Training Survey Statisticians at Statistics Canada

Royce, Don
Statistics Canada, Methodology Branch
150 Tunney's Pasture Driveway
Ottawa, Ontario K1A 0T6, Canada
Don.Royce@statcan.ca

1. Introduction

Statistics Canada is Canada's national statistical agency, publishing a wide variety of statistical information on the economic and social fabric of the nation. Among the agency's 5,000 employees are approximately 320 mathematical statisticians, known locally as methodologists, who work in the Methodology Branch. This is by far the largest single group of mathematical statisticians of any organization in Canada. These employees are responsible for the statistical design of virtually all of the agency's programs, with expertise in subjects such as questionnaire design, sample design, estimation, editing, imputation, quality evaluation, data analysis, time series analysis, and many others. As such, they are a key set of employees, with specialized knowledge and skills that are crucial to the quality of the information produced by Statistics Canada. They also represent a significant investment of over $25 million annually.

Training of this group of specialized knowledge workers is obviously a high priority for Statistics Canada. Approximately 5% of the agency's salary budget goes to formal training, with much more devoted to informal training. Training and development of Methodology Branch staff is a key preoccupation of senior management. Nevertheless the training of this group of employees presents significant challenges. In Sections 2 and 3 respectively of this paper, we describe the nature of these challenges and the solutions we have developed in response to them. Section 4 gives a brief description of the curriculum of in-house courses; for a more complete description we refer the reader to Gambino and Gough (2005). Section 5 concludes the paper with a description of how we evaluate the success of this training.

In this paper, we use the word "training" to refer to a wide variety of approaches to the development of the knowledge and skills of the employee, not just formal classroom training. In fact, within the agency the word has largely been replaced by the concept of "learning," putting the focus on the employee rather than the employer. This is consistent with the agency's philosophy for learning and development for all of its employees: the responsibility for learning and development is that of the employee, while management's responsibility is to create the proper environment. The focus of this paper is primarily on training in statistics and survey methodology, although other types of training relevant to the work of a methodologist are mentioned. The training described here is just one part of the agency's overall continuous learning strategy for all of its employees.

2. Challenges in Training Survey Statisticians

Probably the greatest challenge we face in training statisticians is the wide diversity of educational backgrounds of our staff. Statistics Canada's predominant source of methodologists is the pool of graduates in statistics from universities across the country. On average we hire 20 to 25 recruits per year through our post-secondary recruitment campaign, although there can be considerable fluctuation depending on budgets. These universities have widely different programs in statistics, and while all new methodologists must have a

degree with a minimum number of courses in mathematics and statistics, they can range from someone with a Bachelor's degree who has never taken a course in sampling to a PhD who has written his or her thesis on an advanced topic in the field. As well, the educational backgrounds of recent graduates are much different from those of staff who obtained their degrees many years ago. Thus, there is a need to bring all staff up to a certain basic minimum level of knowledge and skills in the field of survey methods, while at the same time allowing flexibility in the course curriculum to meet individual needs.

A second challenge is the increasing downward pressure on the time available for training, both for the staff who are taking training and for those who are developing and delivering training. In the case of the trainees, the average number of days of training of all types taken per year has fallen from approximately 17 days per employee in 2000 to 13 days today. Although part of this decline is accounted for by a reduction in the number of recruits hired each year (recruits typically spend more time on training), another important factor is pressure from other activities that compete with training for the fixed number of days available in a year. For example, methodologists have recently been awarded significant increases in vacation and other types of leave through collective bargaining. In order to maintain the number of days working on the statistical programs of the agency at a reasonable level, other activities, such as training, have had to be reduced. While recent measures have stabilized the situation and no further reductions are expected, it seems clear that the number of days available for training will not increase significantly in the near future.

In the cases of the trainers, most of the courses and other learning events are developed and delivered by employees of the Methodology Branch itself, doing so on a part-time basis. The formal budget available for developing and delivering courses is quite small, at approximately 1.4 Full-Time Equivalents (FTEs) per year, and does not cover all the activities of the trainers. The staff developing and delivering courses often have to fit this work in over and above their regular duties. Those who develop and deliver courses also work on other projects, whose demands may at times be more urgent. Coupled with this is the fact that the field of survey methodology is constantly developing, requiring old courses to be updated and entirely new courses to be designed. The field of methodology also covers a wide variety of topics, as diverse as questionnaire design and time series analysis, requiring a wide range of training courses.

As well as just finding time and budgets to develop training courses and other learning events, not every methodologist is a suitable trainer. The ability to develop and deliver courses is a specific skill, and while all methodologists are expected to develop good presentation and speaking skills as part of their toolkit, those who are effective trainers need to be very highly skilled in communications. They also, of course, have to be very knowledgeable about the subject they are teaching. Such people are usually very valuable for other work as well, so freeing them up to invest their time in training others can be difficult.

Such challenges would be significant enough, but Statistics Canada also makes training available in both of Canada's official languages, namely English and French. Approximately 45% of the methodology staff is francophone and 55% is anglophone. This implies that two instructors (or sometimes two sets of instructors) have to be found for each course. It also means that the same quality of training must be maintained in both languages. When this is not done, for example because of the quality of the instructor, it is not unheard of for one linguistic group to prefer taking a course in their second language. This, of course, defeats the whole purpose of developing a course in both languages. A solution adopted in about half the cases is to have a bilingual trainer deliver the course in both English and French versions.

Finally, in the last few years there has been increasing competition from types of training other than statistical training. Training in an employee's second official language has become increasingly important for those who wish to advance to any level of management, as staffing processes that used to allow an employee

to become bilingual after they were appointed to the job are being phased out in favour of staffing processes that require bilingualism as a pre-requisite. Courses in informatics, management, subject matter, communications and other topics increasingly compete for an employee's attention as these disciplines become more complex. Furthermore, as a result of increased emphasis on accountability, many government training courses are becoming mandatory. For example, all new government employees must take a two-day Orientation to the Public Service course, and newly appointed supervisors and managers must take basic courses ranging from 3 to 5 days, depending on their level. Statistics Canada itself is in the process of developing mandatory courses for all staff on topics such as confidentiality and quality assurance. While these courses have their own merits, they do reduce the amount of time available for training in statistics.

Table 1 shows the distribution of the type of training (excluding language training) taken by level of employee. As one might expect, at the entry and developmental levels courses in methodology and statistics predominate, although their percentage decreases gradually. The percentage of informatics courses is also fairly high, but decreases rapidly as the level increases. In contrast, the percentage of management courses increases rapidly as employees begin taking the courses needed to achieve their career goals and carry out their new responsibilities. At the management levels, where one manager would supervise between 10 and 50 staff (depending on the level), management courses predominate.

Table 1: Distribution (%) of training by type of course, by level of employee

Type of course	**Level of employee**			
	Entry	**Developmental**	**Autonomous**	**Management**
Methodology and statistics	44	43	30	8
Informatics	37	20	11	4
Subject matter	7	7	8	7
Communications and interpersonal relations	2	8	7	5
Management skills	3	15	32	61
Other career and personal needs	7	6	10	11
Other	1	1	2	4
Total	100	100	100	100

3. Meeting the Challenges

While the challenges in developing and maintaining an effective program of training are many, Statistics Canada has done a relatively good job in training its employees, based on the available indicators described in Section 5. In this section, we describe the ways in which the agency has met the challenge of training a diverse, bilingual workforce to meet the varied demands for statistical methods.

A key factor in this success is that these 320 employees are organized centrally within the agency, making it possible to develop training and development programs specifically for this group. The development and maintenance of these programs are overseen by the Methodology Branch Learning and Development Committee, which consists of representatives of the three divisions within the Branch. Each division also has its own Learning and Development Committee, chaired by a divisional "Learning Champion", who is part of a broader network that extends throughout the entire agency. These divisional committees are responsible, with the divisional director, for developing a divisional learning plan that identifies the learning needs of the division over the coming year. For example, it might identify that the employees of the division need to become more competent with a new piece of software. The Committee also encourages all employees, in collaboration with their supervisor, to develop a Personal Learning Plan.

This plan is based on both the divisional plan and the employee's own career development goals. Approximately 70% of employees have a Personal Learning Plan. To indoctrinate new recruits right from the start with our learning philosophy and to inform them of the resources available to them, the chair of the Branch Learning and Development Committee gives a presentation to all new recruits at a one-day welcoming session held annually.

Under the auspices of these committees and their predecessors, the Methodology Branch has developed a number of guides designed to assist employees in developing a Personal Learning Plan. These documents are available to all employees through an internal website dedicated to career development. Foremost among these is *The Career Path in the MA Group*, which describes the role of the methodologist, the knowledge and skills required at each level, how promotions are made, and a set of recommended training courses at each level. The recommended training is divided into core training and other suggested training. Another key document is the *Methodologists' Training and Development Handbook*, which goes into more detail on the learning philosophy at Statistics Canada and the specific needs and training courses recommended at each level.

As well as their supervisor, new staff can have a mentor assigned to them. This more experienced person, who is not in the line management of the employee, serves to guide the new employee in their career, showing them "how things work" and providing advice on the less tangible aspects of the job. Mentors often assist employees in defining their learning needs and can suggest ways to meet them. It is important to note that the mentor is not involved in the evaluation of the employee's performance; he or she is strictly there as an advisor. The choice of the employee to have a mentor or not is completely voluntary.

The above committees, documents and mentoring program are all aimed at career development in general. For statistical training in particular, most of which is provided by the Methodology Branch itself, there is also a Statistical Training Committee. This committee's role is to identify the needs for statistical training, to coordinate the development of new statistical courses, to find instructors, to coordinate registration and other logistics, and to collect participant feedback and follow up where necessary. The Committee maintains a curriculum of approximately 20 courses, ranging in length from one day to eight days. The courses range from very introductory to very advanced, covering diverse topics such as sampling, longitudinal surveys, record linkage, time series, quality control, data analysis, survival analysis, questionnaire design and small area estimation. This wide variety of courses allows staff to choose the courses that they need to bring themselves up to a basic level or beyond.

These courses usually require a substantial time commitment on the part of both students and teachers, however. Also, they do not cover all the topics relevant to the work of methodologists, and are not always easy to keep up to date. To address these issues, the Branch has recently developed a series of special topic seminars given by recognized experts in the Branch. These seminars, which last only an hour, are designed to give new recruits (and others) a quick introduction to the basics of the subject and to get them up to speed quickly. Topics have included sampling frames, collection methods, editing, non-response and imputation, estimation, variance estimation, quality control, dissemination and confidentiality, record linkage, time series and others. The slides from past presentations are available on the website referred to above and the student can then delve into the topics by reading some of the reference material listed by each presenter. This seminar series has been quite popular with both students and teachers, although it does not substitute for more in-depth courses. The seminars also have the advantage of letting the recruit know whom to contact as the expert in the topic. As well as these seminars on basic survey methodology topics, each division also has a bi-weekly seminar series that covers topics of current interest. These seminars are generally related to the specific survey program that the presenter is working on and are aimed at staff at all levels.

The courses and seminars described above are taken by the student according to their individual needs, as expressed in their Personal Learning Plan. However there is one course which all new professionals (not just methodologists) are expected to take at some point in their first two years on the job. The *Survey Skills Development Course (SSDC)* is a six-week course for all new professionals designed to expose them to all the steps of a survey. The first few weeks are spent in lectures, learning about populations, frames, sampling, data collection, processing and analysis. In the last two weeks, the students design and carry out an actual survey for a real client, including conducting door-to-door interviews, processing the data, and writing an analysis report. At the end of the course the student will have an excellent overview of the survey process and an appreciation for the value of teamwork. It also provides non-statisticians with an appreciation for the job of a methodologist, making future communications easier. The SSDC is in its seventeenth year and has produced over 2,000 graduates. Several other countries have adopted or been influenced by this approach.

As well as courses and seminars provided by the agency, Statistics Canada also makes use of external training resources to the extent possible. There are three universities in the region which provide courses in statistics and related topics. The Methodology Branch will reimburse the student's expenses (tuition fees and books) providing that the student passes the course. For those who wish to pursue a post-graduate degree full-time, education leave with some financial support can be granted, providing that the subject of study is relevant. As well, there are links with professors from these universities. Statistics Canada has been very fortunate to have had the assistance of Dr. J.N.K. Rao for more than thirty years. Professor Rao has often been a teacher of some of the formal courses mentioned above, and is also a consultant one and a half days a week to the staff of the Branch. Visiting professors are another valuable resource. Other external training opportunities include conferences and workshops. Attendance at these is usually limited by budget, and for conferences usually requires the participant to present a paper. For this reason, Statistics Canada has organized its own *International Methodology Symposium*, held each fall, where staff can attend and hear from leading experts in the field. The Symposium is preceded by a day of workshops given by experts in the field; for example in 2006 there were workshops on record linkage and analysis of longitudinal health data.

Self-learning, or knowledge sharing, is another approach taken by the Methodology Branch to meet the challenges described in Section 2, and is an important part of the Personal Learning Plans of employees. There are several resources available to employees to assist them. First, each division has its own library of technical books and journals in addition to those in the Statistics Canada library. Also, each employee receives a copy of the journal *Survey Methodology*, published twice per year and containing the latest articles on the state of the art. Each employee also receives a copy of the Statistics Canada publication *Survey Methods and Practices (2003)*. This is a practical guide to survey planning, design and implementation and now serves as the textbook for the SSDC course mentioned above. The *Analysis Handbook* is another textbook inspired by the many consultations done by Statistics Canada's Data Analysis Resource Centre and contains practical information for analysts of Statistics Canada's data. The Branch has also had a series of Working Papers, stretching back over many years. Most recently, the Branch has created a methodology documentation database, called MEDOC, which makes methodology documents from a wide variety of sources easily accessible to staff. The database currently contains over 1100 documents from sources such as recent Symposia, the Advisory Committee on Statistical Methods, Methodology Branch Working Papers, internal reports, and so on. The Branch also has discussion groups on specific topics, such as imputation practices and longitudinal surveys. One of these, on imputation practices, publishes an *Imputation Bulletin* that describes the latest developments in the area. It is disseminated both internally and externally.

Finally, the agency attempts wherever possible to leverage its investments in training courses that are provided to external clients. The Branch has a range of courses that it provides to external clients on a cost recovery basis on topics such as sampling, questionnaire design and data analysis. Although these courses are

usually less technical than the courses our own staff takes, there are many elements of these courses that can be borrowed for internal use as well. In fact this was the case for the *Survey Methods and Practices* publication mentioned above; it was originally developed for a technical assistance program for the National Bureau of Statistics of China and then revised for our own use.

From the above description, it will be evident that the key principle in learning and development for statisticians at Statistics Canada is flexibility, but with guidance provided. There are very few mandatory courses (SSDC is the principal one); in general it is up to the employee, with support from their supervisor, mentor and the documents provided by the Branch, to decide what their training needs are and the best way to meet them. By making a wide range of learning opportunities available we hope to maximize the utility to the staff as a whole.

4. Current Curriculum of Courses

Table 2 shows the current listing of courses available in the Statistical Training Program. Although most of these courses are aimed primarily at methodologists, they are open to all staff of the agency providing they have the course pre-requisites.

Table 2: Statistical Training Courses at Statistics Canada

Title	Duration (days)	Description
Questionnaire Design Workshop	3	To provide an understanding of the basic concepts, principles, and practices of questionnaire design for the collection of useful and meaningful data.
Introduction to Statistical Sampling	3	To provide a basic understanding of practical sampling. Principles of sampling are covered with a minimum of formulae. Aimed primarily at non-statisticians.
Statistical Sampling Theory	5	To familiarize participants with statistical sampling methods and their application.
Indirect Sampling and Difficult-to-reach Populations	1	To familiarize participants with Indirect Sampling and the Generalised Weight Share Method; and to apply these methods for surveying populations difficult to reach.
Theory and Application of Longitudinal Surveys	4	To introduce the participants to the basic concepts of longitudinal surveys, to acquire the ability to plan, develop and evaluate panel surveys, to learn estimation and analysis techniques to exploit longitudinal data.
Small Area Estimation	5	To familiarize participants with traditional and model-based methods for small area estimation and their application.
Quality Control Methods for Survey Operations	2	To provide an operational overview of the methods of quality control.
Statistical Methods for Quality Control	3	To provide an overview of the concepts of statistical quality control.
Nonresponse and Imputation: Theory and Application	6 or 8	To provide students with the necessary theoretical knowledge in the field of nonresponse and imputation in surveys and censuses.

Introduction to Record Linkage	2	To provide an overview of record linkage, focusing mainly on probabilistic linkage.
Statistical Methods for the Analysis of Data – Introductory Level	3	To familiarize the participants with the basic mathematics and statistics that underlies the techniques of statistical analysis of data.
Statistical Methods for the Analysis of Data – Intermediate Level	4	To familiarize participants with basic and intermediate techniques for the statistical analysis of data. It is a theoretical course with exercises (with/without SAS).
Statistical Analysis of Survey Data – Module 1: Basic Principles and Some Common Methods	4	To provide participants with some necessary statistical methods to analyze survey data.
Statistical Analysis of Survey Data – Module2: Linear, Logistic and Generalized Logistic Regression	4	To provide participants with more statistical methods useful for the analysis of survey data.
Introduction to Exploratory Data Analysis	5	To familiarize data analysts with the principles, techniques, and tools of Exploratory Data Analysis in order to better understand, edit, and analyze their datasets.
Intermediate Exploratory Data Analysis	3	This course is designed for those wishing to do more formal statistical analysis.
Survival Analysis	4	This course covers both the theoretical and practical aspects of the methodology relating to survival data.
The Components of Time Series	1.5	To enable the participants to recognize, understand and interpret the movements present in time series; and familiarize them with the graphical representation of data.
ARIMA Modelling and Forecasting of Time Series	5	To enable participants to describe and explain the required theory and use the PC SAS statistical system to fit ARIMA models, intervention models or transfer function models to time series.
Seasonal Adjustment with X-12-ARIMA	5	To enable participants to use the X-12-ARIMA seasonal adjustment method to estimate the trend-cycle, seasonal, holiday, trading-day and irregular components of a time series.

5. Evaluation of Training

As mentioned in Section 1, Statistics Canada spends approximately 5% of its salary budget on formal learning and development, so it is important to evaluate what benefit the agency derives from this investment. Statistics Canada uses a four-level model to evaluate the effectiveness of training:

Level 1: Reaction – Did the trainees like the course?
Level 2: Learning – Did they learn it? How well?
Level 3: Behaviour - Did they remember? Did they use it on the job?
Level 4: Results - Did this change organizational effectiveness?

As far as formal courses are concerned, evaluation at Level 1 (Reaction) is universal – all courses ask the participants for feedback on the course and for suggestions for improvement. Level 2 (Learning) is also fairly common; for example courses that involve a case study or a real survey (such as SSDC) make it evident as part of the course whether or not the participants learned the material. University courses also

have formal evaluations of whether the student learned the material. Level 3, while less common, is measured by some courses by following up participants after they are back on the job. For example the SSDC course organizes meetings with former students several months after the course is complete. Level 4 is even harder to measure, but by seeing how groups of employees who have followed certain courses perform in related activities on the job, inferences can still be made.

When it comes to the less formal training methods described in Section 3, evaluation becomes much more difficult. Level 1 can usually be inferred by the popularity of the various learning opportunities that are made available, and appropriate adjustments can be made. Level 2 and above is usually not possible – workshops, seminars and reading are not formally evaluated, and behaviour and organizational change are difficult to correlate with the learning activity.

However the Methodology Branch does have two more formal evaluations available that can at least indirectly measure the effectiveness of its overall learning and development program. The first of these is a client satisfaction survey, conducted every four years. Two of the questions, on the quality of results and the relevance of solutions provided, are related to the quality of the statistical methods provided by methodology staff. Both of these questions received ratings of 6.0 out of 7.0 (7.0 being the highest) in the last survey, conducted in late 2004. The other measure is an Employee Opinion Survey, the last of which was held in 2005. One of the questions asked employees whether they felt they had the training needed to do the job; some 94% agreed that they did. Another question asked whether the supervisor helped the employee determine their learning needs; in this case 75% agreed. So from these indicators at least, it appears that the current approach is effective.

REFERENCES

Gambino, J., and Gough, H. (2005) *Teaching Sampling in a Government Statistical Agency: The Canadian Experience.* Proceedings of the 55th Session of the International Statistical Institute.

Statistics Canada (2003). *Survey Methods and Practices.* Catalogue no. 12-587-XPE

RÉSUMÉ

Statistique Canada emploie plus de 300 statisticiens d'enquêtes, aussi appelés des méthodologistes, qui sont chargés de la conception statistique de pratiquement tous les programmes de l'organisme. Afin de donner une formation efficace en statistique et en méthodes d'enquêtes à ce groupe d'employés essentiels, il faut surmonter un certain nombre de difficultés. Celles-ci comprennent la diversité des parcours universitaires du public cible, la pression de plus en plus forte sur le temps alloué à la formation, la pénurie de ressources disponibles pour élaborer et offrir des cours, la nécessité d'offrir la formation en anglais et en français, et l'intensification de la concurrence d'autres types de formation, dont certains deviennent obligatoires. Le présent document décrit certaines des façons par lesquelles Statistique Canada a réussi à relever ces défis, comme l'élaboration d'un guide sur le perfectionnement professionnel, la personnalisation de la formation, l'utilisation de ressources externes, le développement de ressources afin de soutenir l'autoapprentissage et l'offre de cours de formation de l'organisme aux clients externes. Le programme actuel des cours internes est décrit brièvement. Le document se termine par quelques pensées sur l'évaluation de notre programme de formation.

Mots clés : formation, méthodes d'enquêtes, autoapprentissage, cours