New Technologies in Census Data Collection
Part 1: Planning for Mobile Data Capture

Select Topics in International Censuses[1]

Released October 2016

INTRODUCTION

As mobile technology, like tablet PCs or smartphones, become more widely available, many statistical agencies are considering using mobile data capture for the 2020 round of population and housing census data collection. Adopting mobile data capture means changing from paper to electronic questionnaires and affects the entire census life cycle. This is part one of a set of technical notes, *New Technologies in Census Data Collection*, that presents an overview of what to consider when transitioning to mobile data capture. This document discusses planning considerations for mobile data capture. Part two deals with developing an electronic questionnaire.

MOBILE DATA CAPTURE VS. SCANNING OR KEYING

The decision on whether to use mobile data capture should begin in the initial census planning stages.

Computer-assisted personal interviewing (CAPI) is when enumerators use electronic questionnaires on laptops, tablet PCs, smartphones, or other handheld electronic devices to conduct face-to-face interviews.

Mobile data capture allows users to take advantage of added features that can be programmed into mobile devices. These include:

- Integrated maps and Global Positioning System (GPS).
- Computerized case management.
- Automated skip patterns.
- Simultaneous process for data collection, data capture, editing, and consistency and nonresponse checks.
- Automated coding.
- Ability to preload questionnaire responses and customize questions.

There are both advantages and disadvantages to using mobile data capture, which should be weighed carefully. Be sure to consider the practicalities of using mobile devices, such as electricity supply, Internet availability and speed, and how much time is available until census enumeration. Preparing for a census using mobile data capture requires more preparation time than for a paper-based census.

---

[1] This technical note is one in a series of "Select Topics in International Censuses" exploring matters of interest to the international statistical community. The U.S. Census Bureau helps countries improve their national statistical systems by engaging in capacity building to enhance statistical competencies in sustainable ways.

Advantages of Using Mobile Data Capture

- Improves data quality through:
  - º Logical validation rules in the questionnaire. Errors and inconsistencies may be resolved at interview time, closer to the source.
  - º More accurate progression through the questionnaire due to automated skip patterns and mandatory items.
  - º Reduced manual data entry errors.
  - º Reduced number of unlinked forms. Reduces the risk of one section of the form getting separated from the other.
- Improves field management and real time monitoring of enumeration activities.
- Reduces time and costs by eliminating a separate process for data capture, automated coding, and reducing the volume of post-enumeration data checking.
- Decreases time between data collection and release of the results.
- Reduces costs of data capture, questionnaire printing, storage, and transportation. However, the increased equipment costs may outweigh the savings (see section on "Disadvantages of using mobile data capture").
- Can utilize GPS features and digital mapping capabilities.

Disadvantages of Using Mobile Data Capture

- High equipment costs with limited long-term use.
  - º A census generally employs a large number of enumerators. Purchasing a mobile device for each enumerator is costly.
  - º The equipment may become obsolete quickly.
  - º Devices may get lost, stolen, or break during the census enumeration.
- More time needed during the preparation stage.
  - º Need time for programming the application, setting up systems for data transfer and storage.
  - º Requires extensive testing.
- Needs more skilled programmers who are able to do sophisticated programming specific to the mobile device (e.g., Android) and all the necessary components of an electronic questionnaire application.
- Requires technologically skilled enumerators with more training and field support.
  - º Requires that enumerators can use a mobile device.
  - º Requires intensive training of enumerators on use of device. Training may be more complicated than with paper questionnaires.

- º Requires more technical support in the field to replace or repair devices and troubleshooting of hardware or software application problems.
- Segmentation effect may offset improvements in data quality.
  - º Enumerators only see a portion of the questionnaire at a time, so they may miss the bigger picture of the questionnaire. This may be alleviated through design features that allow enumerators to see the overall questionnaire structure.
- Technological problems could interfere with enumeration.
  - º The battery could run out during enumeration.
  - º The equipment could break leading to data loss or lost enumeration days.
- There are more infrastructure constraints to consider.
  - º Electricity needs to be available to charge the devices; challenging in areas with limited electricity.
  - º Limited Internet availability may cause difficulties in data transmission and other functions of the mobile data capture software application that rely on the Internet. A realistic assessment of Internet connectivity in all areas should be conducted during the planning stage.
  - º Needs a system to transfer the data from the handheld devices in a timely and secure manner. This could be the Internet, local area network, or through manual data transfer.
  - º Data security must be assured during data collection, transfer, and storage.
  - º Needs a system for backing up the data to prevent data loss.

Internet-Based Data Capture

Many national statistical agencies also are exploring using Internet-based data capture. The main advantage of Internet-based data capture is that it can be cost effective since enumerators are not used. It can also be used for residents who are hard to reach in person. However, in order for it to be cost effective, there needs to be a large number of people with reliable access to the Internet. Another disadvantage is that a highly sophisticated control system is required to avoid duplications and undercounting and to ensure the security of the data. Since the focus of this brief is on CAPI, we will not go into detail on Internet-based data capture. However, some questionnaire design considerations for CAPI are also applicable to Internet-based data capture.

PLANNING CONSIDERATIONS FOR MOBILE DATA CAPTURE

In order to ensure the success of the census, it is critical to identify all requirements for carrying out the census using mobile data capture and develop plans early in the census life cycle.

Census Timetable

The census timetable should be adjusted to fit the needs of implementing mobile data capture. Although the programming of the questionnaire cannot happen until the questionnaire specifications have been developed, it is a good idea to plan for the programming in the early stages of questionnaire development. With mobile data capture, previously separate processes may be integrated, or may need to be carried out earlier in the census life cycle. For example, data collection, capture, and editing can be done simultaneously in mobile data capture. However, generally, more time is needed to develop and test the application, set up the data transfer and processing systems, and procure, program, and test the mobile devices. If edit checks are to be added to the application, the edit programming must be completed before the enumeration begins, rather than programming those separately in the post-enumeration stage. Further, more time should be allotted for training the enumerators since the training must include the use of the mobile devices. Therefore, it is critical to determine all the steps needed to take place in order to set up a mobile data capture system and allow sufficient time prior to the enumeration in the census timetable.

Budget Considerations

Using an electronic questionnaire may save costs in the printing of the questionnaire and data capture, but the costs of the electronic equipment may be higher than the savings. Carefully consider all costs of conducting mobile data capture and account for them in the budget.

Infrastructure Considerations

Infrastructure issues such as availability of electricity and Internet access can affect the success of mobile data capture. Early in the planning stage, identify places that lack electricity and/or Internet access. If relying on a cellular or Wi-Fi data connection to transmit data, it is important to research the speed of the data transmission and plan accordingly. Mapping the areas without electricity and/or Internet access is helpful in planning. Plans should be developed for charging and backing up the devices. It is critical to have contingency plans for when electricity and/or Internet access are not available.

Differences in Questionnaire Development Process

As with a paper questionnaire, developing an electronic questionnaire is an iterative process. It must be developed, tested, revised, and then tested again, repeating the cycle until the questionnaire works as intended (see Figure 1). With an electronic questionnaire, the technical aspects of the electronic questionnaire application must be tested and revised, in addition to content.

Further, after the subject matter specialists finalize the questionnaire content, specifications are needed to serve as a blueprint for programmers to design the application. Adequate documentation of the questionnaire instrument is also required. Not only does the lack of paper documentation make it difficult to discuss the questionnaire with various stakeholders, it also makes testing more time-consuming and error prone since skip patterns are less obvious. In addition, an electronic questionnaire may contain additional features not included in a paper questionnaire, such as data validation and error messages. These features also must be specified so that programmers can design the application as intended.

When designing an electronic questionnaire, the subject matter specialists may lose control over the wording, layout, and design of the instrument. Therefore, it is critical that the subject matter specialists work closely with the programmers to make sure that there is clear communication regarding the questionnaire content, layout and design, data validation, and other specifications. It is also important to have a good understanding about the timeline, ongoing changes to the questionnaire content, data security, and quality assurance.
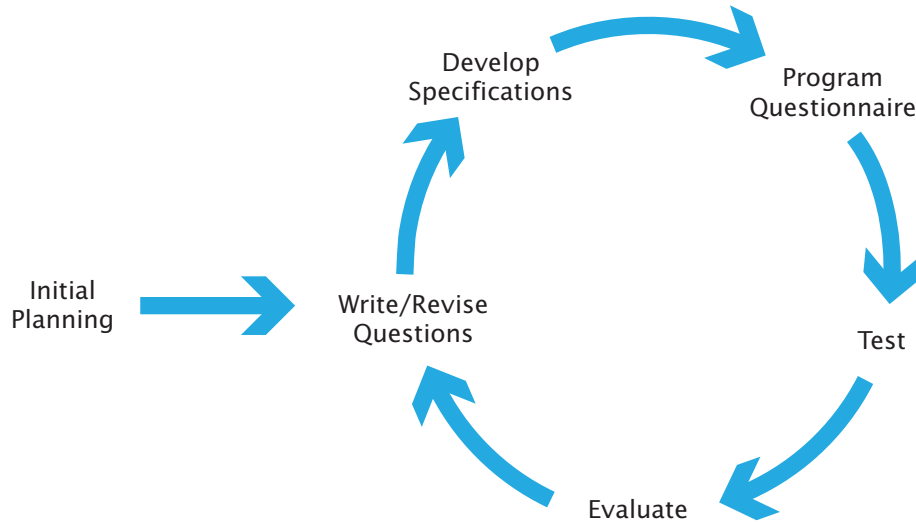
These issues are discussed in more detail in *New Technologies in Census Data Collection Part 2: Developing an Electronic Questionnaire.*

CHOICE OF MOBILE DATA CAPTURE APPLICATION

There are many existing software applications for mobile data capture. One such program, CSPro, is free software developed by the U.S. Census Bureau. It has been used in over 160 countries for censuses and surveys.

Figure 1.
Steps in Electronic Questionnaire Development

Initial Planning → Write/Revise Questions → Develop Specifications → Program Questionnaire → Test → Evaluate → Write/Revise Questions

Some questions to ask when selecting software are:

- How easily can someone modify the software to meet the questionnaire's specific needs?

- How easily can the enumerators use the software?

- How much does the software cost?

- What platform/operating system does the software require? What are the cost implications?

- What are the hardware requirements for the software? What are the cost implications?

- In what format does the software output the data?

- How well does it work with the existing data processing system?

- Is there technical assistance or training available? What is the cost?

DATA SECURITY, TRANSMISSION, AND STORAGE CONSIDERATIONS

An important consideration is how to transmit data securely from mobile devices to a central server. Secure data storage is also critical at every level where data are stored (for example on mobile devices during the enumeration, on supervisors' devices, and on regional and central data servers.) This topic should be discussed with the programmers and IT specialists.

Most mobile devices can store data either locally or on a network. Storing data locally risks data loss if the device is lost or stops working. Storing data on a network requires a way to transfer the data securely from each device to the main server. Some ways to transfer data securely include:

- Secure Internet connection.

- Local area network (LAN), a computer network that connects computers and mobile devices within a limited area (such as an office building).

- Peer-to-peer (P2P) network (for example, tablet PC to tablet PC connection through Bluetooth or Wi-Fi).

- Cable connection (for example, connecting an enumerator's tablet PC to her supervisor's laptop).

- Wide area network (WAN), or a private network that extends beyond the geographic reach of LAN, but is not connected to the Internet (for example, a dedicated mobile network).

It is important to consider what is feasible in all areas of the country when setting up a system. Reliable Internet may be available in cities, but not in rural areas. It may be necessary to establish more than one system according to infrastructure constraints. For example, data may be uploaded to the server through secure Internet connection in cities, but in rural areas, data may be transferred using cable connection. Be sure to allocate time and resources to test the data transfer and storage system. Failure to do so may have costly consequences.

IN-HOUSE PROGRAMMING VS. OUTSIDE CONTRACTOR

One issue to consider in adopting mobile data capture is whether to do the programming in-house or to use an outside contractor. One reason to hire an outside contractor is if the agency does not have the technical skills or expertise to program the software application. Another reason would be if an outside contractor already has developed a software application that can be adapted for the census. On the other hand, an advantage of using in-house programing is that skills and devices used during the census would be

available in the post-censal period. Required programming skills depend on the selected software application.

CHANGES IN STAFF SKILL REQUIREMENTS

Adopting new technologies often requires a new set of skills, which the existing staff may not have. Extensive training for existing staff may be needed or new staff may have to be hired. For enumerators, additional training will be needed to familiarize them with the technical aspects of the data collection process, such as how to operate a tablet PC, transmit data, and navigate through the application. Field supervisors may have fewer quality check responsibilities, since completeness and consistency checks may be done by the data collection software. However, they may have increased technical responsibilities, such as troubleshooting hardware/software problems and data transmission tasks. They may also be required to conduct field case management and reporting using a computerized system. Help desks at headquarters and field technical support staff may be needed to handle technical queries and troubleshoot problems during data collection. In addition, existing programmers may not have the skills to program the questionnaire application and set up extensive data systems, requiring further training or hiring new staff.

WORKING RELATIONSHIPS BETWEEN ORGANIZATIONAL UNITS

In many organizations, computer programmers, subject matter specialists, and field operation staff are often in separate organizational units. Using an electronic questionnaire requires close coordination between organizational units. How different organizational units work together should be considered when planning for mobile data capture. Depending on the design of the case management system, human resources and payroll units may also have to be involved.

COMPUTERIZED FIELD CASE MANAGEMENT

One feature of using an electronic questionnaire is that it can be linked to a computerized field case management system. As data from each device are sent to a central database, the data can be used to monitor the progress of the enumeration and identify which households the interviewers need to visit. Quality metrics and audits also can be incorporated in the field case management system. Field operations and programmers should work together to ensure that the case management system meets the needs of the field operations unit and be able to take full advantage of a computerized case management system.

REFERENCES

Bethlehem, Jelke, and Silvia Biffignandi, *Handbook of Web Surveys*, John Wiley & Sons, Hoboken, NJ, 2012.

Brancato, G., S. Macchia, M. Murgia, M. Signore, G. Simeoni, K. Blanke, and J. Hoffmeyer-Zlotnik, "Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System," European Statistical System, 2006, <www.istat.it/en/files/2013/12/Handbook_questionnaire_development_2006.pdf>.

Dibbs, Ruth, and Hale, Alison, "Questionnaire Design in a Paperless Society," SLID Research Paper Series No. 93-08, Statistics Canada, Ottawa, 1993, <http://publications.gc.ca/collections/Collection/Statcan/75F0002MIE/75F0002MIE1993008.pdf>.

Dillman, Don A., *Mail and Internet Surveys: The Tailored Design Method—2007 Update with New Internet, Visual, and Mixed-Mode Guide*, John Wiley & Sons, Hoboken, NJ, 2011.

Groves, Robert M., Floyd J. Fowler Jr, Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau, *Survey Methodology*, Second ed., John Wiley & Sons, Hoboken, NJ, 2009.

Hansen, Sue Ellen, and Mick P. Couper, "Usability Testing to Evaluate Computer-Assisted Instruments," in Stanley Presser, et al. (eds.), *Methods for Testing and Evaluating Survey Questionnaires*, John Wiley & Sons, Hoboken, NJ, 2004, pp. 337–360.

Kinsey, Susan H., and D. M. Jewell, "A Systematic Approach to Instrument Development in CAI," in Mick P. Couper, et al. (eds.), *Computer Assisted Survey Information Collection*, John Wiley & Sons, New York, 1998.

Tarnai, John, and Danna L. Moore. "Methods for Testing and Evaluating Computer-Assisted Questionnaires," in Stanley Presser, et al. (eds.), *Methods for Testing and Evaluating Survey Questionnaires*, John Wiley & Sons, Hoboken, NJ, 2004, pp. 319–335.

United Nations Department of Economic and Social Affairs Statistics Division, *Census Data Capture Methodology: Technical Report*, United Nations, New York, 2009, <http://unstats.un.org/unsd/demographic/sources/census/2010_phc/docs/CensusDataCaptureMethodology.pdf>.