INTERNATIONAL WORKSHOP ON HOUSEHOLD SURVEY NONRESPONSE

AUGUST 30 – SEPT 1, 2017, UTRECHT

**Introducing mobile device data collection in household surveys**

Authors: Barry Schouten & Peter Lugtig

**Summary:** In 2016, Statistics Netherlands and Utrecht University established a joint research program on ICT innovation in primary data collection (in Dutch: Waarneem-Innovatienetwerk or WIN). In the first few years, WIN will focus on the use of mobile devices (smart phones, tablets, wearables) in primary data collection, both as a survey mode and as an instrument to collect automated measurements through its sensors. This paper motivates the initiation of WIN, presents its main research questions and describes the various research projects. Questions for discussion are included in the final section.

# Motivation

Internet and communication technology have evolved drastically over recent years and have changed the frequency, time and place in which persons exchange information. This trend goes parallel to a more individual, personal life style. As a result persons’ expectations and attitudes about how data is collected for statistical purposes have changed. Traditional data collection modes have not fully lost their attraction; face-to-face interviews still obtain relatively high response. However, new ICT opens up new possibilities and, above all, is relatively inexpensive.

*Figure 1. Percentage of logins through tablets and smart phones*

Statistics Netherlands has changed virtually all its person surveys to multi-mode surveys including Web. Doing so, it paved the way for various forms of online data collection and finds itself in the arena of ICT evolution. In order to stay up to date and to remain linked to communication practice, it has to follow trends closely and adapt, if necessary. Figures 1 and 2 show the percentage of respondents in person surveys that logs in using a mobile device, respectively the percentage of respondents that remains using the device as a function of interview duration. In Figure 1, smart phones and tablets are distinguished over time. In Figure 2, as a benchmark also the break-off in desk/laptop is presented next to smart phones and tablets. Clearly, there is a strong increase in log-in rates but also a strong drop-out during online interviews. To date, only one CBS survey has been made fit for response on all possible devices. This is the Consumer Sentiments Survey, which is short (5 to 10 minutes) and never had a face-to-face option in the past. All other surveys are not designed for response on mobile devices and are usually challenging to fill in on a mobile device. Nonetheless, an increasing part of the respondents tries to complete the survey on their mobile device.

*Figure 2. Percentage of logins per device type that has not broken off as a function of interview time.*



Mobile devices have become standard tools for communication, but they also support a range of automated measurements through their sensors. The range and type of sensors depend on the platform (Android and iOS) and model, but there is a fair list of sensors that is shared by all. The appendix contains an overview of available sensors. Many mobile device applications use at least one of these sensors, such as GPS or accelerometers. Employing multiple sensors and combining with advanced software, mobile devices can be powerful tools to collect data. In brainstorm sessions at Statistics Netherlands, six topics have been identified as promising for a combination of survey and sensor data: 1) Travel and time use, 2) Internet behaviour and use, 3) Expenses and buying behaviour, 4) Health and fitness, 5) Living and working conditions, and 6) Mental and emotional states. These topics are hard or cognitively burdensome to fully measure by means of a questionnaire and may arouse relatively strong measurement error. Furthermore, these topics already attract a lot of research into the potential of sensor data, partially as in-depth, qualitative measurements and partially as big data.

Obviously, mobile device sensors do not allow for measuring all possible survey topics unless they become very intrusive and extensive. It is here, where the biggest challenge is in innovating primary data collection: Communication by mobile devices has become more frequent, but shorter and more diverse, whereas sensor measurement, especially when combined with a survey, usually require commitment and duration.

WIN explores the most promising and relevant avenues for innovation and attempts to prepare for implementation into survey practice.

# Research questions and projects

The three main research questions for employing mobile devices are:

* How to decide when to block, discourage, accept or stimulate responses through tablets and smart phones?
* How to combine mobile devices with traditional modes and channels?
* How to combine mobile device survey and sensor data?

These three questions open up a lot of detailed questions.

The first main research question is central in (re)designing surveys for multiple devices. There are two motives for being more or less progressive in offering mobile device as an option: device effects and additional operational effort to build and maintain responsive questionnaires. Mobile devices are self-administered, visual and computer-assisted just like laptops and desktops, but have smaller screen sizes and are used at a larger variety of times and locations. They may be more similar to desktops and laptops, but they share the volatility and pace of telephone interviews. As a consequence, it is anticipated that surveys should be shorter in order to avoid break-off and lower data quality than in other devices. This is reflected by the (substantially) higher amount of survey break off in mobile devices, especially in smartphones. Furthermore, device selection effects are likely because of preferences and experiences that respondents may have. These potential device effects make survey managers reluctant to embrace the mobile device option, especially for the longer surveys. In order to make devices more attractive, insight in device measurement effects is needed and effective data collection strategies need to be developed.

The other motive, the potential additional effort in multi-device and multi-mode questionnaire design, data collection monitoring and processing further hampers experimentation and implementation. This additional effort is in part, at least at Statistics Netherlands, a legacy of existing systems and processes.

Official surveys are often repeated and have a history in traditional survey modes. Online response rates are relatively low, so that these surveys are mostly sequential mixed-mode. This means that these surveys do not have a mobile device first design, and mobile devices may add more complexity to design and analysis. To date, for this reason, there still is relatively little empirical knowledge about device effects in official, cross-sectional surveys of the general population.

Summarizing, there is a need for effective data collection strategies, responsive questionnaire design and lay-out for mobile device questionnaires.

Given that mobile devices are a complementary mode/option, the second main question is about the inclusion of mobile devices in a multi-mode survey. Apart from the potential device effects and incomparability in time and between population publication domains, there is a general intuition that mobile devices demand for shorter questionnaires. The optimal length depends on the topic and respondent, but drop-out rates per survey item are known to be higher for tablets and even higher for smart phones. As a consequence, multi-device and multi-mode surveys revive the interest in so-called split questionnaire designs and modular survey designs; different sample persons are randomly allocated to different shorter versions of the questionnaires that contain subsets of the survey items or survey modules.

Split questionnaire designs and modular survey designs go back to the 90’s, but have never been implemented at a large scale due to their operational and inferential complexity. The number of possible shorter questionnaires can be (combinatorically) very large and a smaller number has to be chosen in order to become operationally feasible. However, even for a small number of versions, questionnaire development and maintenance may quickly become prohibitive. Split questionnaire and modular survey designs introduce missing data on purpose. The users of the survey data, therefore, have to change their analysis strategy, i.e. through modelling and/or imputation, in order to extract as much information as possible from the data. In other words, the survey analysts pay the price for the new data collection strategy. The split questionnaires or modular designs need to be designed to yield optimal statistical power, while attempting to keep measurement context effects as small as possible. This is methodologically very challenging. Nevertheless, for mobile devices, such an approach may have a much stronger business case, because of the drop-out rates and anticipated lower data quality.

Summarizing, there is a need to explore the boundaries of split questionnaire and modular survey design in a multi-device and multi-mode setting.

The third main research question is relatively wide and concerns the use of mobile device sensor data. Within WIN, only primary data collection is considered, i.e. sensor data are collected for a random sample drawn from the population registers. Sensor data can, however, also be part of the Internet-of-Things (IoT) and are then secondary data collection. IoT sensor data are typically explored within big data research, but also sensor data for a sample may be too big to be handled by conventional methods and software. The main distinction with big data research is that for primary data collection, costs are (primarily) to the statistical institute that is collecting them. Sensor data can be collected through the devices that are owned by respondents or by devices that are (temporarily) provided to them. More advanced and accurate measurements require more expensive sensors that have a lower population coverage or tailor-made sensors. Whether such sensors can be provided depends on the self-perceived role and identity of the statistical institute. Although in recent years Statistics Netherlands has become more pro-active in explaining survey statistics, it still has a more descriptive, exploratory than an analytic, explanatory role. As a consequence, surveys are quantitative rather than qualitative and sensor data per sample unit have to be relatively inexpensive or must lean on devices and wearables that have a high population coverage. Research, therefore, has to find a balance between accuracy and relevance of sensor measurements and costs.

Mobile device sensor data require a different data collection infrastructure and different data processing, i.e. an investment in both systems and researchers/analysts. The strongest business case for sensor data, therefore, comes from questionnaire topics that are demanding in terms of respondent effort, time and/or cognition but that seem within reach for sensors. Out of the six themes that are identified by Statistics Netherlands’ survey coordinators, the most promising are sensor data around mobility/travel using GPS, Wifi and accelerometer, for use in travel surveys and time use surveys. Travel data have already been explored extensively and commercial applications are available. However, since decision rules in such applications are not fully transparent and since additional information needs to be asked about the purpose and context of the travels, WIN is developing its own proof-of-concept using open source components. Two other themes where WIN is actively investigating sensor data measurements are budget expenditure and ICT/internet behaviour. Within Eurostat there is an interest in innovating budget expenditure surveys, given the often low response rates, high response burden and data quality concerns of these surveys. The business case for budget expenditure, therefore, is very strong. Sensor data measurements can be collected using tailor-made applications or using plug-ins in browser-oriented questionnaires. The latter offer the advantage that respondents can visit the regular website and do not have to download and install an application. For ICT/internet behaviour, the browser-oriented approach is explored.

A theme that is considered highly relevant and promising but complex and multi-dimensional is health and life style. In this theme, a lot of (inter)national parties are already active. It is promising, as respondents may see the benefits of collecting such data and learn about themselves. For these reasons, we will start by making an inventory of the opportunities and risks in this theme, together with other institutes. As part of the inventory, a sensor data challenge (a so called hackathon) is organized in September 2017, where teams compete to develop health-related sensor measurements that are relatively inexpensive or use omnipresent sensors.

Obviously, respondents need to consent to provide (and link) sensor data. Some sensors, such as the accelerometer, do not need a consent, technically, and some sensors, such as GPS or Wifi, may already be enabled by the respondent him/herself. Other sensors, such as audio or video, always need explicit consent. However, regardless of the type of sensor, it is deemed ethical to inform respondents and to explicitly ask for consent. In 2017-2018, two sensor data consent surveys are fielded under varying conditions to explore the best data collection strategies and to (de)select themes based on the willingness of respondents to open up sensor data.

Finally, apart from ethical and methodological issues, there is also the confidentiality of sensor data. In theory, when secure data transmission and storage can be assured, it may seem that sensor data are just like traditional survey data. However, the main differences lie in the automation of the data collection, i.e. respondents do not fully know what data is collected and how it is used, and in the availability of the sensors to other applications and data collectors when consent is given. Confidentiality questions have been posed to Statistics Netherlands’ legal officers.

Summarizing, mobile device sensordata research organized around the combination of survey data and sensor data; it investigates ethical and privacy issues, effective consent strategies, applications that integrate sensor measurements and survey data, and explores the trade-off between availability and provision of sensors and costs.

# Current activities

1. Mobile device design and lay-out:
	* Usability pre-tests to compare mobile devices and traditional devices;
	* Analysis of break-off rates and data quality as a function of interview duration and survey item characteristics in the Consumer Sentiments survey and Health survey;
	* Development of effective data collection strategies (QR-codes, advance letters);
	* Mobile device field test using randomized allocation to a number of style options in the Health survey;
2. Mobile device split questionnaire and modular survey design:
	* Using various methods to split questionnaires linked to existing surveys, and come up with one or two promising versions;
	* Qualitative research into respondent experience and context effects;
	* Simulation of statistical power of split questionnaire/modular survey design missing data linked to the Crime Victimization survey and the Health survey;
	* Empirical insight into optimal questionnaire lengths for mobile devices, possibly as a function of survey topics;
	* An investigation into response rates and measurement context effects using a split questionnaire/modular survey design field test;
3. Mobile device sensor data:
	* Development of a tailor-made application for (re)location measurement;
	* Test of (re)location sensor measurements with face-to-face interviewers;
	* Development of browser-oriented sensor measurements linked to ICT/Internet behaviour;
	* Exploration of a tailor-made application for budget expenditure;
	* A sensor data challenge on health and life style subjects;
	* Development of a sensor data collection strategy, using two experimental sensor data consent surveys;
	* Legal advice in confidentiality and privacy issues surrounding mobile device sensor data;

**Discussion**

Questions for discussion:

1. Should we consider a mobile device to be a separate mode in terms of selection and measurement properties? If so, how to estimate and disentangle device effects?
2. How to decide whether to block, discourage, accept or stimulate mobile device response in a survey with Web as one of the modes?
3. Does accepting or stimulating mobile device response imply split questionnaire/modular survey designs for longer surveys?
4. Would you consider split questionnaire/modular survey designs for your organisation (and why (not)), given?
5. How to implement split questionnaire/modular survey designs to minimize the burden for users of surveys?
6. How to find the balance between high quality, but expensive and relatively inexpensive sensor data that are available in larger quantity?
7. What would be sensible strategies to obtain consent for sensor data measurements?
8. Do sensor data measurements have ethical and privacy consequences beyond those of online data collection using a traditional questionnaire?

**APPENDIX: Overview of sensors in mobile devices**

Overview of the most common sensors and feedback mechanisms found in wearables and cell phones.

**3D touch**

This sensor measures the pressure exerted on the screen. Small objects up to 385 grams can be weighted. Only in iPhone 6S and 6S plus.

**Accelerometer, gravity, gyroscope**

A set of sensors measuring motion, acceleration and position of the device. Used for position tracking or step counters.

**Ambient Light**

Measures the intensity of the ambient light. More advanced versions also determine the light colour or –temperature. Commonly used to adapt the screen brightness and colour to the ambient conditions.

## Bluetooth

Wireless communication protocol. Can connect to small low energy devices, such as wireless headphones, key fobs, smartwatches or smart scales. Also detects the presence of Bluetooth beacons.

**Camera**

Takes pictures or videos, or measures light intensity. Usable for image- or pattern recognition, scanning of QR- or barcodes and colour analysis. Can also coarsely measure gamma rays, a form of radioactivity.

**Camera Flash**

Usually used as a flashlight or as illumination for pictures or videos. Can create stimuli that can be picked up by other sensors.

**Cellular**

The core of all cell phones. Used to make and receive calls and text messages. Strength and ID of cell tower broadcasts can be measured. With the knowledge of tower positions the user location can be determined with a precision of ~500 meters. More advanced cell phones – almost all cell phones today – can also connect to the internet. The presence of the internet connection as well as its speed (upload, download, responsiveness/ping) can be determined. Not often found in wearables yet.

**Fingerprint**

Some devices are capable of detecting fingerprints. The raw data is not accessible, but it can be used as identification or simply as a button.

**GPS**

The Global Positioning System. Dozens of GPS satellites circle the earth and broadcast beacon signals. By measuring the time-of-flight of the satellite signals, the distance to that satellite can be calculated. With four or more satellites visible, the position on earth can be triangulated. The precision is ~5 meters outdoors. Indoor performance is poor. The satellite signals also includes time information.

**Heart Rate**

Measures the heart rate, usually optically on the finger (cell phone), wrist (smartwatch) or in-ear (headphones).

**Humidity**

Measures ambient humidity. Not very widely used yet

**Magnetic Field**

Usually used as a compass, but can also measure the strength of magnetic fields or can be used, within limits, as a metal detector.

**Microphone**

Detects speech and sounds that can be saved, streamed or analysed. Can also determine loudness and detect ambient noise. Multiple microphones in one device allow for determining the directionality or distance of sound sources. This is used to filter out ambient noise in phone calls. Microphones can also be used to record the heartbeat or estimate lung function/spirometry.

**NFC (near field communication)**

The same technology as contactless payments. Can be used to pay with the cell phone or smart watch, as “contactless QR code” / “NFC tags” to change phone settings (muting) or trigger the start of certain apps. A phone can be a tag as well, so two devices can identify each other and initiate a data channel for communication.

**Pressure**

This sensor measures the ambient air pressure and functions as a barometer. The precision is so high, that height differences of a few meters (ground floor vs first floor) can be detected. Is also used in combination with GPS for more precise height determination.

**Proximity**

Measures the presence of objects close to the screen, usually binary (object present or not). Switches off screen and touchscreen during phone calls.

**Screen**

Display of static or dynamic images. Can also illuminate the surroundings. Commonly used for user interaction.

**Speaker**

Plays sounds or speech. Can be used for feedback to the user or in combinations with other sensors such as microphones.

**Thermometer**

Usually placed in or near the battery to prevent overheating. Measures the battery/cell phone temperature which might be higher than the ambient temperature.

**Vibration**

Feedback mechanism, induces vibration in the device. Can be used as feedback or combined with other sensors such as the accelerometer.

**Wi-Fi**

Usually used for internet access. Can detect presence and strength of different wireless networks in different frequency bands. Measuring of connection speed (upload, download, responsiveness/ping) is possible.

**Wireless Charging**

Neither sensor nor feedback, strictly speaking. Some devices can be charged wirelessly over very short distances (1-2 millimeters). There are competing non-compatible standards, such as Qi and PMA for cell phones. The device can detect the presence of a certain charging station.

More background information can be found on the following websites:

<https://developer.android.com/guide/topics/sensors/sensors_overview.html>

<https://developer.android.com/training/wearables/apps/index.html>

<http://stackoverflow.com/questions/26769864/is-there-an-nfc-api-for-the-smartwatch-3-swr50>

<https://en.wikipedia.org/wiki/List_of_iOS_devices>

<https://support.apple.com/kb/SP737?viewlocale=en_US&locale=en_US>

<http://abstract.cs.washington.edu/~shwetak/papers/SpiroSmart.CR.Final.pdf>