Component of Statistics Canada Catalogue no. 11-522-X
Statistics Canada's International Symposium Series: Proceedings

Article

Symposium 2008:
Data Collection: Challenges, Achievements and New Directions

Experiences with Multimode Surveys

by Mark Pierzchala

2009

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2008
Data Collection: Challenges, Achievements, and New Directions

Experiences with Multimode Surveys

Mark Pierzchala[1]

Abstract

Surveys that employ simultaneous web, CATI, and paper modes (or any two-way subset) are increasingly common. Mathematica Policy Research, Inc (MPR) has deployed several of these types of surveys in Blaise. This paper reviews MPR's experiences and issues with these efforts by addressing instrumentation, survey management, and other considerations. The paper emphasizes the electronic implementation of these surveys and covers topics that emerge solely from the surveys' multimode nature; that is, material that goes beyond the implementation of a single-mode survey.

Key Words: Multimode instrumentation, Survey specification, Survey management, Blaise

1. Introduction

1.1 Description

The electronic implementation of a multimode survey must provide capabilities that go beyond that of a single-mode survey and recognize the intrinsic differences between modes. For example, each mode has its optimal question presentation, which sometimes differs between modes. In addition, the respondent may start a survey in one mode and then complete it in another mode. Again, a questionnaire completed in one mode (e.g., paper) may have gaps that must be retrieved in another mode, such as Computer-Assisted Telephone Interviewing (CATI). Any survey management strategy should balance the cost of the mode versus the willingness of a respondent to complete it in that mode. Finally, data handling and operational details such as deploying versions of the instrument are important.

In several surveys, MPR has combined all modes into one electronic instrument that manifests itself in different ways. MPR has used such an approach for paper/CATI/web surveys as well as for web/CATI surveys in Blaise. In this paper, we call an approach with all modes implemented in one instrument a *combined implementation*. The advantages and disadvantages of combined implementation are detailed below. It should be noted however, that all additional multimode capabilities described in this paper are required and must be provided one way or another regardless if the implementation is a combined implementation or not.

An important consideration in the planning and implementation of a multimode survey is its specification for instrumentation, survey management, and data flow. While esteemed survey researchers (e.g., Dillman 2008; De Leeuw 2005) have described a holistic approach to multimode specification, survey researchers who typically optimize implementation of a survey in one mode at a time often overlook the importance of a holistic perspective. A challenge to combined implementation would be a specification process that does not consider all modes together.

At the same time some differences between modes are legitimate. A good example is with data edit implementation. Edits may be fully implemented in CATI in Blaise for Windows and partially implemented in web mode while data entry from paper typically does not require the implementation of edits at the time of data capture.

[1]Mark Pierzchala, senior fellow, Mathematica Policy Research, 600 Maryland Avenue SW, Suite 550, Washington, DC 20024, USA (MPierzchala@Mathematica-MPR.com)

1.2 Multimode combined implementation

Figure 1.2-1 depicts images of a Blaise instrument for paper, web, and CATI modes. The web mode is displayed in an Internet browser while the CATI mode and data entry from paper are implemented in Blaise for Windows. The remainder of this paper focuses on the three-way paper/web/CATI combination.

Figure 1.2-1
A questionnaire presentation in three modes

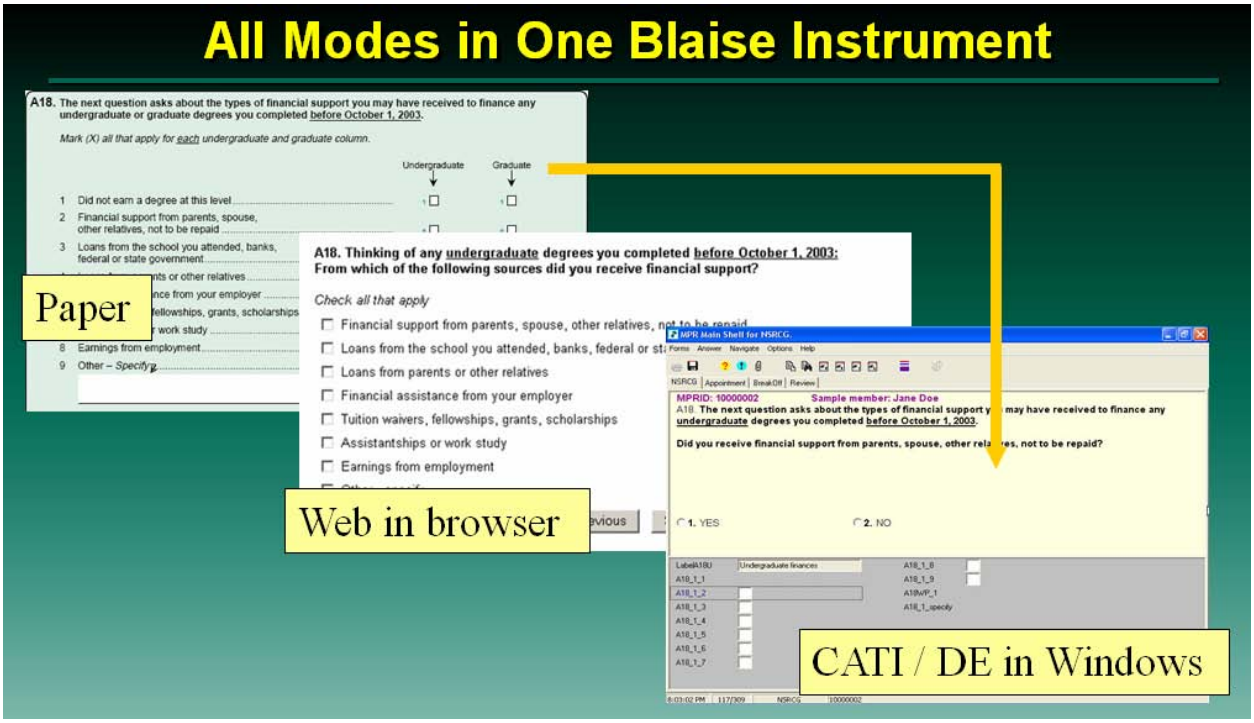


Figure 2.1-1 shows the relationships between multimode survey modules. Some of the modules are specific to a mode; for example, the scheduler is specific to CATI. The figure also explicitly shows one electronic instrument with three presentations, all of which feed into one Blaise database at the same time. All modes operate simultaneously, although MPR emphasizes modes in stages in order to reduce costs. For example, MPR's survey management strategy offers the least expensive modes first--paper and web--and then uses CATI to follow up with non-respondents or with respondents whose paper or web submissions are incomplete (break-offs) or have missing data.

MPR relies on its Sample Management System (SMS) for much of the overall study management. The SMS provides mailing and emailing services as well as paper check-in, locating (tracing), and reports. It communicates with Blaise through an overnight process that exchanges status codes and other information such as newly found telephone numbers from the locating module (SMS to Blaise) or the results of attempts on telephone numbers (Blaise to SMS).

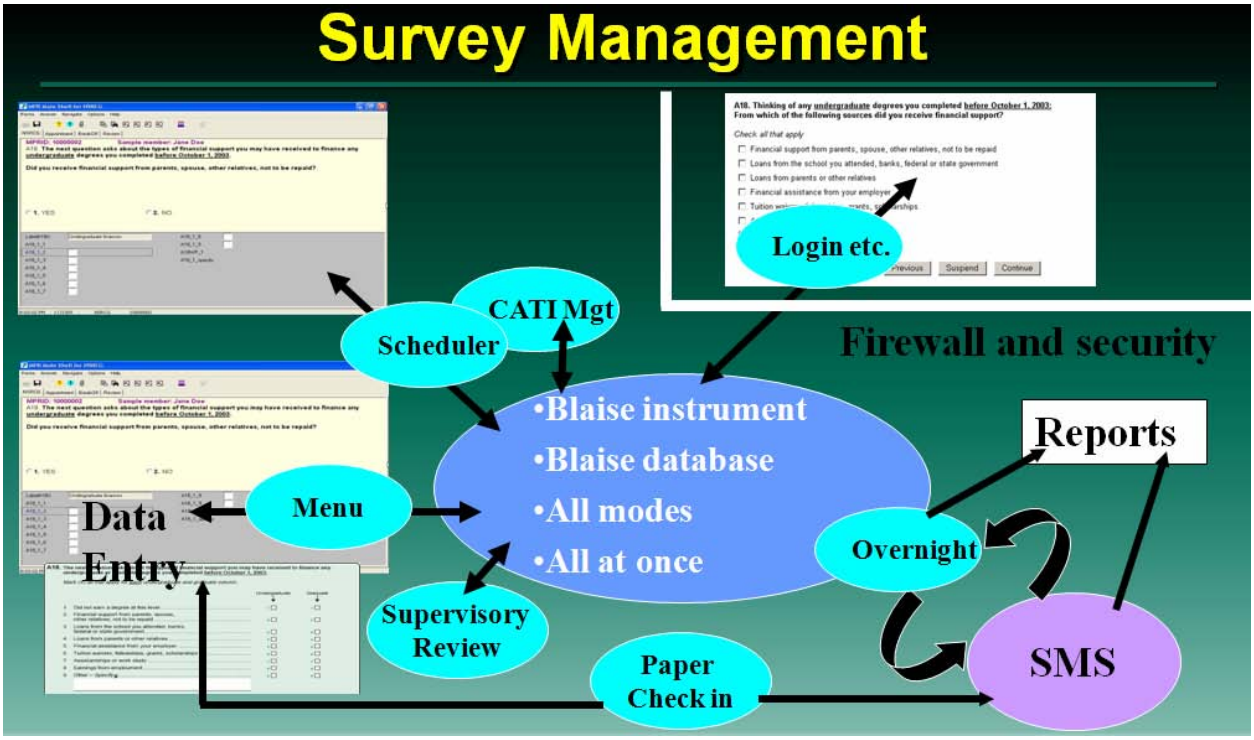
Several reports provide ongoing information on various aspects of the multimode survey. For example, one report presents a complete synopsis of all status codes across all modes. SAS programs easily generate the report from survey management data maintained in the Blaise database and elsewhere.

2. Instrumentation

2.1 Specification and programming

One of the justifications for including all modes into one electronic instrument is to realize savings from the *common aspects* of the different modes. For example, the Blaise field name (variable name), its data definition (valid values), and its place in the flow are usually identical regardless of mode. The goal is to program these commonalities once rather than two or three times in different systems. At the same time, *common mode differences* are inherent to many MPR multimode surveys and must be handled as elegantly as possible. Differences include (1) question wording; (2) allowed field attributes such as EMPTY, Don't Know (DK), and Refusal (RF); (3) edits; and (4) display.

Figure 2.1-1
Survey management relationships for three modes of a survey using combined implementation



Another, more troubling class of differences between modes is *disparate mode differences*, which are profound differences between modes such as different question formats. An example is a "code-all-that-apply" question on paper rendered as a series of Yes/No questions in CATI. The issue of disparate modes is beyond the scope of this paper but is discussed in Pierzchala (2006) and in Pierzchala et al. (2004). A goal of multimode specification should be to reduce or eliminate disparate mode differences, although it is not always possible to do so.

The greater the disparity among modes, the greater is the difficulty to render all modes into one electronic instrument. MPR's experience regarding disparity in three modes ranges from no disparity to up to half the questionnaire (Pierzchala et al. 2004).

2.2 Multimode specification

Specification writers often experience difficulty in thinking about all three modes at one time. Part of the problem is coming up with a specification process and document format that works successfully across all three modes.

Typically, several types of actors are involved in a specification: (1) the client; (2) specification writers; (3) programmers; (4) operations staff, including trainers; (5) testers; and (6) data users.

A single specification format is not likely to meet the needs of all users. On one hand, a simultaneous three-way specification should make it easier for a programmer to implement all modes at one time and for an analyst to use the data; it should also ensure that the modes remain synchronized. The client, however, may have other goals, and the testers and trainers certainly do not want a three-way specification. Even though user needs vary with regard to specification documentation, adherence to standards can expedite the specification of three modes at one time.

2.2.1 Specifying common aspects

A three-way specification handles field name, data definition, and flow as they would be handled for a single-mode CATI survey. MPR uses two complementary approaches to specify flow; first is to state the conditions under which a question is asked. Second is to state Go-To instructions. The first approach, which is essential for the data user, is analogous to Blaise IF conditions in the RULES; the second approach is particularly useful to a client and testers. These two methods of specifying flow are identically equivalent except in the case of an error.

2.2.2 Question text differences

Different Blaise languages handle text differences between web and CATI (one for each mode) while, for data entry, question text is not a concern. However, if another spoken language is involved (e.g., English/Spanish in the United States or English/French in Canada), then at least some questions may require a four-way specification. Avoidance of much multiple-question statement is possible by carefully crafting questions that work well in both web and CATI. The specification document can appear in all these language combinations, appearing one after another with proper annotation. Conditional text and fills also need to be considered.

2.2.3 Allowed field attributes

In CATI, MPR requires an answer to each question even if the answer is DK or RF. On the web, we usually do not display DK or RF but rather allow respondents to continue without answering. Possible exceptions might be critical items or other important questions. In data entry, a DK or RF may be entered if the respondent has indicated such an answer. Blanks are often allowed for data entry items; when a blank is not allowed, a missing value must be entered.

In the specification document, MPR handles these question attribute differences by stating the conventions once at the top of the document. In the Blaise source code for a multimode survey, all fields allow EMPTY, DK, and RF. In CATI, an edit is used for each field to disallow an EMPTY (blank).

2.2.4 Edits

MPR Blaise multimode instruments use three types of edits and two types of edit severity. The edits are (1) data definition; (2) suspicious range edits; and (3) consistency edits (two or more fields). The latter two edits may be soft or hard. The following table shows how the latter edits are typically are used between modes.

Mode	Edit Type	Severity	Note
Paper/data entry	Suspicious: No	N/A	
	Consistency: No	N/A	
Web	Suspicious: Yes	Soft	Apply strategically; Fields should be on the same browser page
	Consistency: Rare	Soft	
CATI	Suspicious: Yes	Hard or soft	Commonly used; May be between any number of fields anywhere in the instrument
	Consistency: Yes	Hard or soft	

The first type of edit, a data definition edit, derives from the field's valid values. In the following field declaration

```
FIELDS
  Age : "What is your age?" : 10..110, EMPTY, DK, RF
```

the user (web respondent, interviewer, or data entry operator) must enter a number between 10 and 110, leave it blank if allowed, or enter a DK or RF if allowed. None of these users would be able to enter a 9. Data definition is explicitly noted as a type of edit because specification writers are often unaware of how Blaise handles this field declaration. This is through a pop-up dialog that closely resembles a hard edit dialog that cannot be suppressed.

The specification of suspicious range edits and consistency edits should indicate applicable modes, where the edit should be invoked, and the severity. For example, it is possible for an edit to be hard in CATI and soft on the web. The programmer handles mode differences through the use of IF conditions around the edit statements.

2.2.5 Display

MPR specification writers tend not to focus on page-based display issues for data entry or CATI but do focus heavily on such issues for the web. Specification of display is one of the most difficult parts of the specification process. In web mode, it is important to declare which fields need to be on the same browser screen and whether a special format is required, such as a tabular appearance for a series of Yes/No questions. If a consistency edit must be applied in web mode, then fields should be on the same browser page.

In Blaise, the concept of a page differs between the Windows and browser platforms (Figure 1.2-1). In Windows, the page is an organizational concept in the bottom part of the screen (the third image in the figure). MPR does not specify page breaks for CATI or data entry though it could; these page breaks would differ from those in web mode.

The programmer easily handles page groupings between modes by using the Blaise 4.8 mode-specific LAYOUT. The Mode Library and/or Blaise Internet Tool, which are beyond the scope of this paper, handle other page-specific layout design issues. MPR also has well-stated question display conventions concerning fonts, font sizes, colors, and use of icons. The specification writers and programmers are trained how to state and interpret these conventions.

3. Survey management

3.1 Overall organization

Figure 2.1-1 shows the overall survey management scheme, especially how specific management modules are tied together through the common Blaise database and common instrument. It also demonstrates that the SMS is part of the overall management scheme. Each module works on an appropriate subset of cases as defined by the multimode status code scheme.

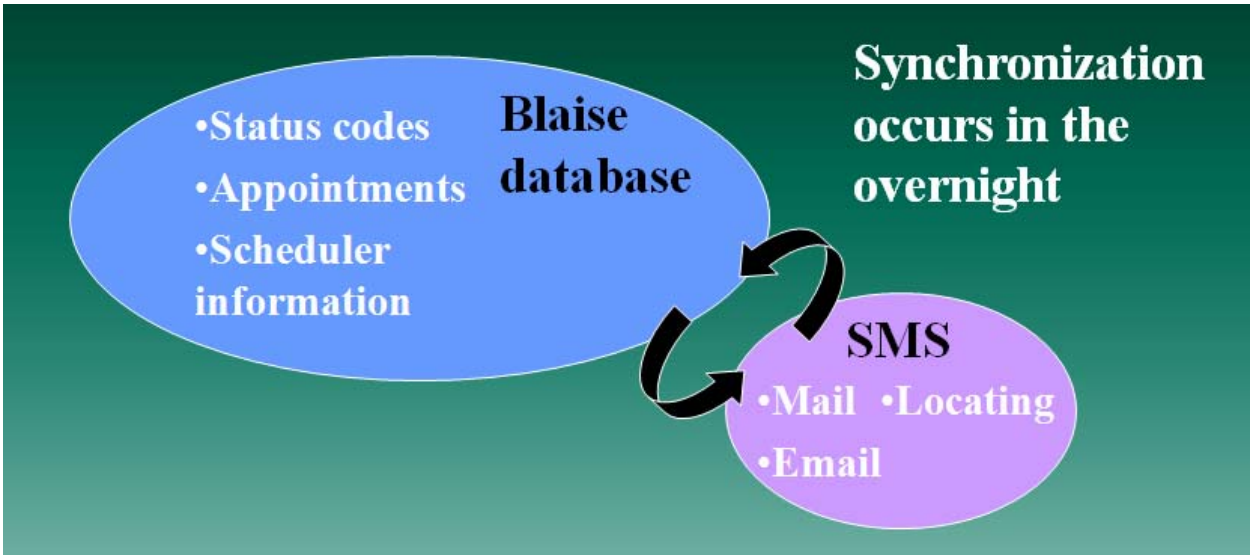
3.1.1 Reports

SAS generates status code reports from data directly retrieved from Blaise. The reports are produced after the daily Blaise/SMS reconciliation and thus reflect complete system status at one point during the day. SMS produces additional reports (not shown in the figure) for locating. Productivity and cost reports are generated from data supplied by Blaise, the SMS, and MPR's time-keeping system.

3.1.2 Location of survey management data

All data collection actions on a record are maintained in the record's case history in Blaise. The case's current disposition status in Blaise reflects its total multimode experience. Survey management data are held in two locations: in Blaise and in the SMS (Figure 3.1.2-1). The management data between the two systems are reconciled on a nightly basis through an overnight process.

Figure 3.1.2-1
Survey management data locations



3.1.3 Multimode combined implementation status code scheme

MPR has a well-developed multimode status code scheme that was established several years ago. The default set contains 102 codes broken into ranges for completes, refusals, barriers/non-completes, locating (tracing), miscellaneous non-contacts, telephone problems, miscellaneous contacts including appointments, and untouched. A consideration with combined implementation is the transition of status codes between modes. For example, if a case started in web mode and then transferred to CATI mode, the system must handle the transition correctly to reflect the new mode. A few codes are explicitly transitional, such as when a case received on paper must be re-contacted to fill in missing data. Such cases receive an explicit transition status code so that MPR can see their number in a report and better manage the re-contacting process.

Status codes may be set by the Blaise instrument, the overnight process, the SMS, or supervisory review. Each code is marked by whether it can be a final status, which module can set it, and to which mode(s) it applies. If necessary, more status codes may be added to a survey. If codes are placed in the proper range, then the overall system knows how to handle them. It should be noted that this status code scheme applies to any MPR survey--whether or not it uses Blaise--and may be and has been applied to manage a survey even when the modes are in different systems. Regardless of specific system implementation, the contribution of the status code scheme to the success of survey management cannot be overstated.

3.1.4 Call scheduling

An important subset of survey management is call management. In many respects, call management plays out just as it would for a single-mode CATI survey, although combined implementation demands some additional considerations. For example, the call scheduler often needs to leave untouched some cases in the database--completes regardless of mode, web partials, receipted paper cases, and special status codes indicating a need to re-contact for data retrieval. In addition, the scheduler should have the capability of holding cases for a period of time immediately following the issuance of email invitations. It is possible to declare special CATI management groups that are unique to multimode surveys, such as a group for data retrieval specialists.

Another feature required of a multimode scheduling operation is the ability to set waiting periods for sets of cases such as those that started on the web but never came back. The re-incorporation of these cases into the schedule after, say, a two-week period should be automated. (The script used by the CATI interviewer for such a case should recognize that the respondent accessed the survey on the web.) It is MPR practice to make every attempt to

complete the interview at the time that a respondent is on the telephone. However, it is also necessary to allow the respondent to promise to complete the survey on the web or on paper and, if the respondent agrees to later completion in another mode, to set automatically another waiting period.

3.2 Survey management strategies

In a typical MPR multimode survey all three modes are available to all respondents at any time. To reduce costs however, MPR encourages web and paper modes before transitioning to CATI, although MPR makes CATI available from the start by providing respondents with the option of calling in their responses.

MPR conducts survey management experiments to determine the least expensive methods of securing a high response rate. In recent experiments, MPR did not mention a paper questionnaire (to boost web responses) and offered varying incentive levels that ranged from no incentive to higher incentives for web completions.

MPR has found that populations differ considerably in their response patterns and that those patterns can change over time for the same population. An instructive case is two MPR surveys of college seniors and of recent college graduates. On the surface, the populations are similar and likely share some characteristics such as Internet access or ability to use a computer. However, the former all had a known email address, residential address, and telephone number, whereas the latter required extensive locating because of their status as recent graduates. Even though both surveys offered all three modes, the web response (invitations sent in email invitations) for the college seniors was so high that paper and CATI saw only sparing use. For the recent college graduates, MPR had email addresses for only a subset of cases, many of which were no longer valid. The result was a lower web response rate and much more intensive locating.

4. Other considerations

4.1 Operational considerations

Other considerations worthy of note include deployment and redeployment, staff training, educating operations staff, follow-up situations, and handling off-path data on paper forms.

4.1.1 Deployment and redeployment

Even with extensive testing, it may be necessary to redeploy an instrument. It can be difficult for CATI personnel to understand why they must take a break to redress a web problem. Considerations for redeployment include the length of time required and the optimal time of day. For example, MPR's data entry staff start work far earlier than do the CATI staff. The commonly used CATI redeployment time of 8:00 a.m. would take the data entry staff offline.

4.1.2 Training staff

Blaise programmers may not be used to single-mode programming. The issue is not necessarily with programming difficulty; the challenge lies in making a cognitive shift from a single- to a multimode paradigm so that the instrument will operate in all modes. For example, IF conditions for flow and edits may need the addition of a phrase that mentions mode.

Testers are always concerned with one mode at a time, even in a multimode project. In fact, the staff responsible for testing CATI may not be the same staff responsible for the web or data entry mode. When requesting changes testers, or a testing coordinator, must do so with all modes in mind. For example, if the question text is inadequate in CATI, it could require modification for the web (or not). A specification writer fully versed in all mode requirements should pass judgment on such situations. It is inefficient if the programmer makes adjustments in

CATI and must then make the same adjustments in the web mode when someone finally notices a problem. Worse is the situation where a change is made in one mode and not in another.

4.1.3 Educating operations staff

When planning a calling strategy, an email broadcast, or any other initiative, **operations** staff must maintain their focus on the multimode survey approach. It may be efficient to hold some cases from CATI, for example, when those cases with emails are sent an email invitation.

4.1.4 Follow-up situation

Follow-up situations include filling in missing critical items (from paper cases) or following up web partials. In the first case, missing critical items need the retrieval of one or two points of information. The problem lies in finding the places in the survey records where the information has been reported. MPR has the ability to execute a within-place/within-record retrieval in CATI that requires some (not a great deal of) special programming as well as training for a few data collection staff. In addition, the introductory scripts should be adapted for use with the respondent. If the collection of a critical item results in a new path, then any critical item on the new path must also be asked. The second case (web partials) is more like a CATI call-back to finish an incomplete telephone call, but with two major differences. First is the CATI-specific front end that manages telephone contact attempts. Second is an appropriate script that acknowledges the respondent's effort on the web. Once the front end and screener are cleared, the interviewer presses the *End* key to get to the point where the respondent left off on the web.

Off-path data entry occurs when the respondent has entered data on an inappropriate path. If the survey policy calls for the entry of all data whether or not the path is valid, then a special effort is needed to provide override values to allow the entry of off-path data. These are provided at questions that control skip patterns.

5. Conclusion

5.1 Concluding remarks

MPR has implemented several multimode surveys (paper/web/CATI) all in one Blaise instrument, with both advantages and disadvantages. To ensure that multimode surveys are a feasible business strategy, MPR must carefully plan for them and train staff in multimode methods. In addition, the underlying infrastructure should be able to accommodate multimode surveys and thus reduce the level of special effort. An advantage of a combined implementation is that it makes you confront, and explicitly handle, differences between modes.

The author believes that the advantages of multimode surveys outweigh their disadvantages. It should be possible to reduce programming costs, ease survey management, and reduce differences between modes. It is essential, however, to develop a highly trained study staff in combination with a survey infrastructure that can accommodate these requirements.

References

De Leeuw, E.D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys, *Journal of Official Statistics*, 21, 233-255.

Dillman, D. A., Smyth, J.D., and Christian, L.M. (2008). *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method.* 3rd edition. New York: John Wiley and Sons.

Pierzchala, M. (2006). Disparate Modes and Their Effect on Instrument Design, *Proceedings of the 10th International Blaise Users Conference 2006.* Arnhem, The Netherlands: Statistics Netherlands, http://www.blaiseusers.org/2006/Papers/207.pdf.

Pierzchala, M., Wright, D., Wilson, C., and Guerino, P. (2004). Instrument Design for a Blaise Multimode Web, CATI, and Paper Survey. *Proceedings of the 2004 International Blaise User's Conference*. Gatineau, Canada: Statistics Canada, http://www.blaiseusers.org/2004/papers/24.pdf.