Modernisation of Statistics Production (MSP2009)
Conference in Stockholm 2–4 November 2009

Data Collection Research using Paradata at Statistics Canada

François Laflamme[1]

Over the past few years, Statistics Canada has conducted several analytical studies using paradata to learn more about various issues surrounding the data collection process and practices. In particular, these investigations have attempted to better understand how data collection progresses through its cycle, to identify strategic opportunities, to evaluate new collection initiatives and to improve the way the agency conducts and manages its surveys. The main objectives of this paper are to present the main results of these past and ongoing investigations describing Statistics Canada's experiences with regards to paradata. Future research plans that focus on identifying viable operational strategies that could improve efficiency or data quality are also discussed.

Key words: Paradata; call transaction history; responsive design; productivity.

1. Introduction

The challenge of any statistical organization is to collect high quality data in a cost effective manner despite many influencing factors such as decreasing response rates, evolving population behaviour and increasing respondent burden. Data collection is definitively a key element of the survey process because it has a direct impact on quality and it is a major component in the cost of many statistical products. Over the past few years, Statistics Canada has been reviewing its data collection processes to evaluate and monitor current practices and new data collection initiatives, as well as to identify strategic opportunities for improvements. Paradata has been the cornerstone of these investigations and will continue to be extensively used in future research.

Much of the research discussed in this paper relies on the Blaise call transaction history of many social and agricultural CATI surveys. The main objectives of this paper are to provide the highlights of this research including some major achievements and to present ongoing and future data collection research plans and priorities. The paper begins with an overview of the data collection and paradata context at Statistics Canada followed by a description of the main objectives of this research. Section 4 presents the highlights of past paradata research that have resulted in an increased understanding of many issues surrounding the data collection process and identified potential areas of improvement that might require further investigations. Section 5 describes ongoing research on survey productivity and costs while section 6 discusses the assessment of new data collection initiatives. Section 7 presents an overview of some major achievements resulting from paradata research which include the initial development of a responsive design strategy. The last section describes future research plans that focus on identifying viable operational strategies which could improve data collection efficiency or data quality.

[1] François Laflamme, Statistics Canada, Business Survey Methods Division, 6th floor Jean Talon Bldg., Ottawa, Ontario, K1A0T6, francois.laflamme@statcan.gc.ca

2. Data Collection and Paradata at Statistics Canada

This section provides an overview of the data collection and paradata organizational context at Statistics Canada including a brief description of data sources and contents of the Statistics Canada paradata warehouse.

2.1. Organization of data collection

Data collection for CATI social, agriculture and business monthly surveys is conducted and managed in Statistics Canada's six Regional Office (RO) call centres located across the country. The mix of surveys varies by site, and on average, each RO CATI centre conducts between 5 and 10 surveys in a given month. In 2007, about 1,000 CATI interviewers made, on average, 700,000 calls per month.

2.2. Paradata context

Data Integration and Production Planning (DIPP), which is Statistics Canada's paradata warehouse, was built to integrate and standardize operational data from three different sources for Management Information System (MIS) reporting and analytical purposes: 1) Blaise Transaction History (BTH) files for CATI surveys, 2) Case Management (Caseman) files for CAPI surveys, and 3) Survey Operations Pay System (SOPS) information for all surveys including non-standard surveys. DIPP includes paradata from most Statistics Canada surveys conducted since 2003 and is updating to include ongoing active surveys on a daily basis. In practice, this information becomes available the day after paradata information is collected or recorded.

A call transaction history record essentially refers to production information for both CATI (BTH files) and CAPI (Caseman files) surveys. A record is created each time a case is opened, either for data collection or other purposes. It contains detailed information about each call or visit made to contact each sampled unit during the data collection process. It includes information on the survey and case identification, the date, the amount of time the case was open, the interviewer who worked on it, the result of the call plus additional relevant information about each call or visit. On the other hand, the Survey Operations Pay System (SOPS) contains administrative and financial information on interviewer pay claims for all collection activities. A SOPS record is generated each time an interviewer enters a claim for a particular survey and task on a given day, either for direct data collection activities (e.g. interviewing, tracing, etc.) or for other purposes (e.g. supervision, specific training, etc.). Each claim includes: interviewer identification, type of interviewer (i.e. regular, senior), survey name, date, task code (e.g. interview, training, tracing, etc.), number of hours claimed and fees for CATI/CAPI surveys and additional information such as expenses and kilometers for CAPI surveys. At the end of 2008, DIPP contained about 60 million transaction records and 9 million administrative records from about 722 survey cycles (e.g. a monthly survey represents 12 survey cycles) which represent 80 unique surveys (19 agricultural, 9 business and 52 social surveys) since 2003.

In addition to the information available through DIPP, sample design information and external information collected by interviewers or recorded for each interviewer can also be considered as important paradata sources. For example, the sample design information of each survey is very often used to enhance the analytical value of these studies or for new collection initiatives such as Active Management. On the other hand, interviewer characteristics (e.g. interviewer experience) and behaviour collected through CATI/CAPI monitoring (or an audit trail) as well as interviewers' subjective information about neighbourhood characteristics and potential survey

cooperation from the sampled cases (e.g. assessment of the likelihood that a case would complete the interview after a first refusal) are also part of paradata. Most this information was not used in these studies, with the exception of interviewer experience.

3. Research objectives

Research that uses paradata can be conducted before, during and after data collection to understand, assess, monitor and improve the data collection process. The goals of all these studies include one or more of the following specific objectives:

- to learn more about the data collection survey process within and across surveys;
- to identify operational efficiency opportunities;
- to evaluate the data collection process including new initiatives and emerging issues;
- to provide timely feedback and customized information for active survey management;
- to maintain and improve data quality; and
- to improve the way data collection is conducted and managed.

All these studies are based on empirical paradata automatically collected throughout the data collection period; not on anecdotal or point-in-time observations. In addition, most of these studies were conducted for many surveys to compare the results across different types of surveys (e.g. longitudinal, cross-sectional and RDD) as well as to validate and generalize research conclusions.

4. Lessons learned on data collection process

This section presents highlights of past paradata research that have significantly increased the understanding of many issues surrounding the data collection process and practices within and between surveys. The vast majority of these studies have also identified potential operational improvement opportunities, some of which have been implemented while others require further analysis. The priority for future research will be given to strategic opportunities that could be operationally viable and lead to cost efficiency or quality improvements. As mentioned, most of the research presented relies on social and agricultural CATI surveys.

4.1 Calls versus time spent

In order to better understand and evaluate data collection, it is necessary to know how data collection time and effort is distributed throughout the collection process. Historically, the focus has been on the reduction of the number of calls to get a first contact. In general, however, a large proportion of collection time is spent after the first contact trying to get cooperation and to complete an interview or to confirm a non-response. It would be worthwhile to pay more attention to the calls after the first contact to investigate the current collection strategy and practice.

4.2 Contact rates

For the purpose of these analyses, a contact is defined as any call for which an interviewer is able to talk to someone in the household even though it is not the selected person. On average, it takes

between two and four calls to make an initial contact with a household. As many other studies have shown, weekday evenings and weekends have been systematically found to have the best contact rates for CATI social surveys while early morning, lunch time and early evening are more appropriate for agricultural surveys. These conclusions were corroborated by the General Social Survey (GSS) - Time Use Survey which provides information to determine the likelihood that an individual will be at home at a specific time based on socio-economic information. This general information could be used at the initial phase of data collection in conjunction with the time slice initiative (see section 6.2) to improve contact rates during the first few calls especially when some socio-economic information of the individual is known prior to the start of data collection as for longitudinal surveys.

4.3 Sequence of calls

While it is important to make the first call at an appropriate time, the majority of calls are subsequent calls. For these calls, the specific history of calls for each sampled case should also be taken into account because the sequence of calls provides specific and targeted information about each sampled case during data collection and could be used in many useful ways.

Initially, the sequence of calls could be used to increase the likelihood of contacting a particular individual within the next few calls. For example, the contact rate after the first two calls depends on the time in which each call was made. For social CATI surveys, the contact rate remains relatively high when at least one of the calls is made in the evening period. The same conclusion applies for longer sequences of calls. As the number of calls increases, the sequence of calls becomes a more important piece of information than the general population information because it provides specific information about pattern for each sampled case. In addition, the sequence of calls could be used in conjunction with the outcome of the previous calls. Previous studies have shown that when contact was made on the first call, the timing of the second call is not as critical as when no contact was made. However, it is beneficial for no-contact cases to be re-attempted in the evening. As opposed to the socio-demographic data that provides general information about the "best time" to call a given household, the sequence of calls provides specific information for each individual case about ways to improve the probability of reaching the household. Secondly, the sequence of calls and outcome codes could also be used for decision making purposes. For example, preliminary investigations for RDD surveys have indicated that cases with the first two or three consecutive calls with a "fast busy signal" are more likely to be identified and classified as out-of-scope cases at the end of the data collection period. Therefore, this information could be used to make early decisions about those cases and to save or redirect resources to other types of in-progress cases. Finally, information about the sequence of calls could be used in conjunction with the elapsed time between calls and the duration of the data collection period to ensure that call attempts are spread evenly throughout the collection period so as to maximize response rates. It is not desirable that cases reach their limit of calls (cap on calls) by the middle of data collection period since these cases have no chance to be attempted after that. In summary, the sequence of calls could be processed, analyzed and used during the data collection process to adjust and improve the collection strategy based upon specific information about each sampled case. This would be similar to what is done with score functions for business surveys. This is a practical example of adaptive data collection strategy defined in the responsive data collection conceptual framework discussed in section 7.

4.4 Contact versus interview

Depending on the type of survey, about 40% to 50% of the total number of CATI respondents is reached on the first contact call. An additional 4% to 7% of respondents are reached on the same day as the first contact attempt was made, for example, an appointment in the morning following by the interview in the evening. Reaching the second half of respondents requires a significant effort. Further investigations of the collection process and practices after the first contact with a household would be advantageous for improving efficiency.

4.5 Interaction between surveys

On any given day, regional offices conduct many CATI surveys which compete for available data collection resources. Some surveys might be at the beginning of their collection period while others are in the middle or close to the end of theirs. In addition, surveys do not always have the same priority level. For example, the Labour Force Survey (LFS), one of the most important Statistics Canada surveys, collects data from about 55,000 households over 10 days each month, requiring almost all collection resources in the first few collection days.

4.6 Staffing versus in-progress workload

Some studies were performed to better understand the relationship between the amount of interviewing effort made and the expected in-progress workload during the data collection cycle. Previous research (Laflamme, 2008a) as well as many field observations resulting from the Active Management initiative (Laflamme et al., 2008b) have suggested that the interviewer staffing level is not always well aligned with the workload sample and the expected productivity. For example, the fact that cases are likely to be called more often during a single day in the second half of the collection period suggests that at that point the interviewer staffing levels are greater than the sample workload. The proportion of in-progress cases attempted on the consecutive days, as well as a rapid decrease in terms of productivity, are also good indicators. Data collection managers, therefore, need interviewer staffing management and planning tools to reduce some of the tension on collection productivity and costs (Cooper et al., 1998). These types of projects will definitely be among the research priorities for the next few years. It should be noted that rules surrounding notice of shift changes for a unionized interviewing workforce need to be factored into any action plans resulting from this research.

4.7 Ad hoc research

In addition to these specific research projects, many ad hoc investigations were conducted on emerging or operational issues that required immediate attention as well as on issues about data collection progress and results raised by senior management. For instance, the relationship between interviewer experience and productivity and investigations into responsive sample design in a multi-mode and multi-site environment impact are only two examples. Most of these ad hoc investigations have also contributed to increase understanding of the data collection process and practices within and across surveys.

5. Ongoing research on survey productivity and cost

The main objective of this type of research is to investigate the relationship that exists between production and cost data during the collection period. The first part of the research project merges and consolidates production (BTH) and cost (payroll) information by creating a single record that contains a summary of both production and financial information by interviewer, survey and day.

This also includes the assessment of the consistency of this consolidated information between the two data sources, the results of which are not presented in this paper. The concepts associated with this summarized information are described in sub-sections 5.1 and 5.2. The following sub-sections present the highlights of this relationship as well as an overview of the initial research with respect to productivity indicators and cost analysis studies.

5.1 Production paradata

Non-Interview System Time represents the amount of time spent on a case to make contact, to try to get cooperation or to confirm a non-response or to determine it is an out-of-scope (i.e. all calls prior to the interview itself including tracing time recorded by the system). On the other hand, Interview System Time represents only the amount of time devoted to conducting interviews. These two variables add up the Total System Time that represents the time logged on the system with an open case. It does not take into account such activities as tracing time performed outside the system,; breaks and the elapsed time between two calls. These activities represent some of the difference between system time and paid time.

5.2 Cost paradata

Direct Collection Payroll Hours represents the amount of time claimed on direct data collection activities (e.g. interviews, contact calls, no contact calls, tracing, etc) which is conceptually comparable to Total System Time. Other Collection Payroll Hours represents the amount of time claimed on indirect data collection such as special training, supervision or other related collection activities or duties. Total Payroll Hours equals Direct Collection Payroll Hours plus Other Collection Payroll Hours and it provides a very good proxy of the survey cost for CATI surveys.

5.3 Relationship between survey production and cost

Figure 1 presents production and cost progress throughout the whole survey cycle for the Survey of Labour and Income Dynamics (SLID 2007) conducted from January to April 2007. All series have a similar pattern suggesting a very good relationship between production and administrative data. In particular, there is a strong correlation between Total System Time and Direct Collection Payroll Hours ($\rho > 0.95$). It should also be noted that the proportion of production time is a very good predictor of the proportion of payroll hours claimed as shown in Table 1. For MIS purposes and survey cost tracking total system time can be used as a reliable proxy for the direct survey costs. These initial findings have generated many ideas for new research, in particular, with regards to productivity and survey cost analysis as briefly described in the next two sub-sections.

Figure 5.3-1
Distribution of System Time and Payroll Hours by Collection Day
Survey of Labour and Income Dynamics (SLID 2007)

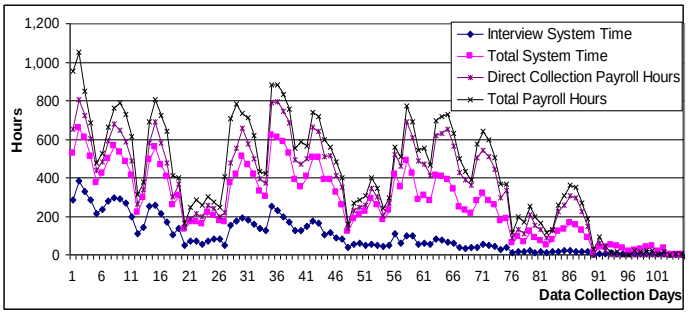

Table 5.3-1
Distribution of Cumulative System Time and Payroll Hours by 10-Days Period – SLID 2007

10-Days Period	Cumulative Response Rate	Cumulative Percent Total System Time	Cumulative Percent Direct Collection Payroll Hours	Cumulative Percent Total Payroll Hours
1-10 days	21.7%	17.9%	16.5%	16.8%
11-20 days	33.8%	30.3%	27.8%	28.4%
21-30 days	41.2%	39.3%	36.3%	37.9%
31-40 days	54.1%	55.3%	51.8%	52.9%
41-50 days	61.6%	67.0%	63.0%	63.6%
51-60 days	67.3%	78.2%	74.0%	74.1%
61-70 days	72.6%	89.2%	87.4%	86.8%
71-80 days	75.2%	95.0%	94.9%	94.5%
81-90 days	76.7%	98.5%	99.7%	99.4%
91 days plus	**77.1%**	**100.0%**	**100.0%**	**100.0%**

5.4 Survey productivity

The development of productivity indicators that could be used to manage survey performance and cost during the data collection process was identified as an important component of the overall paradata research initiative. Among these indicators, the ratio of Interview System Time (i.e. time solely devoted to the interview itself) to Total System Time (which includes all successful and unsuccessful calls) is a good indicator of the evolving survey productivity during the course of data collection because its links the interviewing effort to the results (i.e. the interview) for a given survey and period, generally a day. This productivity indicator decreases over time for all CATI surveys which is understandable since the proportion of more difficult in-progress cases increases towards the end of collection. However, it is important to note that these types of indicators depend directly on both interview length and response rate. For example, surveys with long interviews will have higher productivity ratios than surveys with shorter interviews for a given response rate. Conversely, the productivity indicator defined as the ratio of (Cumulative Non-Interview System Time / sample size) is less affected than the previous indicator by the interview length (but is still dependent on response rate). In practice, it provides a measure of the average time to obtain cooperation from respondents or to confirm a non-response or an out-of-scope for all sample cases. This type of indicator is more suitable to evaluate and compare survey productivity across surveys than the previous one. Finally, it is also noteworthy that the ratio of Total System Time over the number of completed cases (i.e. number of interviews) could also be useful to estimate the expected time required (and then the collection resources) to meet respondent targets in some surveys (Lepkowski et al., 2007). However, since this ratio generally increases over the course of the collection period, it often underestimates the amount of resources required.

5.5 Survey cost analysis

Timely and accessible detailed cost information as well as findings about the relationship between production and payroll data have recently generated survey cost analysis. In particular, a study on the impact of cap on calls on survey costs for some of the most important CATI longitudinal surveys (Survey of Labour and Income Dynamics (SLID) and Youth In Transition Survey(YITS)) and cross-sectional surveys (e.g. Canadian Community Health Survey(CCHS)) was undertaken. This study took advantage of the relationship between production and cost data because there exists no direct link between a call and its cost (e.g. paid time is not linked to any case in particular). Based on preliminary results, the maximum potential cost saving could have been, in theory, in the 3%-4% range under a cap of 40 calls when all production time for calls that exceed the cap on calls is accounted for. In practice, time spent on over-capped cases on a given survey is not automatically "saved" since interviewers continue to work on the other available cases. In other words, there are always cases left to be worked on. The impact of cap on calls on survey costs is still being investigated and the scope of research will be extended to other types of surveys.

6. Evaluating new data collection initiatives

This section provides an overview of two new data collection initiatives: cap on calls and time slices including the preliminary results on their impact on survey progress and costs. This research essentially aims to meet the third goal of the overall data collection objectives described in section 3.

6.1 Cap on calls

Before January 2006, there were no limit on the number of calls that could be made to complete a case, as either a response, non-response or out-of-scope. It was not unusual to observe more than 25 calls made before completing an interview, even more in the case of longitudinal surveys. For example, in 2006, about 11% and 14% of respondents for the Survey of Labour and Income Dynamics (SLID) and the Youth In Transition Survey (YITS) were reached after 25 calls. Starting in 2007, a policy on cap on calls was gradually implemented for all CATI surveys. The policy aims to limit the number of calls permitted per case to control respondent burden and to improve the cost benefit of the calls made. In practice, the maximum numbers of calls were set to 20 and 5, respectively for listed and unlisted telephone numbers for RDD surveys; 25 for targeted respondent surveys (except for the Canadian Community Health Survey (CCHS) whose maximum is set to 40); 40 for longitudinal surveys and to 15, 25 or no limitation for agricultural surveys depending on the priority of the case.

Although the cap on calls has now been implemented for most surveys and has eliminated a very large number of calls, its impact on response rates, survey estimates and costs (section 5.5) has raised concerns. For instance, many studies were conducted to assess the impact of caps on calls on response rates using paradata available prior to its introduction in 2006. It is estimated that 1.6%, 2.1% and 2.5% of respondents would have been lost respectively for CCHS, SLID and YITS under a cap of 40 calls. In the case of longitudinal surveys such as SLID and YITS, the cap on calls could have a negative impact on the quality of the estimates given the cumulative effect from wave to wave. Currently, the number of cases capped in most surveys is assessed throughout the Active Management initiative to monitor the impact of the cap on calls on response rates. On the other hand, its impact on survey estimates and cost is still being investigated.

6.2 Time slice

Due to the limit on calls, it is important to ensure that all calls are handled in the best possible manner. The time slice feature in the CATI call scheduler was utilized to assist in managing the new cap on calls policy. In practice, time slices ensure that a specific number of calls are attempted at different times of the day, and on different days of the week, before a case is finalized. The introduction of time slices has shown some encouraging signs of slightly reducing the average number of calls to get a first contact for RDD surveys. While about two to four calls may seem reasonable, it is important to note that between 40% to 50% of respondents are reached on the first contact and thus many others require a much higher number of calls. Finally, it should be mentioned that the assessment of the impact of time slice throughout the data collection period is not straightforward because it involves many operational factors.

7. Summary and achievements

This section provides a summary of paradata investigations in terms of the research objectives including a brief description of data collection process changes and developments resulting from these research findings.

Past, ongoing and ad hoc paradata analyses have definitively provided a better understanding of the data collection process and practices within and across CATI surveys. Changes to data collection have already been implemented to take advantage of these findings. For example, the proportion of evening shifts versus day shifts has gradually increased throughout the collection

period to improve contact rate and productivity. As well, time slices were customized for some survey types in order to use information available prior to and during collection, for example sample design and household socio-economic information collected from the roster. In addition, the impact of new data collection initiatives was assessed and continues to be monitored and this has resulted in changes in the cap on call definition for some major surveys.

The research findings have also stressed the need to develop a more flexible and efficient data collection strategy, not only to reduce collection costs, but also to make better use of the calls allowed under the new cap on calls policy. The approach should evolve during the collection period to take account of survey progress, productivity and cost. This collection strategy essentially refers to the responsive design approach as initially defined by Groves and Heeringa (2006). Recently, some work was undertaken at Statistics Canada to develop a responsive design conceptual framework (Laflamme and Mohl, 2007) which includes two main components: Active Management ((Laflamme, Maydan and Miller 2008)) and adaptive collection. The main idea is to constantly assess the progression of data collection (Active Management) using the most recent information available, and adjust data collection strategies in order to make the most efficient use of remaining available resources (adaptive collection). More specifically, a timely Active Management tool is needed to closely assess and monitor survey progress, effort and cost during the course of the data collection process to predict, identify and correct operational problems if necessary. Timely information and production indicators are also needed to determine critical data collection milestones that require more significant changes to data collection. In other words, an adaptive collection approach is required to allow for more important changes in the data collection strategy at different points in time during the collection period. The approach would take into account survey progress, the time and effort already put on the remaining in-progress cases and characteristics such as the proportion of the sample in the refusal or tracing groups or the average number of calls for the in-progress cases. Of course, the approach would also consider impacts on data quality and cost. The Active Management initiative has been implemented for most CATI social and agricultural surveys and has been developed for CAPI surveys.

8. Future Work

Among the opportunities for improvements already identified and discussed, some require further investigations to better assess their feasibility and operational advantages and benefits on the data collection process in a responsive design context. In particular, the priority of future paradata research will be given to opportunities that could be operationally viable and lead to cost-efficiency, timeliness or quality improvements. For example, a more detailed analysis of the sequence of calls (including the elapsed time between calls) to obtain a first contact as well as the time spent to achieve cooperation after a first contact is required to maximize the likelihood of making fewer calls to make contact and receive cooperation. Another example would be a study on the relationship between productivity and survey progress indicators to identify distinct data collection phases from which important data collection changes (adaptive collection) are required for responsive design purposes. Finally, there is a need to investigate tools to better plan the effective use of data collection resources (interviewers) during the collection period based on observed progress because it directly impacts survey productivity and costs. One of the most important challenges would be to phase in, integrate and consolidate this series of opportunities

into a responsive design data collection strategy to improve the cost efficiency of the data collection process in the long run.

9. Conclusion

Paradata has been the cornerstone of data collection research at Statistics Canada and continues to be extensively used. In fact, timely and easy access to an exhaustive paradata database has many advantages. Firstly, results of the research are based on objective and empirical measures about the collection process. Secondly, it allows for comparisons across different types of surveys as well as for the validation of research findings since the same investigations can be reproduced on many types of surveys or survey cycles. Thirdly, accessibility of historical data provides the opportunity to conduct trend analysis over time and to assess the impact of new initiatives. Fourthly, paradata are automatically obtained during collection representing almost no collection cost and a very low interviewer burden. The main cost of paradata analysis is associated with the creation and maintenance of the paradata database and MIS reports. There is also a human cost in analyzing and interpreting the data which includes the learning and training aspects. Even though paradata research requires some upfront investment, the benefits of this type of analysis are of significant importance given that data collection represents a very large proportion of the overall survey cost of many statistical programs.

References

Couper, M. P., Baker, R.P., Bethlehem, J., Clark, C.Z.F., Martin, J. Nicholls W.L. and O'Reilly, J.M. (1998), "Computer Assisted Survey Information Collection", Wiley series in survey methodology section, Chapter 15 by Edwards, Suresh and Weeks, p. 301-306.

Groves, R.M. and Heeringa, S.G. (2006), "Responsive design for household surveys: Tools for actively controlling survey errors and costs". Journal of the Royal Statistical Society, Series A. Volume 169, Part 3.

Laflamme, F. and Mohl , C. (2007), "Research and Responsive Design Options for Survey Data Collection at Statistics Canada". 2007 American Statistical Association, Proceedings of the Section on Survey Research Methods.

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

Laflamme, F., (2008b), "Understanding Survey Data Collection through the Analysis of Paradata at Statistics Canada". American Association for Public Opinion Research 63rd Annual Conference, 2008 American Statistical Association, Proceedings of the Section on Survey Research Methods.

Lepkowski James M. et al., (2007) "Advance s in Telephone Survey Methodology", Second International Conference on Telephone Survey Methodology, Miami 2006, Wiley series in survey methodology section, p. 363-367