Multi-mode/Multi-site Collection

Canadians seem to have less time and interest in completing surveys. Increased efforts are required to collect data and maintain the quality of information being gathered by STC collection personnel. We recognize the need to offer more modes of collection to our respondents in order to maintain high response rates and we also wish to better respond to the evolving needs of our clients. STC also needs to have more flexibility in case of capacity problems or regional natural emergency. For example, there was a hurricane in Halifax during the LFS collection period in September 2004. We believe certain surveys may also have better response rates in a certain mode as opposed to another. For example, I believe the Youth in Transition Survey with, in addition to CATI, an Internet mode would attain a higher response rate. Reverse Record Check of the 2006 Census was to be strictly CATI. Because of difficulty attaining the required response rate of 100%, we added a paper questionnaire and CAPI to our collection efforts. A final response rate of 88% was attained in Sherbrooke.

In the Fall of 2007, Statistics Canada asked one of Canada's leading public opinion and market research firms to determine, among other things, the public's willingness to participate in future STC surveys and mode preferences. An RDD sample of 33,000 across the country was used to collect the information. Of those willing to participate in a survey and all age groups combined, 54% said they would participate by Internet, 25% by mail, 15% by telephone, 5% by in-home interview and 1% replied no preference.

When asked about future Census completion mode preference, the responses by age groups were as follows:

For the 18 to 29 age group (a notoriously difficult age group to contact): 85% said they would respond by Internet. 78% of the 30 to 44 year old age group said they would respond by Internet, 62% of the 45 to 59 age group and 30% of the 60 and over age group said they would respond by Internet. It is fair to assume that the same would apply to those willing to participate in a STC survey.

In the near future (September 2009 for CCHS collection), Statistics Canada is planning multi-mode and multi-site collection of Health, Labour and Household surveys. At the beginning, we are mainly looking at transfers of cases from CATI to CAPI (or vice versa) and from one CATI collection site to another, since many of our surveys still don't have on-line questionnaires. The challenge was to be able to move these cases between formats (Caseman for CAPI and GIRO (Blaise) for CATI) without losing any data and without affecting the performance of these two applications. For security reasons, there are two networks in use at Statistics Canada. GIRO functions on network A (internal only) and Caseman functions on network B (external only), thus the complexity of bridging these two applications.

Statistics Canada is developing the Survey Master Control System (SMCS). This is a master control module allowing for the management, monitoring and dynamic allocation of work to existing collection systems (CATI, CAPI, EDR (Electronic Data Reporting, Internet mode)). The SMCS will connect the existing systems (Caseman, GIRO and EDR) and allow cases to be transferred between them and between sites, in an automated fashion, without loss of case information. The SMCS limits itself to the dispatching of cases between modes of collection and various collection sites. For the first phase of development, it attempts to stay out of the collection process itself, leaving legacy systems (existing collection applications) to function as they currently do with minimal change. These collection systems need only to keep the SMCS informed of the status of the cases that have been assigned to them. The SMCS also needs to receive information on any cases that have been created during the collection process (LFS, SLID).

The SMCS will significantly improve the operational effectiveness of collection.

The SMCS will also generate reports allowing the managers to monitor collection with real time information across modes, sites and surveys and is built to allow for the addition of new modes of collection and new collection software. These reports will inform the managers of the status of all surveys in collection in a particular mode/location and the status of a particular survey across modes/locations. These reports will be used to support decisions about the allocation and reallocation of survey work in response to evolving operational situations. Optimal use of collection resources is not possible without accurate data on the current state of the collection effort. We are also aiming to standardize question modules across agricultural, business and social surveys.

Para data will be standardized and stored centrally. It will remain available for analysis.

With the SMCS, the managers of surveys such as LFS (Labour Force Survey) and CCHS (Canadian Community Health Survey) will be able to move cases between CATI and CAPI without IT personnel needing to make a copy of the cases, as it is now required. We will simply use an outcome code to indicate which cases need to be transferred to another mode of collection and the SMCS will transfer the case, along with any data from a partially completed interview. The transfer is done automatically without action being needed from IT personnel. We will also be able to move CATI cases from one collection site to another in the event of capacity problems, emergency situations such as power outages or to deal with language cases.

The SMCS will solve the technical problems of moving sample. Collection and survey managers in conjunction with the clients are responsible for deciding when a case or whole or part sample will be transferred to another location or to a different collection mode. Close collaboration between all managers and clients is necessary for success.

The following questions concerning the operational aspects of collection are being discussed. There will be many more to come as we start working with the SMCS in a multi-mode, multi-site setting.

How do we allocate initial budgets and move money around if the sample or part of the sample has moved?

It's been suggested that H.O. could allocate the budget based on what is sent out initially, and hold back a certain amount to be distributed based on any reallocations that occur. This is my preferred option. Regional offices must have some financial information to schedule the interviewers and plan collection.

Another option might be to hold the money centrally in H.O. Head Office would then pay out to the Regional offices at the end of the survey based on the hours worked; this would require close monitoring to ensure time is being used productively and would also require that hours be charged to the correct project code.

Effective utilization of collection staff. If more staff is available in one site or in one mode, it will be possible to unburden the over burdened site or mode. To consider: CAPI collection is much costlier than CATI collection. How do we manage these extra costs?

SMCS will allow us to make these changes in an automated fashion. However, the decision making is still a negotiated process that has to take place between the client, the regions and head office. SMCS will allow everyone to see where the capacity exists. There are different costs associated with different modes and who assumes these extra costs has to be clear during the decision making process.

Once the decision is made, sample can be easily moved around, but it will have to be decided as to whether some CATI cases will be finalized early as non responses (those cases with very little chances of success), freeing up money and making the change cost neutral, or if more money will be added to the budget. If there are a significant number of non-contact cases in CATI, it could be decided to close these cases and to prioritize those with the greater chance of success by working on these in another mode (CAPI).

What kind of performance indicators will we have? Response rates may not be very meaningful at a site or mode level if sample is moving between sites or modes.

We will need (and plan to have) more detailed information in order to relate hours worked to progress; For example, if we can measure when cases have been contacted (and rostered), and whether we have a partially completed questionnaire, then we can produce new indicators that will allow us to distinguish between time spent on calls where some progress is being made vs. calls where nothing new is being collected. We could then produce some "state of completeness" indicators. Reports indicating how many cases have reached these various states would be available, regardless of mode/site. This level of detailed information would have to be added to the BTH (Blaise Transaction History). Once added, the information would need to be available in a report form.

How do we evaluate interviewer performance?

In a full SMCS world, we would be better able to measure interviewer performance across surveys since all the data would be available from a single source. We might also want to expand our definition of interviewer performance to include not just the number of completed cases they do, but also cases where progress is made. This will be possible even though the case might be completed by another interviewer on another day.

Quality control. Our usual monitoring system remains for CATI. If a CAPI case becomes a CATI case after a transfer, does our regular monitoring system work?

The plan is that once a case is transferred to CATI, it would look exactly like any other CATI case and would be eligible to be monitored.

In CAPI, the POINT (Pace of Interview) system will be used.

Both these systems monitor interviewer performance only, not quality control of the data.

For EDR, what is our quality control system?

The internet application would have edits built in. Since it is a self-completed application, the edits (and the question wording) may be different from what would appear in a CATI or CAPI application. If the case is later transferred to CATI (either for failed edit follow-up or just to complete a partial) then the case would be subject to whatever edits were built into the CATI application and could be monitored as any other CATI case could.

What happens to the time slices if we move the case to a different time zone? Does a case coming from another mode arrive in CATI as a new case therefore not subject to time slices?

The time slices would be based on respondent time, so even if the case is moved to a different office (time zone), it would be handled by the call scheduler in an appropriate manner. If a case is transferred from another mode into CATI, it would take on all the requirements of the call scheduler once it arrives in the CATI environment. It would be subject to time slices but it would be starting as a fresh case (i.e. no calls have been made yet). Appointments would be respected (again, based on respondent time).

Caps on calls (attempts) in CATI still apply but do the efforts deployed in the previous collection mode apply? Example: If there were 4 tries in EDR, do these count in the 15 tries in CATI?

This needs to be discussed more widely. I believe that the caps on calls (attempts) should only apply within a mode. At most, it could be cumulative across CATI and CAPI since in both cases an interviewer is trying to contact the respondent.

In response to changing survey conditions, managers want to be able to change collection sites to improve response rates or better control collection costs.
Multi-site collection would allow collection capacity in the various offices to be managed more effectively. Time zones and language requirements must be considered.

This is part of the argument for having more sites rather than consolidating into one large CATI site. Multiple sites protect against changes in labour force conditions in one part of the country and respond to the need for conducting surveys in many different languages where interviewers may not be available if only one site was used.

SMCS is a tool that will give managers more flexibility. However, all changes done during collection still need to be negotiated.

A CATI no contact case, after how many attempts do we wish to transfer the case to CAPI if indeed that is what we want to do?

It is expected that this will be a survey-specific decision. SMCS will allow surveys to have "business rules", and this could be defined there. So, a survey might say that after x number of attempts (or x days of collection) in mode A, the case should be transferred to mode B. The rules might be more elaborate -- for example, it could be specified that a valid telephone number (or address) must be present or that only cases in urban areas should be transferred. The surveys will have to decide on what is best and most cost-effective for them. And this must be discussed with all players before the business rules are finalized.

An initial collection plan of a specific survey will need to indicate:

- The modes of collection to be employed
- The location where a particular mode of collection will be employed

- Any default conditions to be used in a particular mode of collection (e.g. number and periodicity of reminder e-mails for an EDR collection, cap on calls settings for a CATI collection).
- The rules for allocating cases among the modes/location of collection
- The rules for transferring uncompleted cases from one mode of collection to another.
- The rules for resolving ambiguity (e.g. for a single case, responses are received from two different modes of collection.

It will be possible to modify the collection plan during the collection period. Cases will be able to be shifted from one location to another, the addition of modes of collection will be possible as will be the resetting of default conditions

Surveys with e-Questionnaires. If no response, after how long are the cases transferred to a collection site and treated as a CATI case?

This would be a survey-specific decision as to when (or if) cases get transferred to another mode, and how many reminders get sent.

For Internet surveys, if we do not have an Internet address for the respondent, a letter is sent with the website address and an access code. Right now, we don't have a lot of experience with Internet surveys but as we gain experience we would develop guidelines and best practices concerning things like the number of reminders to send and how long to wait before trying CATI follow-up.

Capacity issues for large sample movement. What is the minimum time needed to call in interviewers if any are available? From experience, 24 to 48 hours are required. If staff is not available, the transfer of cases is not a viable option. If staff is available, large staffing changes require at least 48 hours. The number of computers available must also be considered since just about all social surveys have the same best time to call.

Large sample movements would have to be discussed in advance of any actual movement, so in most cases there would be time to plan and schedule additional staff.

If there was an actual emergency that required moving sample, then we would prioritize what surveys to work on. Interviewers already scheduled to work on one survey will be asked to work on the cases of the newly arrived survey. Example: a regional office being shut down during LFS week and the other offices being asked to take on their caseload to get the LFS done. Other surveys in the receiving offices might not get worked on. This kind of situation would be unusual.

Very close collaboration between CATI and CAPI managers and between managers in different sites is necessary for this to function.

One thing that has been proposed is for one person in a region to act as the overall regional collection manager for a survey; a data collection manager in each site would still need to manage the interviewers to make sure people are scheduled and showing up to work. But the single manager would be responsible for actually managing the collection efforts since they would have access to progress information for all sites. This needs to be discussed further. Many other management options exist.

How much of a say does the client have when the decision is made to move a case or part of sample to another site or mode collection?

Small-scale movements should be defined in the business rules for the survey, so no additional discussions would be needed. Those moves would happen automatically. For large scale moves that are not covered by the business rules, there should be a discussion between the client, regional managers and head office to ensure that everyone knows what is going on and the implications of any decisions. If the movement is between sites within CATI, the part of the organization responsible for collection has the ultimate decision.

When do we offer the respondent another collection mode? Do we offer or do the managers decide when to move a case from one mode to another?

This will have to be a survey-specific decision. If other modes are available, and the respondent offers to respond in another mode, then it is expected that we will do the transfer. However, we want to avoid interviewers offering to transfer a case to another mode to avoid working on the difficult case! Generally, the mode transfers will be done according to a set of business rules.

Some modes (like internet and maybe paper/mail) could be available at the same time as CATI or CAPI, so the respondent could then have the choice of doing the survey over the phone when called, in person with a field interviewer or completing it on the internet or to fill out a paper questionnaire. Whichever one is done first, the case will be automatically disabled in the other mode.

If we transfer a case to another CATI site, which office becomes responsible for the case? Who tracks its progress?

Whichever office a case is assigned to, it becomes responsible for the case until it gets transferred out. That office would track its progress while it is assigned there. However, we would also want someone to be responsible for monitoring collection as a whole (across all modes). That could be someone at Head Office or more likely someone in one of the regions who would assume overall responsibility for the survey. This is something that has been suggested but it would need to be discussed with all involved.

How does the SMCS track the case? Do we get this information in a report by putting in the case number and the survey code?

SMCS will need a set of unique identifiers - usually survey ID, cycle ID, and a sample ID. These ID's would be used across all modes. A report will also inform us of the route taken by the case during the whole collection period. At the end of a survey, it will be interesting to see how the sample has been dispatched (evolved across modes and sites) so this information can be used when the survey is repeated.

There will be a learning curve for managers when the SMCS comes into operation and multi-mode/multi-site collection becomes a reality. Just as time slices required adjustments to managerial practices, multi-mode/multi-site collection will too. The introduction of time slices is now in the past and the benefits are real. The LFS in Sherbrooke with time slices in use is approximately $5,000 a month less to collect than before. This is a $60,000 saving a year in one location.

Multi-mode/Multi-site collection may not lower the costs of surveys to such an extent but should facilitate the reaching of our required response rates therefore lowering, somewhat, our costs.

The main goal is to answer respondent and client needs and to maintain our high response rates.