

CONFERENCE OF EUROPEAN STATISTICIANS

**Workshop on Statistical Data Collection**

10-12 October 2017, Ottawa, Canada

WP.4-1

22 August 2017

---

## **Active Management Framework to Monitor and Manage Data Collection**

François Laflamme, Dominique Chabot-Hallé, Sylvie Bonhomme and Dennis Galante (Statistics Canada)

*francois.laflamme@canada.ca*

### ***Abstract***

Survey data collection managers need effective and practical tools to assess and actively manage data collection processes and performance. The decreasing response rates in household surveys, the increasing complexity of collection strategies, the growing need for timely and factual information, and a renewed collection vision at Statistics Canada have clearly established the need to develop better adapted tools to actively monitor and manage data collection throughout the collection period. This paper describes the essence of the Active Management (AM) framework, a solution developed and implemented at Statistics Canada to monitor and manage data collection for all types of household and agricultural surveys during collection periods.

# **Active Management Framework for Monitoring and Managing Data Collection**

François Laflamme, Dominique Chabot-Hallé, Sylvie Bonhomme and Dennis Galante  
Collection Planning and Research Division, Statistics Canada

## **1. Introduction**

In the data collection context, **Active Management (AM)** consists in a set of plans and tools used to monitor and manage survey data collection while collection is still in progress. AM activities generally include planning, monitoring, performing timely analyses, identifying problems, implementing corrective actions, communicating, and evaluating.

## **2. Reasons behind the Active Management (AM) initiative**

Several drivers led to the development and implementation of the AM initiative at Statistics Canada. These include decreasing response rates, the frequent need for timely and factual information on collection progress, the increasing complexity of survey data collection strategies, the lack of common collection management processes, and changes to the organization's collection vision.

### **2.1 Decreasing response rates**

Like many statistical organizations, Statistics Canada has observed a downward trend in household survey response rates. Fortunately, recent experience with many surveys has clearly demonstrated the effectiveness of a strong, well-established AM program in alleviating downward pressure on response rates.

### **2.2 Data collection progress information requirements**

While regular monitoring reports generally meet the needs at aggregated levels, survey data collection managers typically do not receive enough detailed real-time, factual and evidence-based information about in-progress cases, or about certain key features of the data collection process. In many instances, regular data collection reports no longer meet the increasing and varying requirements of survey managers and diverse stakeholders.

### **2.3 Complexity of surveys**

The growing complexity of survey data collection strategies and processes creates additional pressures (Laflamme et al. 2016). For example, the implementation of Responsive Collection Design (RCD) for all CATI surveys, the introduction of web and multi-mode surveys, and the increasing complexity and diversity of data collection procedures have made survey monitoring more and more difficult over time. Consequently, it is important to monitor survey progress constantly, using the most appropriate tools available.

### **2.4 Lack of common collection management processes**

The lack of common collection management processes, including coordination between regional offices (ROs), is another reason behind the AM initiative. A process for sharing best practices and collection strategies would allow each RO to take advantage of lessons learned by other ROs. For example, production plans, which specify the planned work intensity and the results expected from that work intensity, often vary greatly from RO to RO, and this variation usually affects response rates. Coordination of production plans between ROs will become even more important under the Integrated Collection and Operation System (ICOS).

## **2.5 Changes to the data collection vision**

ICOS has provided a common solution to meet the collection needs of all surveys, including the most recent (2016) census. According to the ICOS vision, ROs will no longer be responsible for their own samples. Rather, there will be a national sample, available to any interviewer at any RO in Canada. From a data collection management perspective, this is a very important change. It will require new indicators and metrics for monitoring and managing collection efforts and productivity against results, by RO.

## **3. Data collection management**

The challenge for managers with regard to data collection management is to make the appropriate trade-offs between quality, timeliness and cost in order to meet clients' needs. Other dimensions have an impact on survey management information needs and outputs as well.

In practice, data collection management information requirements are driven by three main factors: the management levels involved, the point in time in the data collection process, and the subject being managed (e.g., production units, collection site and interviewer performance).

**Management levels involved.** The organizational structure comprises various levels of management, supervision and oversight, each of which has its own requirements in relation to the depth and breadth of the information needed. Higher management levels need a wider view that reflects the various active surveys within the manager's responsibility whereas, at lower levels, the focus narrows in breadth but increases in depth.

**Point in time in the data collection process.** Data needs vary according to the point in time in the data collection process. This pertains not only to the frequency with which data are refreshed and outputs are provided, but also to the nature of the data elements of interest.

**Subject being managed.** Survey managers will need to focus on interviewers, regional offices or production units, depending on the circumstances, and the management information system (MIS) will need to encompass the relevant details of these aspects (Laflamme et al. 2008). For regional offices, the challenge is to identify the right indicators and metrics for monitoring and actively managing collection efforts and productivity against expected results. For example, results could mean the number of completed interviews by RO, because it will not be possible to calculate response rates by RO in ICOS surveys.

## **4. Active Management (AM)**

Active Management (AM) can be defined as a set of plans and tools for monitoring and managing data collection while collection is still in progress. The purpose of an AM program is to provide timely, factual, topical and relevant data on survey performance and progress throughout collection, so that problems with collection are identified early and decisions on how to correct problems are based on cost, effort and data quality, as well as the response rates attained so far.

AM has four main objectives. The first objective is to determine whether, at any given point during collection, the observed key indicators are aligned with the key assumptions and milestones identified in the planning phase. For example, does the relationship between the observed and expected response rates, and between the proportion of the budget observed as having been spent and the proportion expected to be spent, align with the production plan—that is to say, the survey

progress planning assumptions about work intensity and response rates—at various times during collection? The second objective is to be proactive in identifying problems through timely analysis; and the third objective is to correct these as early as possible, before collection is finished. The fourth objective, which is more global in scope, is to make effective use of collection resources in order to strike the most appropriate balance between data quality, timeliness and survey costs.

In an effective collection strategy, AM is a key element of the decision-making process because, very often, timely changes need to be implemented during the data collection period further to current empirical observations.

The key elements of AM include planning, monitoring progress, timely analysis (identifying problems and implementing corrective action), communication, and evaluation.

#### **4.1 Planning**

Active Management (AM) often requires planning during application design and development. An effective AM plan must address several elements.

The plan must include the key survey planning assumptions that will be monitored throughout the collection period. Those key assumptions are generally common to all types of surveys, subject to a few exceptions. The most important assumptions that need to be part of the planning phase include the expected response rate (and/or the number of completed interviews), the expected hit rate, the time per responding unit, the average interview time, the budget figures (e.g., hours and/or money) and the production plan (i.e., survey progress assumptions about work intensity and expected response rates at different points in time during collection). For surveys where the risk of not achieving the target response rate is high, AM should be involved as early as possible in the decision-making process, to ensure that the proposed collection strategies, and the key planning assumptions, are realistic and achievable.

In addition, the plan must assess the level of needs with respect to MIS-based reporting, survey progress monitoring, and management support, with respect to the level of risk that the survey will not achieve the target response rate. In practice, the AM plan should identify the most appropriate progress indicators for the survey, as well as the types of analysis to be performed and the frequency of such analysis.

The plan should determine the types of actions or interventions that will be carried out in cases where key observed indicators and measures diverge from key planning assumptions (e.g., observed versus target response rates) or where other problems arise.

Lastly, the plan should include a communication plan describing the governance and communication protocols to be followed in managing collection and operations activities.

AM planning can account for the particularities of a given survey. For example, strategies to deal with hard-to-reach cases, to follow up on non-response cases, to move cases from one collection mode to another, or to monitor specific aspects can be addressed.

In summary, planning involves all aspects of data collection at all stages required to meet a survey's objectives.

#### **4.2 Survey monitoring**

The progress of surveys is generally monitored through reports. The reporting plan for a survey should identify the type of information needed as well as any variables, indicators or metrics required for reporting at various levels of aggregation. It should also ensure that this information is available in a timely manner during the data collection period. The plan should identify the most important indicators for monitoring progress and, consequently, the survey-specific report(s) that may be required for a given survey.

There are four types of Active Management (AM) reports:

- standard reports, suitable for the vast majority of surveys
- survey-type reports, applicable to a particular type of survey (e.g., electronic questionnaire (EQ) transition report or flow metrics for initial and current / final collection modes for web and multi-mode surveys)
- survey-specific reports that respond to a particular need or address the unique nature of a given survey (e.g., response rate by domain of interest).
- ad hoc reports generally created once, in response to a specific issue (e.g., list of promising cases close to the end of collection).

Several variables, indicators, measures and metrics are included in these reports with a view to monitoring progress. Examples include

- response rate
- EQ take-up rate
- effort (e.g., attempts such as calls or visits)
- time spent
- refusals
- conversion effort
- performance (e.g., productivity, efficiency)
- cost (e.g., proportion of budget spent) at various levels of aggregation, such as collection site, date, type of case (respondent, non-respondent, out-of-scope), time slice
- attempt outcomes
- domains of interest
- strata.

In some situations, reports to monitor collection effort and productivity at the interviewer level, and reports to assess progress and status at the sample-unit-level close during collection, are required. In practice, an AM plan should identify the series of relevant reports that needs to be distributed at different points in time during collection, with a view not to overburden survey managers.

#### **4.3 Timely analysis**

Given the increasing complexity of survey design and survey collection strategies, it is essential to add value to survey monitoring reports by providing a timely, factual and evidence-based analysis with respect to the collection progress and by including performance highlights from the collection period. With that consideration in mind, an AM plan should specify the type and frequency of analysis that will need to be performed. For example, short surveys with short collection periods (e.g., less than 14 days) need to be monitored daily, while surveys with long collection periods often require only weekly monitoring.

The first objective of this analysis is to determine whether the observed key indicators are aligned with the key planning assumptions throughout the collection period, such that the expected results—i.e., either the expected response rate or the number of completed interviews in specific cases—are likely to be achieved. For example, analysts often compare the expected response rate against the proportion of the budget spent at a given point in the collection period, in order to detect issues that need flagging. Many comparisons can be made across multiple observed and expected values for these key indicators. However, it should be noted that many interactions can occur between these key indicators; for this reason, they need to be taken into account during the analysis. These include operational constraints, which generally involve external information coming directly from the field data collection managers.

The second objective of this analysis is to identify data collection problems as soon as possible during collection. The idea is to find problems when, or even before, they occur, not when collection has ended. This is another instance where a timely analysis plan that quickly identifies the nature of a problem is needed. The range and sources of possible problems are numerous. Problems can be related to inaccurate planning assumptions, an unrealistic production plan, unexpected difficulties relating to the frame, ineffective allocation of data collection effort, interactions with concurrent surveys, etc.

For regular and ongoing surveys, there is also an opportunity to compare the results of the current survey cycle with those of previous cycles. This often makes it easier to understand the nature of a problem. The monitoring reports are designed to identify possible problems and pinpoint the source of such problems: they do not presume to give all the answers relating to an issue. From that perspective, AM also provides mechanisms for handling unanticipated or emerging problems—for example, preparing ad hoc reports that investigate situations as they develop.

The third objective of the analysis is to take the most appropriate corrective action. The goal at this stage is to determine whether action is required to correct a problem and to decide what steps to take should action be required. The decision-making strategy and its implementation should have been determined at the initial planning stage. However, since the ability to take corrective action may depend on the application's design and on operational constraints that might reduce the possible range of actions (e.g., time lag before changes can be made to computer-assisted personal interviewing (CAPI) surveys), it is important to include contingency planning in the overall collection management plan.

Currently, the AM team in the Collection Planning and Research Division performs a weekly analysis of surveys that pose a high or medium risk of not achieving the target response rate. The AM team conducts this first analysis using all relevant existing reports, then summarizes the highlights in a survey monitoring template. The team meets once a week to raise and discuss any problems that might require immediate attention or further assessment. Sometimes, an ad hoc report is prepared on the matter. Results, observations and recommendations are then communicated to, and discussed with, field data collection managers so that the most appropriate corrective action can be taken should action be warranted.

This overall AM exercise requires experienced and well-trained staff, not only at every step of data collection, but also in all other aspects that can affect the achievement of survey objectives—for example, sample design. Since its full implementation, AM has clearly demonstrated its effectiveness in alleviating the downward pressures on response rates. In fact, response rates for many household surveys have increased since the AM initiative was put in place.

#### **4.4 Communication**

Many experiences, some recent and others less so, have highlighted the importance of developing and implementing an effective communication plan. AM is a multidisciplinary team effort resting on an extensive collaborative approach between collection managers at head office and those in regional offices. Therefore, an ongoing and well-thought-out communication plan must be implemented and maintained in order to identify collection issues and achieve agreement on any needed changes to collection strategies in a timely manner. The decision-making strategy should have been determined, and its implementation should have been planned, at the initial planning stage. The communication strategy should clearly identify who needs to be contacted in the event of a problem, who has the authority to make decisions, and what the process for taking corrective action should be. The strategy should also identify who is responsible for monitoring different aspects of the collection process.

#### **4.5 Evaluation**

The collection process should be evaluated to establish best practices and share experiences so that problems are not repeated with other surveys. The evaluation should document the types of problems identified, the decisions made as to how to correct these problems, and the impact of the corrective actions on the data collection indicators. The results of the evaluation should be included in the survey post-mortem so that regional offices can share collection best practices with each other.

#### **Conclusion**

Given the increasing complexity of survey design, data collection managers need better timely and factual tools to manage and assess effort and performance against expected result collection while data collection is still in progress. This need has never been more acute. Active Management is a key element of a data collection framework: it integrates a wide variety of components and processes of differing levels of importance into different management levels at different points in time in the collection cycle. The effectiveness of this framework depends not on any one process, but on the collective effect of many interdependent and often complex measures, indicators or metrics. As with other aspects of the survey progress, the efficiency of data collection depends on Statistics Canada's ability to adapt and evolve.

Accordingly, the agency needs to continue supporting this type of analysis and to encourage innovation so that a professional operational infrastructure can be maintained. To be true data collection experts, staff should not only develop technical knowledge; they should also be aware of data collection issues, and be able to monitor and analyze the data collection process.

#### **Acknowledgements**

The authors wish to thank the staff of the Collection Planning and Research Division and the Collection and Regional Services Branch of Statistics Canada—especially Sylvain Hamel, Greg Moreau, Joshua Finn, Laura Simpson and Said Akif—for their contributions to, and comments on, this research.

#### **References**

Laflamme, F., S. Hamel, and D. Chabot-Hallé. 2016. *Statistics Canada International Symposium Series: Proceedings – Statistics Canada's Experiences in Using Paradata to Manage Responsive Collection Design for Computer-Assisted Telephone Interview Household*. Proceedings of the 2016 Statistics Canada

Symposium "Growth in Statistical Information: Challenges and Benefits." Statistics Canada Catalogue no. 11-522-X.

Laflamme, F., M. Maydan, and A. Miller. 2008. "Using Paradata to Actively Manage Data Collection." *2008 Proceedings of the Section on Survey Research Methods*. American Statistical Association.