Canada's General Social Survey: Five Years of Experience

D.A. NORRIS and D.G. PATON[1]

ABSTRACT

The Canadian General Social Survey is an annual survey that aims to provide data on the demographic and social characteristics of Canadians. This paper provides an overview of the program, based on the experience of the first five surveys. The objectives of the program, the methodology used, the themes and issues addressed, the program outputs and the plans for the future are all discussed.

KEY WORDS: Social surveys; Telephone surveys; Random digit dialing; Time use surveys; Health surveys.

1. INTRODUCTION

Statistics Canada's social statistics program is concerned with providing information on the demographic and social characteristics and conditions of Canadians. The program's output sustains the development of policy on many critical social issues.

The Census of Population, held every five years, is the cornerstone of the social statistics program, providing benchmark information on the demographic, social, and economic conditions of the population and the basis for future sample surveys of the population. In addition to the Census, activities include on-going surveys and other statistical programs, many based on administrative data sources, in the areas of Health, Education, Culture, Justice, Public Finance, Employment and Unemployment, Income and Expenditures and Demography.

While household surveys have long been an important part of the social statistics program, the regular survey program has historically been directed mainly at labour market and income related issues and there have been no regular ongoing surveys in areas such as health, education, justice or culture. In order to partially fill this data gap Statistics Canada established in 1985 a General Social Survey (GSS) program.

The purpose of this paper is to outline the nature and scope of the GSS program and to describe its evolution over the past five years. Included is a description of the methodology and the content of the five surveys that make up the program. Finally there is a brief discussion of some future directions for the program.

2. GSS PROGRAM OBJECTIVES AND STRUCTURE

The period 1930-1980 witnessed a rapid rise in the number and size of social programs in Canada. Whereas in the early 1930's all government expenditures on social programs accounted for about 10% of GNP, by the early 1980's this expenditure had climbed to about 30%. Along with this rise came an increased demand for and use of data and information to monitor and analyze social trends, and over the years, Statistics Canada expanded its social statistics program to meet growing requirements. Nonetheless, the more extensive use of available data in recent

[1] D.A. Norris and D.G. Paton, Statistics Canada, Ottawa, Ontario, Canada, K1A 0T6.

years revealed major areas of weakness where relevant data were too narrow and restrictive for the effective planning of policy programs, products, and services, or for determining the allocation of resources between competing alternatives.

In the early 1980's, a shortcoming of the social statistics program was that aside from the labour market and income areas, most other social data were derived from administrative records or surveys of institutions. These data sources provided only limited information on the population who came in contact with social institutions and no data on the need for, or impact of, social programs on the general population. Such data can only be obtained through a general population survey.

While a case could have been made for regular, frequent and large scale surveys in a variety of fields (*eg.* health, education, criminal victimization), resources to mount such a large scale program were not available. Instead Statistics Canada initiated a much more modest annual General Social Survey which over five years would cover major topics of importance and which would in the long term serve as a vehicle for monitoring social change. In the short term it could also serve as a vehicle to collect limited data on topics of current social policy interest. The total annual budget for the GSS was originally set at about one million dollars (CAN) and the program was funded by an internal reallocation of Statistics Canada resources derived from efficiency gains in the Labour Force Survey program.

The objectives of the GSS program are two-fold:
• To gather data with a degree of regularity on a broad range of social trends in order to monitor temporal changes in the living conditions and well-being of Canadians; and,
• To provide information on specific social policy issues of current or emerging interest.

To meet these objectives, the GSS program was established with an annual survey cycle. In order to cover the wide range of social issues for which data are required, the GSS program consists of five survey cycles, each covering a different core topic. The collection of data for these topics is thus repeated every five years. The core topics identified for the five cycles are:
1. Health
2. Time Use
3. Personal Risk (accidents and criminal victimizations)
4. Education and Work
5. Family and Friends.

An additional objective in planning content was to include questions that would be useful in deriving indicators of the quality of life, for example, measures of life satisfaction, attitudes, perceptions, or beliefs.

The content of a GSS cycle consists of the following three modules:
• Core content, which is repeated every five years in order to gather information to monitor trends in living conditions and well-being.
• Focus content, which varies from survey to survey and is aimed at the second survey objective of providing information on specific policy issues of particular interest to certain federal departments or policy groups.
• Classification content, which is collected in every cycle and consists of a set of basic demographic and socio-economic variables that enable the delineation of various population groups to facilitate the analysis of core and focus content.

While core and classification content are funded by Statistics Canada, costs associated with focus content are recovered from sponsors.

The target population for the GSS consists of the non-institutionalized population aged 15 and over living in the ten provinces. It was decided that the Labour Force Survey would not

be used as a vehicle for the GSS in order to avoid placing an excessive response burden on LFS respondents and to allow the GSS to use sampling and collection methodologies and sample allocations that differ from those of the LFS. The target sample size for each cycle is 10,000 individuals which was arrived at as a compromise between the competing demands of precision of estimates, budget and length of interview. However, there exists within the GSS program the potential for sponsors to expand the sample for a target population or geographic area. The first survey on health was conducted in late 1985 and the other surveys followed at approximately one year intervals. The fifth cycle on the family was conducted in early 1990 and data collection for Cycle 6 began in January 1991.

The themes and research issues which are covered by each of the surveys are discussed in more detail below. However, before considering these the methodology of the survey is examined in more detail.

3. METHODOLOGY

3.1 Requirements and Constraints

The following are the principal methodological requirements of the GSS: i) it should allow for extensive analysis of the adult Canadian population at a national level and somewhat less detailed analysis at the regional level (this has implications on both the sample size and on the amount of data collected from each respondent); ii) it should have an acceptable cost; iii) it should have a design appropriate for a multipurpose survey; and iv) it should provide public use microdata sets that could be used for analysis by researchers outside Statistics Canada without too much difficulty.

These requirements all interact with the choice of data collection mode, sample design and sample size, but the last two were principally responsible for the choice of sample design, while the sample design and the first two requirements were largely responsible for the choice of data collection mode and sample size.

The last requirement suggests that the sample design be simple as the design information that would be necessary to analyse complex survey data cannot generally be made available on public use files. Requirement iii) suggests that the design not be highly optimized for specific variables.

3.2 Mode of Data Collection

The choice of data collection mode involved balancing a number of competing factors: cost per interview, length of interview, response rate, accuracy of information collected and sample size. The level of detail required in the data collected meant that interviews were expected to last 20 to 30 minutes per respondent. To reduce response burden at the household level and to avoid a cluster effect at the household level it was decided that only one person per household would be interviewed. The principal data collection methods considered for the survey were: self-completed mail-back questionnaire; personal interview; and telephone interview. The high non-response rates experienced with self-completed mail-back questionnaires were felt to be unacceptable (in terms of potential biases) given the heterogeneity of the target population. Personal interviews were felt to offer a number of advantages that would improve the quality of the data collected such as low non-response rates and low item non-response rates, but suffered from the disadvantage of high cost. In addition, many designs used to reduce the cost of personal interviewing have multiple stages of selection and are highly optimized for a few variables. (To not use a design and frame currently used for personal interviewing would have

been unreasonably costly.) These complicated designs make analysis of the resulting datasets difficult and the optimization leads to high design effects for some variables. These high design effects make such designs less appropriate for multipurpose surveys like the GSS. Experience with telephone surveys at Statistics Canada indicated that fairly high response rates could be achieved at reasonable cost. In addition, random digit dialing (RDD) sampling methods allow the efficient selection of samples that are simple random samples or nearly simple random samples.

For these reasons, the GSS has used telephone sampling (RDD) methods and telephone interviewing for most of its sample in all cycles conducted to date. When there has been a need to focus on special target groups its main sample has been supplemented with individuals selected from list frames. In Cycle 1 it was felt that face to face interviews should be used for many of the interviews with elderly respondents.

3.3   Target Population

The target population of the GSS is all persons over the age of 14 permanently living in Canada, with the following two exclusions: i) residents of the Yukon and Northwest Territories, and ii) residents of institutions. This target population is different from that of the Labour Force Survey, which in addition excludes residents of Indian Reserves and full-time members of the Canadian Armed Forces.

3.4   Sampled Population

The sampling methods used for the GSS exclude some members of the target population from the sample. During weighting, these exclusions are implicitly assumed to be similar to the sampled population (missing at random) and the final weights produce estimates for the target population.

When telephone interviewing methods are used, those persons living in households without telephones are excluded from the sample. This affects less than 2% of Canadian households covered by the Labour Force Survey (Statistics Canada 1989, 1990b). This high rate of telephone penetration is not uniform across age and income groups and varies from province to province: 95.4% of households in the province of Prince Edward Island have telephones while 99.2% of those in the province of Ontario do; 99.1% of households with incomes between 20 and 25 thousand dollars have telephones while only 93.9% of those with incomes less than 10 thousand do. Some subpopulations have much lower rates of telephone ownership than the average; for instance, only 86.7% of low income persons under 65 living alone have telephones.

The GSS does not in general accept proxy responses and so individuals who cannot use a telephone (those unable to hear or unable to speak) or who cannot be reached by phone during the survey period or who do not speak either English or French are excluded from the responding population. (For the sixth GSS cycle (on the health of Canadians) it was decided to accept proxy responses in those situations where the selected respondent could not complete the interview due to a health problem.)

When supplementary samples are drawn from lists of households interviewed by the Labour Force Survey (as was done for GSS Cycles 1, 5 and 6), residents of Indian Reserves and full-time members of the Canadian Armed Forces are excluded from these samples. These exclusions represent less than 0.5% of the population over the age of 65. (This is the only age group that has been sampled this way.)

3.5 Stratification

The stratification used by the GSS is determined by estimation requirements, operational requirements, restrictions on the definition of strata imposed by RDD sampling, weighting problems specific to RDD sampling and the special needs of sponsors. Since some estimates are required at the provincial level the GSS strata never cross provincial boundaries. For operational reasons a stratum must be interviewed from a single Regional Office, thus strata never cross Regional Office boundaries. The RDD sampling method used requires that strata be defined as aggregations of telephone exchanges. During weighting, accurate estimates of the sizes of strata are needed, thus the strata (defined on the basis of telephone geography) need to correspond closely to aggregations of units for which accurate population data or estimates were available. These accurate data are available in intercensal years at the Census Metropolitan Area (CMA) level.

The basic stratification based on these requirements starts with the provincial boundaries as stratum boundaries. In Cycles 1 to 5, Saskatchewan and Ontario were each covered by two Regional Offices, so they were both divided in two by a stratum boundary. Further, within each of the areas thus obtained, the CMA's formed a stratum and the non-CMA areas another stratum. In addition, the two largest CMA's, Montreal and Toronto, were each separate strata. For Cycles 1 to 5 this gives us a total of 25 strata: one in Prince Edward Island (there is no CMA in PEI), two in each of Newfoundland, Nova Scotia, New Brunswick, Manitoba, Alberta and British Columbia, three in Quebec, four in Saskatchewan and five in Ontario. For Cycle 6 there were 21 strata: one in Prince Edward Island, two in each of Newfoundland, Nova Scotia, New Brunswick, Manitoba, Saskatchewan, Alberta and British Columbia, and three in Quebec and Ontario.

This is the basic stratification used by the GSS, but modifications to this basic common stratification to accommodate the particular needs of the subject matter or of the sponsors are possible and have been implemented. In Cycle 2 the special interest in language use indicated that separate strata with higher sampling fractions should be used in "contact regions" in which there were thought to be large numbers of both anglophones and francophones. In Cycle 5 the interest of a client in producing estimates for certain sub-provincial regions of Ontario led to the definition of a special stratification.

3.6 Allocation

The target sample size of the GSS is 10,000 completed interviews. This sample has been allocated to provinces in proportion to the square roots of their population sizes. The allocation to strata within provinces has been in proportion to their sizes. The square root allocation is a method of increasing the sample sizes for the smaller provinces (when compared with a proportional allocation) without compromising the precision of Canada level estimates as much as an equal allocation. The method of Kish (1976) for arriving at an allocation that explicitly balances the need for provincial and Canada level precision has been investigated, but the resulting allocations yield little improvement in the precision at the Canada level while changing the allocations to some provinces dramatically and in a way felt to be undesirable.

3.7 Telephone Sampling Method

Except for supplemental samples of the population over 65 selected using lists of households interviewed for the Labour Force Survey, the GSS samples have been selected using random digit dialing methods. Two methods of sample selection have been used, the Waksberg (1978) method and the elimination of non-working banks method. Both methods use information obtained from telephone companies to improve the success rate of reaching households. The choice of methods depends on the level of detail of the information available.

Telephone numbers in Canada are ten digit numbers that can be decomposed into a three digit "Area Code", a three digit "Prefix", and two two digit fields, the first of which we refer to as a bank identifier. Thus within each "Area Code-Prefix" (ACP) there are ten thousand possible numbers and within each "Area Code-Prefix-Bank" (ACP-Bank or simply bank) there are one hundred possible numbers. For example, here is a fictitious telephone number and its components:

<div align="center">216-357-4675</div>

216	Area Code
357	Prefix (exchange)
46	Bank Identifier
75	Number
216-357	ACP
216-357-46	ACP-Bank (bank).

When the only information that is available is a list of ACP's, the GSS uses the Waksberg method of generating the sample. In this method, banks are selected with probability proportional to size, where the size measure is the number of residential telephone numbers in the bank. Within each selected bank a simple random sample of residential numbers is selected. When the sample size is the same in each bank, this method yields an equal probability sample of residential telephone numbers. The sample size within banks used by the GSS has been 6. This method has the advantage of improving the success rate of selecting residential numbers with the disadvantage of producing a clustered sample. For instance, in some rural areas of western Canada only approximately 6% of the numbers generated using lists of ACP's are residential, while the success rate during the second stage of seletion is about 50%. The design effects due to clustering are small for many variables, on the order of 1.0-1.3.

When more detailed information is available that allows the creation of a list of banks containing one or more residential numbers ("working banks") the method which we call the elimination of non-working banks method (ENWB) is used. A simple random sample of numbers within the working banks is selected and non-residential numbers are rejected, yielding a simple random sample of residential numbers. Since the first GSS in 1985, sampling has shifted more and more to the ENWB method as more information has become available from the telephone companies. For Cycle 6 (conducted in 1991) the ENWB method was used for the entire sample.

A system of computer programmes for the Regional Offices of Statistics Canada has been written to implement these two sampling schemes and to monitor the progress of the survey. Within a stratum, the entire sample must be generated using the same sampling method.

After a household has been reached by telephone, a list of the names and ages of all household members is collected and, using this list and a set of random numbers printed for each questionnaire, one person 15 years of age or older in the household is selected to be interviewed. This is the method of Kish (1949).

3.8 Special Samples

Sponsors of the GSS have the opportunity to fund additional interviews. These additional samples can be simple increases in the RDD sample size for one or more strata or they can be drawn from other sampling frames.

In Cycles 2 and 5 the RDD samples in strata of special interest to sponsors were increased.

In Cycles 1, 5 and 6 additional samples of special interest groups were used to supplement the RDD sample. In these cases samples of persons aged 65 and over were selected using lists of households that had recently been part of the LFS sample.

3.9 Response Rates

One disadvantage of telephone surveys is that respondents seem to find it easier to refuse to participate in a telephone survey than in a survey with personal interviews. Telephone soliciting is being used regularly by businesses to sell products and services and everyone has to learn to say no over the phone. In addition, new technologies such as answering machines and special features being added to telephone systems are making it possible and easy for people to screen their incoming calls.

Table 1 gives response rates for the first five cycles of the GSS. The categories "Other Household Non-Response" and "Other Respondent Non-Response" include non-interviews due to language problems, illness, death in the family and absence for the survey period; some of these non-responses are undoubtedly refusals in disguise. In all cycles except Cycle 2 interviews were conducted as soon as possible after contacting the households. In Cycle 2 there was a gap of about a month between the initial contact with the households and the interviewing; there is a component of non-response that can be directly attributed to this time lag. From the table it seems that there may be a trend toward lower response rates over the five cycles.

If we consult Table 2, which presents response rates for individual Regional Offices for Cycles 3 to 6, we see that the situation is not so simple, with many offices (Halifax, Montreal, Winnipeg) showing little change in response rate over these cycles. In fact if we exclude the results obtained in Toronto, the response rate declined only slightly between Cycle 3 and Cycle 5. We have observed that more experienced interviewers tend to be more successful at achieving high response rates. The dramatic change in response rates over three cycles experienced by the Toronto office may in large part be due to the difficulty in hiring and retaining staff in a city that at the time had a booming economy. It is also possible that some of the change is due to a change in the population sampled from the Toronto office.

Preliminary results from eight (January to August 1991) months of data collection for Cycle 6 indicate (see Table 2) that it was possible to reverse the trend to lower response rates. There were a number of changes made between Cycles 5 and 6, the most important ones being a change to monthly data collection and the reassignment of the sample from offices not used for data collection in Cycle 6: St. John's' sample was transfered to Halifax, Toronto's to Sturgeon Falls and Edmonton's to Winnipeg.

During data collection for Cycle 3 it was noted by interviewers that an increasing number of calls were answered by answering machines. This raised the concern that respondents might use these machines to screen their calls, resulting in higher non-response rates. We are not able

Table 1

Response and Non-response Rates (%) by Cycle and Type

Result	Cycle				
	1	2	3	4	5
Household Refusal	6.2	6.2	6.0	7.2	10.3
Other Household Non-Response	4.4	6.8	6.6	6.4	7.2
Respondent Refusal	1.3	2.8	1.3	1.7	2.4
Other Respondent Non-Response	4.8	3.5	3.2	3.9	4.3
Special Cycle 2 Non-Response		1.9			
Response	83.4	78.9	82.9	80.7	75.8

Table 2

Response Rates by Cycle and Regional Office
(results for Cycle 6 are preliminary – indicates offices not conducting interviews)

Regional Office	Cycle			
	3	4	5	6
St. John's	84.1	82.8	90.9	–
Halifax	84.7	84.1	85.9	82
Montreal	83.0	79.6	81.2	82
Sturgeon Falls	76.5	81.1	71.5	71
Toronto	87.0	75.4	63.0	–
Winnipeg	84.3	87.0	84.3	89
Edmonton	83.2	79.4	76.8	–
Vancouver	75.3	80.2	79.6	82
Canada	82.9	80.7	75.8	81
Canada (without Toronto)	82.1	81.8	80.1	–

Table 3

Response and Non-response Rates (%) by Type and Contact with Answering Machines

	Did any calls reach an answering machine?			
	Cycle 4		Cycle 5	
	No	Yes	No	Yes
Household Refusal	7.18	8.17	10.34	9.74
Other Household Non-Response	6.45	5.89	7.19	7.41
Respondent Refusal	5.05	4.55	4.39	3.27
Other Respondent Non-Response	1.68	2.01	2.31	3.15
Responses	79.64	79.38	75.76	76.44
Number of Records	10,981 (93.6%)	747 (6.4%)	16,611 (90.6%)	1,715 (9.4%)

Table 4

Response and Non-response Rates (%) by Type and Type of First Contact

	Was the first contact with an answering machine?			
	Cycle 4		Cycle 5	
	No	Yes	No	Yes
Household Refusal	7.24	7.19	10.46	7.90
Other Household Non-Response	6.46	5.40	7.15	8.06
Respondent Refusal	5.08	3.96	4.38	3.07
Other Respondent Non-Response	1.70	1.80	2.43	1.92
Responses	79.52	81.65	75.58	79.05
Number of Records	11,172 (95.3%)	556 (4.7%)	17,023 (92.9%)	1,303 (7.1%)

to identify those calls that were answered by a machine for cycles 1 to 3, but we are for subsequent cycles and so can analyze to some extent the effect of their use on response rates. Table 3 compares the response rates for those households for which none of the calls were answered by a machine with those for which at least one call was. No important effect of answering machines is indicated by this table; however the increase in contacts with answering machines, from 6.4% to 9.4% of households, is dramatic (Table 3). Table 4 compares the response rates for those households for which the first answered call was answered by a machine with those for which it was not. If any effect of answering machines is indicated by this table it is that response rates are higher for those households with a first contact by answering machine. There appears to be no evidence that the use of answering machines is seriously reducing response rates.

3.10 Data Capture and Processing

The data for all five cycles were captured directly into the mini-computers in Statistics Canada's regional offices. Some simple edits to check the validity of data as captured were made at the time of capture, but these could in most cases be overridden using special functions. Following transmission of the raw data to Ottawa an exhaustive set of edits was applied to find, and correct if possible, invalid or inconsistent responses. When a response was missing, invalid or inconsistent with other responses and the approriate value could not be inferred from other responses on the questionnaire an 'unknown' code was assigned. Exceptions to this rule were three variables needed for weighting purposes: age, sex and number of telephone lines. In cases where these variables were missing the questionnaires themselves were consulted to assist in the imputation of values.

3.11 Weighting

3.11.1 Initial Weights

Both the Waksberg and ENWB methods of selecting RDD samples yield self-weighting samples of residential telephone numbers. The Waksberg method does not provide an estimate of this weight, but for GSS weighting purposes it is sufficient to use an initial weight of one (1) for telephone numbers in those strata where that method is used. In ENWB strata the initial weight is the reciprocal of the probability of selection of the telephone number. This probability is simply:

$$\frac{n_c}{100 \times N_B},$$

where:

n_c is the number of telephone numbers selected, and

N_B is the number of working banks in the frame.

3.11.2 Non-response Adjustment

The initial weight is adjusted for non-response using adjustment "strata" based on telephone geography. These are typically banks in Waksberg method strata and ACP's in ENWB strata. The initial weights are inflated by the following factor:

$$\frac{n_R + n_{NR}}{n_R},$$

where:

n_R is the number of responding households in the non-response "stratum", and

n_{NR} is the corresponding number of non-responding households.

3.11.3   Telephone Adjustment

Since households with more than one telephone line have a higher chance of being selected by an RDD survey, the initial weight adjusted for non-response (a weight for telephone numbers) is further adjusted by dividing by the number of telephone lines for the household to yield a household weight.

3.11.4   Initial Person Weight

Since only one eligible respondent per household is interviewed, the household weight must be adjusted by multiplying by the number of eligible respondents to yield a person weight.

3.11.5   Poststratification

At this point populations projected from the census are used as reference totals in the poststratification of the person weights, first to the stratum population sizes and then to the provincial age-sex populations. (It should be noted that it is only after the first stage of poststratification that the weights in Waksberg strata actually sum to a population estimate. Until this step they differ from a set of weights based on the inverses of the selection probabilities by an unknown constant of proportionality.) These two sets of reference totals are then used as the margins for a raking ratio adjustment to the weights.

4.   THEMES AND RESEARCH ISSUES COVERED BY THE GSS

As indicated above, in order to cover a wide range of social issues, the GSS examines a different core topic each year for five years and then the topics are repeated. The core topics were chosen to fill perceived data gaps in the social statistics program. The five core themes are discussed in more detail below.

4.1   Health

The core content of the health cycle is directed at providing a range of measures of health status, including short and long term disability, the prevalence of common chronic conditions, such as high blood pressure or diabetes, and the use of various health care services. In addition, data are collected on life-style such as, smoking, drinking, and physical exercise. When linked to health status, these data provide information on the barriers (*e.g.* smoking, drinking) and bridges (*e.g.* physical exercise) to positive health for various population groups.

For the first GSS health cycle, the add-on focus content was directed at older Canadians and covered social networks, support given and received, as well as participation in a range of social activities. The sample size for the elderly population was also increased to allow for more in-depth analyses.

4.2   Time Use

The GSS time use survey consisted of a "24 hour time budget" generally for the day preceding the interview. Respondents provided information on each primary activity engaged

in during that day, the start time and duration of each activity, and associated information on where the activity took place and who was with the respondent at the time (*e.g.* spouse, children, friends, *etc.*). These data provide information on the frequency with which people participate in activities such as paid work, household work, attending cultural events, watching television, and the time spent on these activities.

The survey provides information on how Canadians allocate their time to activities such as paid work, housework and other non-market work and leisure activities. The data can be used to show constraints that limit a person's choice of the use of time and how these are distributed among different population groups. The inclusion of a battery of questions on satisfaction with various dimensions of life allows such measures to be correlated with patterns of time use for different population groups.

The 1986 GSS time use cycle also included a small module on intergenerational social mobility that allows for the analysis of movement on an occupational or educational hierarchy between the respondent and his or her parents.

The add-on focus content for the time use cycle was a detailed set of questions on language knowledge and use. While focus content is generally expected to be related to and complement core content, there was a demand for much more detailed language data than could be included in the population census. The information collected included data on language use at various stages of life (*e.g.* first learned, during childhood, at school) and in various settings including at home, at work, with friends, watching television, and in dealing with federal agencies. In order to allow a more detailed analysis in bilingual regions of the country, sample size was also increased in these geographic areas.

4.3 Personal Risk

The third GSS cycle was based around the topic of personal risk, including both criminal victimizations and accidents. Traditionally, information on these topics has been derived from administrative sources, such as police statistics and hospital records. However, these data provide very little information about the victim and, in addition, there are many crimes (the GSS estimates more than half) and accidents which are not reported to authorities.

The personal risk survey conducted in early 1988 asked respondents about criminal victimizations and accidents that they had experienced during calendar year 1987. Data were also collected on several life-style measures, such as alcohol consumption and frequency of night outings to allow these to be correlated with criminal victimizations and accidents. For each reported crime or accident incident, data were collected on the nature of the incident, the consequences in terms of activity restriction, medical attention and financial loss. In addition, respondents were asked to report their perceptions of crimes and accidents and about precautions taken to prevent these events.

The add-on focus content for the personal risk cycle was a set of questions on contact with the criminal justice system (*e.g.* police, courts, lawyers) and on the awareness and use of services by victims of crime.

4.4 Education and Work

While the monthly Labour Force Survey and other labour related surveys provide a wealth of information about the labour force, none of the existing surveys provides much information on the social aspects of work or the perceived quality of working life. The GSS cycle on education and work, conducted in early 1989, was designed to partially fill this data gap.

The survey was developed around three main themes that reflect fundamental changes in Canadian society: patterns and trends in work and education; new technologies and human resources; and work in the service economy. The themes reflect a range of issues on which more information is required. For example, the accelerating rate of technological innovation demands detailed knowledge about the utilization of and training for computers. Concerns about the effective utilization of the nation's human capital require a better understanding of the links between the labour force and the educational system. We also must anticipate future demands on educational institutions, and changing relationships between educational attainment and socio-economic outcomes. This round of the GSS also augments existing data sources by providing new information about the elderly population as well as some of the socio-economic implications of the baby boom generation entering middle age.

The survey collected a partial work and education history. It also included information on technology training and the use of computers, and on future plans for education. Subjective information also was sought in the form of a series of questions about satisfaction with retirement and other dimensions of life, as well as a block of questions on attitudes to science and technology.

4.5 Family and Friends

The fifth cycle of the GSS was based around issues related to family and friends and was completed in early 1990. While the Census and other household surveys provide family-based data, changes in family life have resulted in a need for new types of information. One shortcoming of existing data is that generally they are based on a rather narrow concept of the family, in particular a nuclear family of parents and children or perhaps an economic family of related individuals living in the same household. This survey looks at the family in a broader context and collects information on the extent and nature of kinship networks and related questions of patterns of informal help and support among family and friends.

A second major theme of the survey is a result of the trends in marriage, divorce, and the increased frequency of common law unions. Increased numbers of Canadians are living in more than one union during their life time. The impact of such changes on family life and children is substantial and can best be studied by an analysis of marital and family history data. Such data were retrospectively collected in a special Family History Survey conducted in 1984 (Burch 1985). The GSS family cycle incorporates the collection of these data on a regular basis. Specific issues that can be addressed include changing patterns of union formation and dissolution, the situation of single parent families, and home leaving patterns of young adults.

A third but more minor theme of the cycle is concerned with the division of household labour.

5. PROGRAM OUTPUTS

The GSS results are disseminated in a variety of ways. For each survey there are one or more publications that present the results of data analysis with respect to particular social issues and the monitoring of conditions and trends. The results of Cycle 1 are reported in Statistics Canada (1987) and Stone (1988); the results of Cycle 2 are presented in Harvey *et al.* (1991) and Creese *et al.* (1991); and the Cycle 3 results are reported in Sacco and Johnson (1990) and Millar and Adams (1991). Publications containing the results of other cycles are in preparation. The general public are made aware of GSS results through the publication of reports in the media which are often based on articles published in *Canadian Social Trends*, a quarterly Statistics Canada publication that is targeted to a general audience.

A second product is a public use microdata file and associated documentation to enable university and other researchers to carry out their own analysis of the data. These data are also useful for teaching purposes. Microdata files from the first five survey cycles are now available.

In addition to the product outputs, the GSS program has developed a survey capacity. This is not simply a system for data collection and processing, but includes other major components. Content research and development, related data specification, analysis of survey and other relevant data, dissemination of informative results as well as the development and use, where applicable, of improved methods of collection, processing, analysis and dissemination are all components of the evolving survey capacity of the GSS group.

6. FUTURE PLANS

As the GSS program moves into the second round of surveys, attention has shifted from the problems of developing and fielding five new surveys to further building the survey program through partnerships with others. The first round of surveys has had a modest success with obtaining buy-ins of additional sample and/or focus content. Only Cycle 4 had neither focus content nor increased sample size. For the first time, the 1990 survey had provincial participation, with the Ontario government funding an increase in sample size.

A new initiative for the GSS program is an investigation of the potential for expanding the scope of the survey to include interviewing a sub-sample of respondents again in future cycles. In the short term, this could provide an enriched data set by linking content from different cycles. In the longer term, it could serve to provide longitudinal data by interviewing respondents on the same topic five years later. A feasibility study was conducted in 1990 and the possibility of interviewing a sample of respondents from a previous cycle is now offered to interested sponsors.

The GSS will also continue to undertake a range of more general research and development activities. Core content of the first set of cycles will be reviewed and input sought from users as to possible improvements for future cycles. While new and alternative survey designs and approaches will be considered, any potential changes will have to be balanced against the impact on data comparability that is required for the long term goal of monitoring change. In addition, the content from the first round of surveys will be reviewed from the point of view of consistency and integration across GSS survey cycles and between the GSS and the 1991 Census and other household surveys. On-going development of the GSS infrastructure will also continue. Consideration was given to changing to monthly data collection (and monthly data collection was implemented for Cycle 6) and will be given to supplemental collection methods (*e.g.* mail). Attempts are also being made to shift processing of the survey to a micro-computer environment to further improve timeliness. Finally, new procedures, such as computer-assisted telephone interviewing, will be considered as these become available as part of a larger Statistics Canada survey development program.

In summary, the GSS Program during the coming years will focus on building on the firm foundation that has been established during the first round of surveys. The primary objective will continue to be the measurement of social conditions and the gradual development of a time series to monitor trends. In addition, flexibility will be maintained in order to quickly respond to new and emerging social information needs.

REFERENCES

BURCH, T.K. (1985). *Family History Survey: Preliminary Findings*. Catalogue 99-955, Statistics Canada.

CREESE, G., GUPPY, N., and MEISSNER, M. (1991). Ups and downs on the ladder of success. *General Social Survey Analysis Series*. Catalogue 11-612E, No. 5, Statistics Canada.

HARVEY, A.S., MARSHALL, K., and FREDERICK, J.A. (1991). Where does time go? *General Social Survey Analysis Series*. Catalogue 11-612E, No. 4, Statistics Canada.

KISH, L. (1949). A procedure for objective respondent selection within the household. *Journal of the American Statistical Association*, 44, 380-387.

KISH, L. (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society A*, 139, 80-95.

MILLAR, W., and ADAMS, O. (1991). Accidents in Canada. *General Social Survey Analysis Series*. Catalogue 11-612E, No. 3, Statistics Canada.

SACCO, V., and JOHNSON, H. (1990). Patterns of criminal victimization in Canada. *General Social Survey Analysis Series*. Catalogue 11-612, No. 2, Statistics Canada.

STATISTICS CANADA (1987). Health and social support, 1985. *General Social Survey Analysis Series*. Catalogue 11-612, No. 1, Statistics, Canada.

STATISTICS CANADA (1989). *Household facilities and equipment*. Catalogue 64-202, Statistics Canada.

STATISTICS CANADA (1990a). *Overview of Special Surveys 1989*, Household Surveys Division, Statistics Canada.

STATISTICS CANADA (1990b). *Household facilities by income and other characteristics*. Catalogue 13-218, Statistics Canada.

STONE, L. (1988). *Family and Friendship Ties Among Canada's Seniors*. Catalogue 89-508, Statistics Canada.

WAKSBERG, J. (1978). Sampling methods for random digit dialing. *Journal of the American Statistical Association*, 73, 40-46.