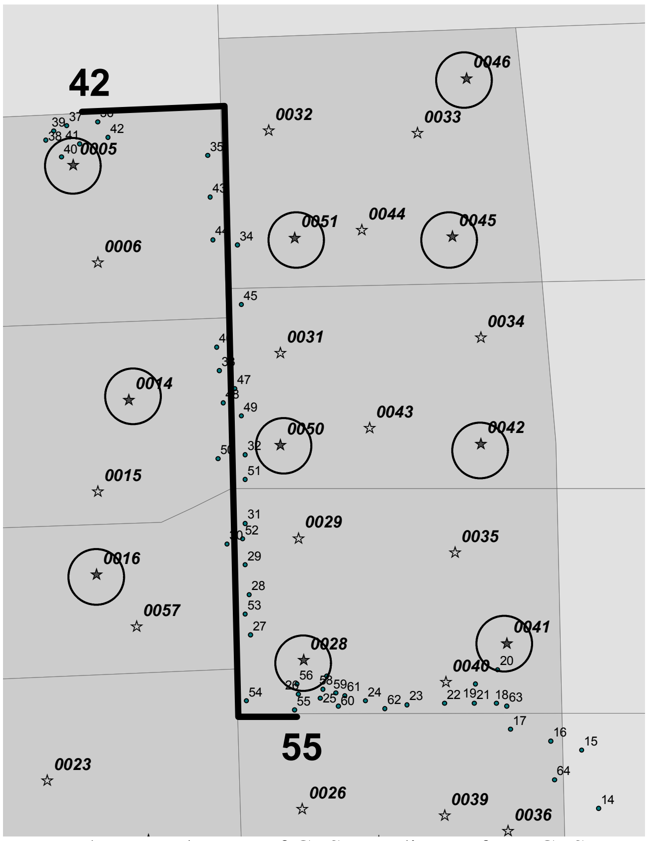

Figure 4. Example map of GPS coordinates from GPS Logger app

4 Discussion

In this paper, we have examined results from a field test of a potentially useful new source of data provided by asking interviewers to record their movements using a GPS logging application on a smartphone. Are GPS-enabled smartphone apps a feasible method for examining interviewer travel? This feasibility test demonstrates the answer is a qualified yes. There are numerous limitations and challenges related to interviewer compliance. Use of the logging app had only modest effects on field outcomes, primarily on days that the app was used. The interviewers who failed to use the logging app tended to do so because of technical problems with the phone, not due to privacy concerns about having their movements recorded, or because they were less satisfied with the phone technology. We have anecdotal evidence that the interviewers who were less familiar with technology or smartphones had more difficulty with the phones, a problem likely to decrease as smartphone use increases. In general, these data appear to usefully supplement, but not replace, other current sources of data. The data were collected via a publicly available smartphone app for Android phones, not a proprietary GPS device. This means that other survey organizations could easily use this app.

The use of smartphones in the field to collect GPS data opens additional opportunities for using these devices for other purposes. Sample management systems can be programmed into smartphones for delivery of sample information without having to turn on a large laptop (e.g. Kennet & Gfroerer, 2005). Interviewers can take discrete observations of sampled housing units with a smartphone, and even take photographs of the housing unit (e.g. Dekker et al., 2013; Keating et al., 2014). The increased use of mobile devices for sample management procedures may improve the quality of call records when they can be recorded without logging into a laptop. If use of a smartphone with a GPS-enabled app makes interviewers aware of being monitored, use of mobile devices may also reduce curbstoning. GPS systems could be integrated with sample management systems to evaluate whether the interviewer is at the correct housing unit, or whether they passed a sampled housing unit without making a call record, and even alert them by beeping or other methods when they are near a sampled housing unit that needs to be screened or interviewed. Further, having both sets of data collected by a single device would greatly improve the linkage of the GPS data to the call records. Future research could evaluate all of these issues.

There are limitations of these data. First, although use of smartphones is increasingly common in in-person field surveys, the phones and monthly data plan are a non-trivial additional study cost. Second, we found moderate levels of compliance with the instruction to record these data, with great variability across interviewers. Full compliance would maximize the utility of these data, but such compliance rates would require removing interviewer decisions to turn the app on and off, potentially compromising interviewer privacy. Alternatively, the CAPI instrument could be programmed such that GPS recordings are automatically taken when the interviewer starts and ends the interview. This approach, however, would not capture movements in the field. Third, there are errors in the data (e.g., extremely long files) that limit their usefulness and require substantial post-collection processing. Errors in conventional paradata are relatively unstudied, and as such, we do not know if the GPS errors are greater than the errors in other paradata such as call records. Fourth, the off-the-shelf app had software updates that we did not control and resulted in app errors. Custom-built apps would not have this problem, but would require substantially greater resources. Fifth, the interviewers in this survey were paid by hour, not by the complete. Interviewer payment schemes can affect field outcomes (e.g. Stoop, Billiet, Koch, & Fitzgerald, 2010; Tourangeau, Kreuter, & Eckman, 2012), and would likely also affect travel decisions. We cannot directly evaluate this hypothesis from these data. Finally, these data were collected in one large-scale national (US) survey. We do not know how these findings would apply to other smaller surveys or to surveys in other countries. Of course, GPS data collection of interviewer movement in other countries may have other legal requirements that should be examined prior to launching this kind of study.

This paper did not examine the quality of the GPS data themselves beyond the length and time of the files. An initial examination has shown that some of the data are excellent quality and some are unusable (see Wagner et al., 2013). We hypothesize that poor quality data result when interviewers are inside conducting an interview, from certain geographic areas that may make acquiring a GPS signal difficult (e.g., urban canyon, dense tree cover), or when traveling at high speeds.

This study found substantial variability in interviewer compliance with the instruction to use the GPS app. In this study, the staffing capacity of human monitors varied over the quarters. Future research could automate reminders to the interviewers to use the app during the day and automatically inform the interviewer when their GPS information is not transmitted each night in order to increase compliance. Alternatively, the daily reports used by field supervisors could include each interviewer's GPS submission rate, allowing supervisors to assist with monitoring without automating the reminders. The study also found variability that may be related to the interviewers' own familiarity with smartphones and other mobile technologies, although we cannot empirically evaluate this hypothesis. Additional training refresher modules on the phone itself could be planned throughout data collection to help increase compliance and to assist those interviewers having technical difficulties.

Despite the limitations, GPS data have a number of potential uses. These new data can be compared with existing data generated from timesheets and call records. Analyses of this type may further illuminate the weaknesses (and strengths) of call record data. For instance, GPS data may help researchers identify underreporting of call records by noticing when interviewers pass by sampled housing units without making a call record (e.g. Wang & Biemer, 2010). If the GPS data highlight weaknesses of call records, then data collection organizations may want to improve the quality of call records or routinely supplement the call records with GPS data.

These data also may be used to identify interviewing strategies that are more efficient. For instance, a key question is whether interviewers should follow a route through area segments that minimizes the distance traveled, or whether interviewers should pay attention to clues in the neighborhood about who may be at home and tailor their route to these clues, thereby increasing the distance traveled. Such strategies may be identified using the GPS data, and compared to conclusions made from geocoding the call record data. We are currently linking these two systems to understand how different types of travel behaviors in the field are related to field outcomes such as contact and cooperation rates.

Acknowledgements

An earlier version of this paper was presented at the American Association for Public Opinion Research Annual Meeting, May 2013, Boston, MA. The authors thank the NSFG senior staff for funding the smartphone data collection and Minako Edgar for her GIS expertise.

References

Bates, N., Dahlhamer, J., Phipps, P., Safir, A., & Tan, L. (2010). Assessing contact history paradata quality across several federal surveys. *Proceedings of the American Statistical Association, Survey Research Methods Section*, 91–105.

Biemer, P., Chen, P., & Wang, K. (2013). Using level-of-effort paradata in non-response adjustments with application to field surveys. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, *176*, 147–168.

Bienias, J. L., Sweet, E., & Alexander, C. (1990). A model for simulating interviewer travel costs for different cluster sizes. *Proceedings of the American Statistical Association, Survey Research Methods Section*, 20–27.

Billiet, J. & Loosveldt, G. (1988). Improvement of the quality of responses to factual survey questions by interviewer training. *Public Opinion Quarterly*, *52*, 190–211.

Bricka, S. (2008). *Nonresponse challenges in GPS-based surveys*. Paper presented at Non-Response Challenges in GPS-based Surveys Workshop Annecy, France. May 2008.

Campanelli, P., P.and Sturgis & Purdon, S. (1997). *Can you hear me knocking: an investigation into the impact of interviewers on survey response rates*. London: Survey Methods Centre at SCPR.

Cecchi, R. & Marquette, R. (2012). *2010 census: Global Positioning System evaluation report.* Washington, DC: US Census Bureau.

Chen, B.-C. (2012). *Simulating nhis field operations.* Proceedings of the 2012 Research Conference, Federal Committee on Statistical Methodology. Office of Management and Budget. Washington, DC. Retrieved from https://fcsm.sites.usa.gov/files/2014/05/Chen_2012FCSM_II-A.pdf

Chung, E.-H. & Shalaby, A. (2005). A trip reconstruction tool for GPS-based personal travel surveys. *Transportation Planning and Technology*, *28*, 381–401.

Couper, M. (1998). Measuring survey quality in a CASIC environment. *Proceedings of the American Statistical Association, Survey Research Methods Section*, 41–49.

Couper, M. (2013). Is the sky falling? New technology, changing media, and the future of surveys. *Survey Research Methods. 7*, 145–156.

Couper, M. & Lyberg, L. (2005). *The use of paradata in survey research*. Proceedings of the 55th Session of the International Statistical Institute, Sydney (CD-Rom).

Dekker, K., English, N., Winfrey, K., & Seeger, J. (2013). *Cost implications of new address listing technology: efficiency and data quality*. Paper presented at the Federal CASIC Workshops, Washington, DC. Retrieved from https://fedcasic.dsd.census.gov/fc2013/ppt/06___DEKKER___FedCASIC_Dekker.pdf

Durrant, G. & Steele, F. (2009). Multilevel modelling of refusal and non-contact in household surveys. Evidence from six UK Government surveys. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, *172*, 361–381.

Ellis, C., Sikes, N., Sage, A., Eyerman, J., & Burke, B. (2011). *Technological advances to reduce survey error*. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Phoenix, AZ.

Fowler, F. & Mangione, T. (1990). *Standardized survey interviewing: minimizing interviewer-related error*. Newbury Park: Sage Publications.

Gong, H., Chen, C., Bialostozky, E., & Lawson, C. (2012). A GPS/GIS method for travel mode detection in New York City. *Computers, Environment and Urban Systems*, *36*, 131–139.

Goodchild, M. (2007). The Morris Hansen lecture 2006: statistical perspectives on spatial social science. *Journal of Official Statistics*, *23*, 269–283.

Haddaway, S. (2013). *CAPI surveys on android devices in the developing world*. Paper presented at the Federal CASIC Workshops, Washington, DC. Retrieved from https://fedcasic.dsd.census.gov/fc2013/ppt/FedCASIC_Haddaway.pdf

Judkins, D., Waksberg, J., & Northrup, D. (1990). Cost functions for NHIS and implications for survey design. *Proceedings of the American Statistical Association, Survey Research Methods Section*, 34–43.

Kalsbeek, W., Botman, S., Massey, J., & Liu, P.-W. (1994). Cost-efficiency and the number of allowable call attempts in the National Health Interview Survey. *Journal of Official Statistics*, *10*, 133–152.

Keating, M., Loftis, C., McMichael, J., & Ridenhour, J. (2014). *New dimensions of mobile data quality*. Paper presented at the Federal CASIC Workshops, Washington, DC. Retrieved from https://fedcasic.dsd.census.gov/fc2014/ppt/Keating%20New%20Dimensions%20of%20Mobile%20Data%20Quality_v1.4%20-%20Evaluation%20FedCASIC2014%20031814.pptx

Kennet, J. & Gfroerer, J. (2005). *Evaluating and improving methods used in the National Survey on Drug Use and Health*. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies.

Kerr, J., Duncan, S., & Schipperjin, J. (2011). Using Global Positioning Systems in health research: a practical approach to data collection and processing. *American Journal of Preventive Medicine*, *41*, 532–540.

Krenn, P., Titze, S., Oja, P., Jones, A., & Ogilvie, D. (2011). Use of global positioning systems to study physical activity and the environment: a systematic review. *American Journal of Preventive Medicine, 41*, 508–515.

Kurkowski, K. (2013). *Using Internet and hand-held computers for data collection in Poland*. Paper presented at The 5th International Workshop on Internet Survey and Survey Methodology, Daejon, Republic of Korea.

Laflamme, F. & Karaganis, M. (2010). *Assessing quality of paradata to better understand the data collection process for CAPI social surveys*. Paper presented at the 2010 European Quality Conference, Helsinki, Finland.

Lemmens, M. (2011). *Geo-information: technologies, applications and the environment*. Dordrecht: Springer Science+Business Media.

Levinsohn, J. R., Medeiros, D. D., Duke, J. C., Yost, P. A., Litavecz, S. D., Zhang, Y., & Karlsen, R. (2010). *The general survey system (GSS): a mobile technologies system for collecting and managing study data*. Paper presented at the The 138th Annual Meeting of the American Public Health Association, Denver, CO. Retrieved from http://www.rti.org/pubs/apha10_medeiros_pres.pdf

Morton, K., Iannacchione, V., McMichael, J., Cajka, J., Curry, R., & Cunningham, D. (2007). Linking mailing addresses to a household sampling frame based on census geography. *Proceedings of the American Statistical Association, Survey Research Methods Section*, 3971–3974.

Morton-Williams, J. (1993). *Interviewer approaches*. Cambridge: Cambridge University Press.

Nusser, S. (2007). Using geospatial information resources in sample surveys. *Journal of Official Statistics*, *23*, 285–289.

Nusser, S. & Fox, J. (2002). *Using digital maps and GPS for planning and navigation in field surveys*. Iowa State University Department of Statistics, Iowa State University.

O'Muircheartaigh, C. & Campanelli, P. (1999). A multilevel exploration of the role of interviewers in survey nonresponse. *Journal of the Royal Statistical Society, A*, *162*, 437–446.

Ongena, Y. & Dijkstra, W. (2006). Methods of behavior coding of survey interviews. *Journal of Official Statistics*, *22*, 419–451.

Pickery, J. & Loosveldt, G. (2002). A multilevel multinomial analysis of interviewer effects on various components of unit nonresponse. *Quality and Quantity*, *36*, 427–437.

Purdon, S., Campanelli, P., & Sturgis, P. (1999). Interviewers' calling strategies on face-to-face interview surveys. *Journal of Official Statistics*, *15*, 199–219.

Rainham, D., Krewski, D., McDowell, I., Sawada, M., & Liekens, B. (2008). Development of a wearable global positioning system for place and health research. *International Journal of Health Geographics*, *7*.

Seeger, J. (2011). *A mobile, GPS-enabled listing application*. Paper presented at the Federal CASIC Workshops, Washington, DC. Retrieved from https://fedcasic.dsd.census.gov/fc2011/ppt/09_seeger.pdf

Sikes, N. (2009). *Current trends in mobile technology for survey research*. Paper presented at the Federal CASIC Workshops, Washington, DC. Retrieved from https://fedcasic.dsd.census.gov/fc2011/ppt/09_sikes.pdf

Sonck, N. & Fernee, H. (2013). *Using smartphones in survey research: a multifunctional tool, implementation of a time use app; a feasibility study*. The Hague: The Netherlands Institute for Social Research SCP.

Stoop, I., Billiet, J., Koch, A., & Fitzgerald, R. (2010). *Improving survey response: lessons learned from the european social survey*. West Sussex, UK: Wiley.

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

Tourangeau, R., Kreuter, F., & Eckman, S. (2012). Motivated underreporting in screening interviews. *Public Opinion Quarterly*, *76*, 453–469.

U.S. Census Bureau. (2014). GPS coordinates. Retrieved January 27, 2015. Retrieved from http://www.census.gov/about/policies/privacy/data_protection/gps_coordinates.html

Wagner, J., Olson, K., & Edgar, M. (2013). *Using GPS and other data to assess errors in level-of-effort data in field surveys*. Paper presented at the Joint Statistical Meetings, Montreal, Canada.

Wang, K. & Biemer, P. (2010). *The accuracy of interview paradata, results from a field investigation*. Paper presented at the Annual Conference of the American Association for Public Opinion Research, Chicago, IL.

Wolf, J., Guensler, R., & Bachman, W. (2001). Elimination of the travel diary, experiment to derive trip purpose from global positioning system travel data. *Transportation Research Record, Journal of the Transportation Research Board*, *1768*, 125–134.

Wu, J., Jiang, C., Liu, Z., Houston, D., Jaimes, G., & McConnell, R. (2010). Performances of different global positioning system devices for time-location tracking in air pollution epidemiological studies. *Environmental Health Insights*, *4*, 93–108.

Zandbergen, P. (2009). Accuracy of iphone locations: a comparison of assisted GPS, wifi and cellular positioning. *Transactions in GIS*, *13*, 5–26.