THE 1993 GENERAL SOCIAL SURVEY - CYCLE 8

PERSONAL RISK

Public Use Microdata File Documentation

and

User's Guide

Summer 1994

Aussi disponible en fran‡ais

The General Social Survey - Cycle 8

Public Use Microdata File Documentation

and User's Guide

Table of Contents

1. INTRODUCTION

This package is designed to enable interested users to access and manipulate the microdata file for the eighth cycle of the General Social Survey (GSS), conducted from February through December, 1993. It contains information on the objectives, methodology and estimation procedures as well as guidelines for releasing estimates based on the survey.

Appendix A contains the Guidelines for Working with Cycle 8 Normalized Record Structure Files and Its Multiple Weighting Factors. Appendix B contains the approximate variance tables.

Working Paper #7 entitled 'Overview of the 1993 GSS on Personal Risk', complete with Cycle 8 and 3 questionnaires are presented in Appendix C. Appendix E contains the data dictionary for the microdata file, the major part of this documentation package.

Excluding the appendices, this package is available in machine readable form.

2. OBJECTIVES OF THE GENERAL SOCIAL SURVEY

Increased pressure, during the past decade, to operate more efficient government funded programmes, has led to a related increase in the information needed for policy formulation, programme development and evaluation. Many of these needs could not be filled through existing data sources or vehicles because of the range or periodicity of the information required. The two primary objectives of the GSS aim at closing these gaps. These objectives are: to gather data on social trends in order to monitor temporal changes in the living conditions and well-being of Canadians; and to provide immediate information on specific social policy issues of current or emerging interest. The GSS is a continuing program with a single survey cycle each year.

To meet the stated objectives, the data collected by the GSS are made up of three components: Classification, Core and Focus.

Classification content consists of variables which provide the means of delineating population groups and for use in the analysis of Core and Focus data. Examples of classification variables are age, sex, education, and income.

Core content is designed to obtain information which monitors social trends or measures changes in society related to living conditions or well-being. The eighth cycle of the GSS marks the first

repeat of the GSS core subject on Personal Risk related to criminal victimization and accidents (see GSS 1988 - Cycle 3). One important application of data on personal risk is to measure the incidence of accidents and criminal victimizations in order to complement the officially reported data for these incidents.

Focus content is aimed at the second survey objective of GSS. This component obtains information on specific policy issues which are of particular interest to certain federal departments or other user groups. In general, focus content, is not expected to be repeated on a periodic basis.
Focus content for Cycle 8 covered alcohol and drug use. In addition to being of interest in relation to some of the core content, these data provide an update to information first collected in the 1989 National Alcohol and Drug Survey. Focus content was sponsored by the Health Promotion Directorate of Health Canada.

3. CONTENT AND SPECIAL FEATURES OF THE GSS CYCLE 8

Cycle 8 was the first time the GSS collected data using Computer Assisted Telephone Interviewing (CATI). With CATI, the survey questions appeared on a computer monitor. The interviewer asked the respondent the questions, and entered the responses into the computer as the interview progressed. Built-in edits and fewer processing steps resulted in better quality data. CATI methodology also eliminated the need for paper and pencil questionnaires. As a result, these forms were produced as reference documents only (see Appendix C). In Cycle 8, the CATI system provided the interviewer with four main "components" which can be imagined to represent four paper questionnaires.

QUESTIONNAIRE AGE GROUP	TITLE
GSS 8-1 All age groups	Survey Control Form
GSS 8-2 Age 15 and over	Personal Risk Questionnaire

GSS 8-2F
Age 15 and over
Accident Report

GSS 8-2G
Age 15 and over
Crime Incident Report

The GSS 8-1 was completed for each telephone number generated in the sample. When a private household was contacted, all members of the household were enumerated and basic demographic information (e.g. age, sex and marital status) was collected for each household member. A respondent, 15 years of age or over, was then randomly selected by the computer. The relationship of each household member to the selected respondent was collected and the main questionnaire (GSS 8-2), any Accident Reports (GSS 8-2F) and Crime Incident Reports (GSS 8-2G) were completed for this person. Proxy interviews were not accepted.

The main component of the survey (GSS 8-2) collected information on the following topics: the respondent's attitudes towards various components of the judicial system, satisfaction with various aspects of the judicial system, perception of risk with regard to accidents and crime incidents (Section A); information on alcohol and drug use (Section D); basic background information on the respondent (Section E); information on the kind and number of times the respondent had been involved in an accident (Section B) or a crime incident (Section C) during the past 12 months. Section B also collected information regarding accidents involving one of the respondent's randomly selected children (natural, adopted or step-), aged less than 15 years and living in the same househld.

The screening questions in Sections B and C of the GSS 8-2 determined whether an Accident Report or Crime Incident Report should be completed. Only accidents which occurred in the past 12 months (from the date of respondent contact) and interrupted the respondent's normal activities for at least half a day, or caused expenses of $200 or more, or required medical attention from a doctor or a nurse, were to be reported. Similarly, only crime incidents which occurred in the past 12 months (from the date of respondent contact), were to be reported. Crimes to be included

were defined by the screening questions in Section C and were of two general types - crimes committed against the respondent and crimes committed against their household.

It should be noted that there are slight differences between the appended forms and the CATI version of the survey. Specifically, random selection of the respondent was done by a computer algorithm rather than taken from a pre-printed Selection Grid Label as in previous pencil and paper questionnaires. Furthermore, the CATI version asked the respondent for information regarding the relationship of each household member to the selected respondent. In previous cycles and in the version appended, relationship to a designated reference person for each household economic family was collected (reference Z9 of GSS 8-1). Other differences involve items which appear on the forms but do not appear on the CATI version. For example, interviewer check items are visible on the questionnaire but exist only as internal edits in the CATI system. Similarly, skip patterns are visible on the questionnaire but exist internally in the CATI system. Additionally, a few questions, such as date of birth, are asked in a different manner using CATI (eg. instead of asking date of birth, CATI asks three separate questions - year of birth, month of birth and day of birth).

4. SURVEY AND SAMPLE DESIGN

Data for Cycle 8 of the GSS were collected monthly from February 1993 to December, 1993 inclusive. The sample was evenly distributed over the 11 months to counterbalance seasonal variation in the information gathered. All of the sample was selected using the Elimination of Non-Working Banks technique. A description of this method is provided in Section 4.3. The target population is discussed in Section 4.1. Stratification used in the survey design is outlined in Section 4.2, and Section 4.4 discusses sample size.

4.1 Target Population

The target population for the GSS was all persons 15 years of age and over in Canada, excluding:

1. residents of the Yukon and Northwest Territories;
2. full-time residents of institutions.

The survey employed Random Digit Dialling (RDD), a telephone sampling method. Households

without telephones were therefore excluded, however, persons living in such households represent less than 2% of the target population. Survey estimates have been adjusted (weighted) to represent persons without telephones.

4.2 Stratification

In order to carry out sampling, each of the ten provinces was divided into strata or geographic areas. Generally, for each province, one stratum represented the Census Metropolitan Areas (CMAs) of the province and another represented the non-CMA areas. There were two exceptions to this general rule:

- Prince Edward Island has no CMA and so did not have a CMA stratum
- Montreal and Toronto were each separate strata.

4.3 Elimination of Non-working Banks RDD Design

The Elimination of Non-Working Banks (ENWB) sampling technique is a method of Random Digit Dialling in which an attempt is made to identify all working banks for an area (i.e., to identify all banks with at least one household). Thus, all telephone numbers within non-working banks are eliminated from the sampling frame.

For each province, lists of telephone numbers in use were purchased from the telephone companies and lists of working banks were extracted. Each bank was assigned to a stratum within its province.

A special situation existed in Ontario and Quebec because some small areas are serviced by independent telephone companies rather than by Bell Canada. The area code prefixes for these areas were identified by matching the Bell file with a file of all area codes and prefixes. Area code prefixes from Ontario and Quebec and not on the Bell file were identified. All banks within these area code prefixes were generated and added to the sampling frame. Use of the Waksberg method (an alternate RDD method) was not possible for these areas since it requires that an

accurate population estimate be available for the survey area. Such an estimate was not available for the parts of Ontario and Quebec not covered by Bell.

A random sample of telephone numbers was generated in each survey month for each stratum (from the working banks). An attempt was made to generate the entire sample of telephone numbers on the first day of interviewing. Therefore, a prediction of the percentage of numbers dialled that would reach a household had to be made (this is known as the "hit rate"). The hit rate for February, the first survey month, was estimated using information from previous RDD surveys. Hit rates for subsequent months were revised as required based on February's experience.

For Cycle 8 of the GSS, 46.4% of the numbers dialled reached households. An attempt was made to conduct a GSS interview with one randomly selected person from each household.

4.4 Sample Size

The sample consisted of 10,385 people and a GSS 8-1 was completed for each telephone number generated in the sample. The main questionnaire (GSS 8-2), and the Accident Report(s) (GSS 8-2F) and Crime Incident Report(s) (GSS 8-2G) as appropriate were then completed for the selected person.

5. COLLECTION

Data collection for the GSS was conducted by Random Digit Dialling methods and involved four possible questionnaires. Respondents were interviewed in the official language of their choice. The French and English versions of the main questionnaire were identical with the exception of question E14 "What language did you first speak in childhood?". Respondents were not asked if they still understood the language in which they were being interviewed. Interviews by proxy were not allowed. Paper and pencil questionnaires and procedures were field tested in August, 1992 in Halifax and Montreal. The CATI system was tested in November in Halifax. Data collection began in February 1993 and continued through the second week of December 1993. The sample was evenly distributed over the 11 months. All interviewing took place using centralized telephone facilities in four of Statistics Canada's regional offices with calls being made

from approximately 09:00 until 21:00, Monday to Saturday inclusive. The four regional offices were: Halifax, Montreal, Winnipeg and Vancouver. Interviewers were trained by Statistics Canada staff in telephone interviewing techniques using CATI, survey concepts and procedures in a four day classroom training session. The majority of interviewers had no previous computer experience but had telephone interviewing experience.

It would be too lengthy to include all the survey manuals as part of this documentation package. However, they can be purchased (see Chapter 9). Shown below is a list of the manuals used in the survey:

- GSS Personal Risk - Introduction to Computer-assisted Telephone Interviewing (CATI)
- GSS Personal Risk - Content Manual
- GSS Personal Risk - Computer-assisted Telephone Interviewing (CATI) Interviewer's Manual
- GSS Personal Risk - Home Study Program
- GSS Personal Risk - Training Guide
- GSS Personal Risk - Regional Office Procedures Manual

6. PROCESSING

The following is an overview of the processing steps for Cycle 8 of the GSS.

6.1 Data Capture

Using CATI, responses to survey questions were entered directly into computers as the interview progressed. The CATI data capture program allowed a valid range of codes for each question and built-in edits and automatically followed the flow of the questionnaire. The data were transmitted to Ottawa electronically.

6.2 Edit and Imputation

All survey records were subjected to computer edits throughout the course of the interview. With CATI, built-in edits identified invalid or inconsistent information as the interview progressed. As a result, such problems could be immediately resolved with the respondent.

The system principally edited the main questionnaire for possible flow errors, out of range values and missing values. Edits on the 8-1 were limited to a few edits for the respondent's age and sex. The CATI system implemented such edits throughout the course of the interview. If the

interviewer was unable to correctly resolve the detected errors, it was possible for the interviewer
to bypass the edit and forward the data to head office for resolution.

Head office edits performed the same checks as the CATI system as well as more detailed edits.
Records with missing or incorrect information were assigned non-response codes and in a small
number of cases corrected from other information from the respondent's questionnaire. In most
cases editing was 'bottom-up', meaning that specific related information following a question with
a branching pattern was employed to ensure that the branching was correct.

With CATI, a 'Don't know' and 'Not stated' response category were required for every question.
In the edits, 'Don't know' responses were treated as a 'No' response, rather than a 'Not stated'.

Due to the nature of the survey, imputation was not appropriate for most items and thus 'not
stated' codes were usually assigned for missing data. In some cases, the answer was not known
but could be obtained deterministically by the questions which followed or from information from
other areas of the survey.

Non-response was not permitted for those items required for weighting. Values were imputed in
the rare cases where any of the following were missing: age, sex, number of residential telephone
lines and the type of crime (personal or household). The imputation was based on a detailed
examination of the data and the consideration of any useful data such as the ages and sexes of
other household members, and the interviewer's comments.

DVTEL (number of residential telephone lines) was derived from questions E5 to E11 of the
Personal Risk Questionnaire (GSS 8-2). When adequate information to derive DVTEL was not
obtained, it was assigned a value of one (1).

6.3 Coding

Several questions allowing write-in responses had the write-in information coded into either new
unique categories, or to a listed category if the write-in information duplicated a listed category.
Where possible (e.g., occupation, industry, language, education, country of birth, religion and
relationship of offender to victim), the coding followed the standard classification systems as used
in the Census of Population.

6.4 Creation of Combined and Derived Variables

A number of variables on the file have been derived by using items found on the GSS 8-1, 8-2, 8-2F and 8-2G Questionnaires. Derived variable names generally start with DV and are followed by characters referring to the question number or subject. In some cases, the derived variables are straightforward and involve collapsing of categories. In other cases, several variables have been combined to create a new variable. The data dictionary provides comments indicating the origin of these variables.

6.5 Amount of Detail on Microdata File

In order to guard against disclosure, the amount of detail included on this file is less than is available on the master file retained by Statistics Canada. Variables with extreme values have been capped and information for some variables have been aggregated into broader classes (e.g., occupation, religion, country of birth).

The measures taken to cap, group or collapse data have been indicated in the data dictionary.

7. ESTIMATION

When a probability sample is used, as was the case for the GSS,the principle behind estimation is that each person selected in the sample 'represents' (in addition to himself/herself) several other persons not in the sample. For example, in a simple random sample of 2% of the population, each person in the sample represents 50 persons in the population.

For analysis of the adult information on the Personal Risk Screening Questionnaire a weighting factor, WGHT_PER, was placed on the microdata file. This factor represents the number of persons in the population that the record represents. It refers to the number of times a particular record should contribute to a population estimate. For example, to estimate the number of adults who are very worried while waiting for or using public transportation the value of WGHT_PER is summed over all records with this characteristic.

Similarly, for analysis of the child data on the Personal Risk Questionnaire a weighting factor, WGHT_CHD, was placed on the file. This represents the number of children in the population that the record represents. For example, to estimate the number of children who were taken

to a hospital emergency department as the result of an injury or poisoning the value of
WGHT_CHD is summed over all records with this characteristic.

Accident Reports also have a weighting factor, WGHT_ACC, that represents the number of
incidents of a given type. For example, to estimate the number of accidents in the home the
value of WGHT_ACC is summed over all records with this characteristic.

Finally, Crime Incident Reports have a weighting factor, WGHT_CRI, that represents the number
of crime incidents of a given type. For example, the number of assaults can be estimated by
summing the value of WGHT_CRI for all records with this characteristic.

The process of deriving the adult weight, WGHT_PER, and the child weight, WGHT_CHD, is
described in Section 7.1. Section 7.2 describes the accident weight, WGHT_ACC and Section 7.3
describes the crime weight, WGHT_CRI.

For a description of the file layout, contents and correct interpretation of data, microdata users
must read Appendix A (Guidelines for Working with Cycle 8 Normalized Record Structure Files
and Its Multiple Weighting Factors).

7.1 Weighting of Personal Risk Questionnaire

A self-weighting sample design is one for which the weights of each unit in the sample are the
same. The GSS sample has such a design, each household within a stratum having an equal
probability of selection.

This probability is equal to:

$$\frac{\text{Number of telephone numbers sampled within the stratum}}{\text{Total number of possible telephone numbers within the stratum}}$$

(The total number of possible telephone numbers for a stratum is equal to the number of working
banks for a stratum times 100).

Where possible, each survey month was weighted independently. This was done in an attempt to
ensure that each survey month contributed equally to estimates. If monthly sample sizes were not
large enough, two or more survey months were combined in certain steps of the weighting.

1) Basic Weight Calculation

Each household (responding and non-responding) in the sample was assigned a weight equal to
the inverse of its probability of selection. This weight was calculated independently for each
stratum-month group as follows:

$$\frac{\text{Number of possible telephone numbers in each stratum-month group}}{\text{Number of sampled telephone numbers in each stratum-month group}}$$

2) Non-Response Adjustment

Weights for responding households were adjusted to represent non-responding households. This
was done independently within each stratum-month group. Records were adjusted by the
following factor:

$$\frac{\text{Total of the household weights of all households in each stratum-month group}}{\text{Total of the household weights of responding households in each stratum-month group}}$$

Non-responding households were then dropped.

3) Multiple Telephone Adjustment

Weights for households with more than one residential telephone number (i.e. not used for
business purposes only) were adjusted downwards to account for the fact that such households had
a higher probability of being selected. The weight for each household was divided by the number
of residential telephone numbers that serviced the household.

4) Adult Weight and Child Weight Calculation

An adult weight was then calculated for each respondent to the survey, by multiplying the
household weight by the number of persons in the household who were eligible to be selected for
the survey (i.e. the number of persons 15 years of age or older).

A child weight was also calculated for each selected child by multiplying the adult weight by the
number of respondent's children less than 15 years of age living in the household and dividing
by the number of parents the child has living in the household.

5) Adjustment of Adult Weight to External Totals.

The adult weights were adjusted several times using a raking ratio procedure. This procedure ensured that, based on the survey's total sample, estimates produced of the sizes of strata or of province-age-sex groups would match external references. The two sets of groupings used for these adjustments were stratum-month and province-age-sex. The age groupings used were:

15-19, 20-24, 25-29, 30-34, 35-39, 40-44,
45-49, 50-54, 55-59, 60-64, 65-69, 70+.

Sample sizes were too small to allow the province-age-sex adjustments to be made at the survey month level. Also due to small sample sizes, there were cases where two or more adjacent age groups in the same province-sex group or two adjacent months in the same stratum, were collapsed before the adjustments were made.

The reference totals for the stratum-month adjustments were one eleventh of the population projections for each month. The reference totals for the province-age-sex adjustments were the average of the population projections for each month. At each stage in the adjustment process the weights were adjusted by the factor:

$$\frac{\text{reference total for group}}{\text{sum of adult weights for group}}$$

The groupings used for the adjustments alternated between province-age-sex and stratum-month until the weights converged.

It should be noted that adults living in households without telephone service are included in the reference totals even though they were not sampled.

6) Adjustment of Child Weight to External Totals.

The child weights were adjusted to external totals using the same procedure as for the adult weight. The two sets of groupings used for the adjustments were stratum-quarter and province-age-sex. The age groupings used were: 0-4, 5-9, 10-14.

Sample sizes were too small to allow adjustments to be made at the survey month level. Also due to small sample sizes, there were cases where two or more adjacent age groups in the same

province-sex group or two adjacent quarters in the same stratum, were collapsed before the adjustments were made.

The reference totals for the stratum-quarter adjustments were one quarter of the population projections for each quarter. The reference totals for the province-age-sex adjustments were the average of the population projections for each month. At each stage in the adjustment process the weights were adjusted by the factor:

$$\frac{\text{reference total for group}}{\text{sum of child weights for group}}$$

The groupings used for the adjustments alternated between province-age-sex and stratum-quarter until the weights converged.

It should be noted that children living in households without telephone service are included in the reference totals even though they were not sampled.

IMPORTANT: See Appendix A, Guidelines for Working with Cycle 8 Normalized Recored Structure Files and Its Multiple Weighting Factors. This section gives a detailed description of how to work with adult and child data on the microdata file.

7.2 Weighting of Accident Reports (GSS 8-2F)

The Accident Reports were weighted after the Personal Risk Questionnaire had been weighted. The final adult weight, WGHT_PER, was the final accident weight:

WGHT_ACC = WGHT_PER

IMPORTANT: Each Accident Report can represent either a single accident or a series of accidents. See Appendix A, Guidelines for Working with Cycle 8 Normalized Recored Structure Files and its Multiple Weighting Factors. This section gives a detailed description of how to work with Accident Reports and how to interpret series report information.

7.3 Weighting of Crime Incident Reports (GSS 8-2G)

The Crime Incident Reports were weighted after the Personal Risk Screening Questionnaire had been weighted. The final adult weight, WGHT_PER, was the basic crime weight. The final crime weight, WGHT_CRI, was calculated following the procedures given below.

Each of 3,740 Crime Incident Reports was classified as being either a "person" or "household" crime based on the value of DVMSCRIM. DVMSCRIM is a derived variable for the most serious crime reported by a given report. This variable is derived based on a hierarchical assignment of incident seriousness described in field 464 of the data dictionary (Appendix E).

In total, there were 1,550 "person" crime reports, with DVMSCRIM = 101, 202, 203, 304, 609 or 610 while there were 2,042 "household" crime reports, with DVMSCRIM = 405, 406, 507, 508, 711, 712 or 813. 148 records originally had an unclassifiable value for DVMSCRIM (i.e. DVMSCRIM = 999). When possible the 'screen question' which gave rise to these incidents was used to classify them as either personal or household. This was not possible when the screen question was either 'C4C' or 'C11'. In these cases, records were manually reviewed and an assignment made. Of the 148 records which had DVMSCRIM = 999, 92 were classified as personal crimes and 56 were classified as household crimes.

Weighting Person Crimes

An adjustment to the basic crime weight is necessary to account for the probability that the incident could have been reported by the other persons harmed or threatened in the incident. The variable DVCRIVIC is the number of persons 15 years of age or older victimized in the incident.

The basic crime weight, WGHT_CRI, is then calculated as:

```
WGHT_CRI = WGHT_CRI / DVCRIVIC
```

Weighting Household Crimes

Household crimes could have been reported by any member of the sampled household so an adjustment to the basic weight was made for all household members. The variable DVELLIG is the number of household members who are 15 years old or older.

The basic crime weight, WGHT_CRI, is then calculated as:

```
WGHT_CRI = WGHT_CRI / DVELLIG
```

IMPORTANT: Each of the Crime Incident Reports can represent either a single crime incident or a series of crime incidents. See Appendix A, Guidelines for Working with Cycle 8 Normalized Recored Structure Files and its Multiple Weighting Factors. This section gives a detailed description of how to work with Crime Incident Report data and how to interpret series report information.

7.4 Weighting Policy

Users are cautioned against releasing unweighted tables or performing any analysis based on unweighted survey results. As was discussed in Section 7.1, there were several weight adjustments performed independently to the records of each province. Sampling rates as well as non-response rates varied significantly from province to province.

Contact was made or attempted with 12,722 households during the survey. Of these, 1,250 (9.8%) were non-responding households. The non-responding households included 717 household refusals, 380 households that could not be reached during the survey period, 151 cases where a response could not be obtained due to language difficulties, illness, or other problems, and 2 cases where the household had already been interviewed. An interview was attempted with a adult randomly selected from the eligible household members of the 11,472 responding households. Usable responses were obtained from 10,385 respondents. The difference consists of 454 person-level refusals, 277 persons that could not be reached during the survey period, 355 cases where the interview could not be completed due to language difficulties, illness, or other problems, and 1 case where the person had already been interviewed. A response rate of 81.6% was obtained, when it is assumed that all of the households for which there was no response were "in scope" (i.e., had at least one eligible member).

It is known that non-respondents are more likely to be males and more likely to be younger. In the responding sample, 3.2% were males between the ages of 15 and 19, while in the overall population, approximately 4.3% are males between 15 and 19. Therefore, it is clear that the sample counts cannot be considered to be representative of the survey target population unless appropriate weights are applied.

7.5 Types of Estimates

Two types of 'simple' estimates are possible from the results of the General Social Survey. These are qualitative estimates (estimates of counts or proportions of people possessing certain characteristics) and quantitative estimates involving quantities or averages. More complex estimation and analyses are covered in Section 7.6.

7.5.1 Qualitative Estimates

It should be kept in mind that the target population for the GSS was non-institutionalized persons 15 years of age or over, living in the ten provinces. Qualitative estimates are estimates of the number or proportion of this target population possessing certain characteristics. The number of women who reported at least one crime incident is an example of this kind of estimate. These estimates are readily obtained by summing the final weights of the records possessing the characteristic in question.

7.5.2 Quantitative Estimates

Some variables on the 1993 General Social Survey microdata file are quantitative in nature (e.g., yearly volume of alcohol consumed - DVD5D6VL). From these variables, it is possible to obtain such estimates as the average yearly volume of alcohol consumed by people reporting any accidents. These estimates are of the following ratio form:

$$\text{Estimate (average)} = X / Y$$

The numerator (X) is a quantitative estimate of the total of the variable of interest (say, number of drinks per year) for a given sub-population (say, persons with accidents, 0 < B7 < 999). X would be calculated by multiplying the adult weight, WGHT_PER by the variable of interest when it is known, i.e. not equal to '9999', (value for nondrinker, '9997', should be made equal

to '0') and summing this product over all records with accidents. The denominator (Y) is the
qualitative estimate of the number of participants within that subpopulation (those reporting any
accidents and for which a yearly volume of alcohol was known). Y would be calculated by
summing the adult weight, WGHT_PER, over all records for persons who reported at least one
accident. The two estimates X and Y are derived independently and then divided to provide the
quantitative estimate. The average yearly number of drinks for persons with at least one accident

is then estimated to be:

$$200.2 = \frac{769,433,829}{3,843,629}$$

7.6 Guidelines for Analysis

As is detailed in Chapter 4 of this document, the respondents from the GSS do not form a simple random sample of the target population. Instead, the survey had a complex design, with stratification and multiple stages of selection, and unequal probabilities of selection of respondents. Using data from such surveys presents problems to analysts because the survey design and the selection probabilities affect the estimation and variance calculation procedures that should be used.

The GSS used a stratified design with significant differences in sampling fractions between strata. Thus, some areas are over-represented in the sample (relative to their populations) while some other areas are relatively under-represented. This means that the unweighted sample is not representative of the target population.

The survey weights must be used when producing estimates or performing analyses in order to account for this over- and under-representation. While many analysis procedures found in statistical packages allow weights to be used, the meaning or definition of the weight in these procedures often differs from that which is appropriate in a sample survey framework, with the result that while in many cases the estimates produced by the packages are correct, the variances that are calculated are almost meaningless.

For many analysis techniques (for example linear regression, logistic regression, estimation of rates and proportions, and analysis of variance), a method exists which can make the variances calculated by the standard packages more meaningful. If the weights on the data, or any subset of the data, are rescaled so that the average weight is one (1), then the variances produced by the standard packages will be more reasonable; they still will not take into account the stratification and clustering of the sample's design, but they will take into account the unequal probabilities of selection. This rescaling can be accomplished by dividing each weight by the overall average weight before the analysis is conducted.

For example, if an analysis of all respondents who were crime victims is required then the
following steps are required:

- "Select" all respondents from the file with one or more crimes reported (0 < C12 < 999
and FLAG_PER = 1)
- Calculate the Average Weight equal to the average of WGHT_PER for these records
- For each of these respondents calculate a "working" weight equal to WGHT_PER /
Average Weight;
- Perform the analysis for these respondents using the "working" weight.

The calculation of truly meaningful variance estimates requires detailed knowledge of the design
of the survey. Such detail cannot be given in this microdata file because of confidentiality.
Variances that take the sample design into account can be calculated for many statistics by
Statistics Canada on a cost recovery basis.

8. RELEASE GUIDELINES AND DATA RELIABILITY

It is important for users to become familiar with the contents of this section before publishing
or otherwise releasing any estimates derived from the General Social Survey microdata file.

This section of the documentation provides guidelines to be followed by users. With the aid
of these guidelines, users of the microdata should be able to produce figures consistent with
those produced by Statistics Canada and in conformance with the established guidelines for
rounding and release. The guidelines can be broken into four broad sections: Minimum Sample
Sizes for Estimates, Sampling Variability Policy, Sampling Variability Estimation and
Rounding Policy.

8.1 Minimum Sample Size For Estimates

Users should determine the number of records on the microdata file which contribute to the
calculation of a given estimate. When using WGHT_PER (i.e. the analysis is of adults) or
WGHT_CHD (i.e. the analysis is of children) this number should be 15 or more. If dealing with
WGHT_ACC (i.e. the analysis is of accidents) this number should be 25 or more, and if dealing
with WGHT_CRI (i.e. the analysis is of crimes) this number should be 40 or more. When the

number of contributors to the weighted estimate is less than these recommended numbers the
weighted estimate should not be released regardless of the value of the Approximate Coefficient
of Variation.

Note that the number of records used for an adult weight (WGHT_PER) estimate or a child
weight (WGHT_CHD) estimate is the number of actual respondents (i.e. adults or children) while
the number of records used for Accident or Crime incident estimates (WGHT_ACC or
WGHT_CRI) may be greater than the number of respondents who reported that type of
incident.

8.2 Sampling Variability Guidelines

The estimates derived from this survey are based on a sample of households. Somewhat different figures might have been obtained if a complete census had been taken using the same questionnaire, interviewers, supervisors, processing methods, etc. as those actually used. The difference between the estimates obtained from the sample and the results from a complete count taken under similar conditions is called the sampling error of the estimate.

Errors which are not related to sampling may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the questionnaire and errors may be introduced in the processing and tabulation of the data. These are all examples of non-sampling errors.

Over a large number of observations, randomly occurring errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Considerable time and effort was made to reduce non-sampling errors in the survey. Quality assurance measures were implemented at each step of the data collection and processing cycle to monitor the quality of the data. These measures included the use of highly skilled interviewers, extensive training of interviewers with respect to the survey procedures and questionnaire, observation of interviewers to detect problems of questionnaire design or misunderstanding of instructions, procedures to ensure that data capture errors were minimized and coding and edit quality checks to verify the processing logic.

A major source of non-sampling errors in surveys is the effect of non-response on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response. Total non-response occured because the interviewer was either unable to contact the respondent, no member of the household was able to provide the information, or the respondent refused to participate in the survey. Total non-response was handled by adjusting the weight of households who responded to the survey to compensate for those who did not respond.

In most cases, partial non-response to the survey occurred when the respondent did not understand

or misinterpreted a question, refused to answer a question, could not recall the requested
information, or could not provide proxy information.

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error,
sound statistical practice calls for researchers to provide users with some indication of the
magnitude of this sampling error.

Although the exact sampling error of the estimate, as defined above, cannot be measured from
sample results alone, it is possible to estimate a statistical measure of sampling error, the standard
error, from the sample data. Using the standard error, confidence intervals for estimates (ignoring
the effects of non-sampling error) may be obtained under the assumption that the estimates are
normally distributed about the true population value. The chances are about 68 out of 100 that
the difference between a sample estimate and the true population value would be less than one
standard error, about 95 out of 100 that the difference would be less than two standard errors, and
virtually with certainty that the differences would be less than three standard errors.

Because of the large variety of estimates that can be produced from a survey, the standard
deviation is usually expressed relative to the estimate to which it pertains. The resulting measure,
known as the coefficient of variation of an estimate is obtained by dividing the standard error of
the estimate by the estimate itself and is expressed as a percentage of the estimate. Before
releasing and/or publishing any estimates from the microdata file, users should determine whether
the estimate is releasable based on the guidelines shown below.

	Type of Estimate	Coefficient of Variation	Policy Statement
1.	Unqualified	0.0 to 16.5%	Estimates can be considered for general unrestricted release.
2.	Qualified	16.6 to 33.3%	Estimates can be considered for general unrestricted release but should be accompanied by a warning cautioning users of the high sampling variability associated with the estimates.
3.	Not for Release	33.4% or over	Estimates should not be released in any form under any circumstances. In such statistical tables, such estimates should

be deleted.

Note: The sampling variability policy should be applied to rounded estimates.

8.3 Estimates of Variance

Variance estimation is described separately for qualitative and quantitative estimates.

8.3.1 Sampling Variability for Qualitative Estimates

Derivation of sampling variabilities for each of the estimates which could be generated from the survey would be an extremely costly procedure, and for most users, an unnecessary one. Consequently, approximate measures of sampling variability, in the form of tables, have been developed for use and are included in Appendix B "Approximate Sampling Variability Tables".

Variance tables for estimates were produced using each of the four weighting factors; the Adult Weight (WGHT_PER), Child Weight (WGHT_CHD), Accident Weight (WGHT_ACC) and Crime Weight (WGHT_CRI), at the Canada level. Corresponding tables for each province, the Atlantic Region, and the Prairie Region are available upon request. Cut-off points for regional and provincial data are provided in Appendix B, however. It should be noted that all coefficients of variation in these tables are approximate and, therefore unofficial. Estimates of actual variance for specific variables may be purchased from Statistics Canada. Use of actual variance estimates may allow users to release otherwise unreleasable estimates, i.e. estimates with coefficients of variation in the "Not for Release" range (see the policy regarding the release of the survey estimates on preceding pages).

The Approximate Variance tables have been produced using the coefficient of variation formula based on a simple random sample. Since estimates for the General Social Survey were based on a complex sample design a factor called the Design Effect was introduced into the variance formula. The Design Effect for an estimate is the actual variance for the estimate (taking into account the design that was used) divided by the variance that would result if the estimate had been derived from a simple random sample. The Design Effect used to produce the Approximate Variance Tables has been determined by first calculating Design Effects for a wide range of

characteristics and then choosing among these a conservative value which will not give a false
impression of high precision. These Design Effects are specified in the table below.

GENERAL SOCIAL SURVEY CYCLE 8
CANADA LEVEL DESIGN EFFECTS

Weighting Factor	Design Effect
WGHT_PER	1.48
WGHT_CHD	1.62
WGHT_ACC*	2.38
WGHT_CRI*	4.25

* Design Effects given here for WGHT_ACC and WGHT_CRI are calculated for a series
factor of 3. See Appendix A entitled Guidelines for working with Cycle 8 Normalized
Record Structure Files and Its Multiple Weighting Factors. This section gives a detailed
description of how to work with Crime Incident Report data and Accident Report data and
how to interpret series report information.

8.3.2 Sampling Variability For Quantitative Estimates

Approximate variances for quantitative variables cannot be as conveniently summarized. As a
general rule, however, the coefficient of variation of a quantitative total will be larger than the
coefficient of variation of the corresponding qualitative estimate (e.g., the number of persons
contributing to the quantitative estimate). If the corresponding qualitative estimate is not
releasable, then the quantitative total will in general not be releasable.

8.4 Rounding

In order that estimates produced from the General Social Survey microdata file correspond to those produced by Statistics Canada, users are urged to adhere to the following guidelines regarding the rounding of such estimates. It is improper to release unrounded estimates, as they imply greater precision than actually exists.

8.4.1 Rounding Guidelines

(1) Estimates of totals in the main body of a statistical table should be rounded to the nearest thousand using the normal rounding technique (see definition in Section 8.4.2).

(2) Marginal sub-totals and totals in statistical tables are to be derived from their corresponding unrounded components and then are to be rounded themselves to the nearest thousand units using normal rounding.

(3) Averages, proportions, rates and percentages are to be computed from unrounded components and then are to be rounded themselves to one decimal using normal rounding.

(4) Sums and differences of aggregates and ratios are to be derived from corresponding unrounded components and then rounded to the nearest thousand units or the nearest one decimal using normal rounding.

(5) In instances in which, due to technical or other limitations, a different rounding technique is used, which results in estimates being released which differ from the corresponding estimates produced by Statistics Canada, users are encouraged to note the reason for such differences in the released document.

8.4.2 Normal Rounding

In normal rounding, if the first or only digit to be dropped is 0 to 4; the last digit to be retained is not changed. If the first or only digit to be dropped is 5 to 9, the last digit to be retained is raised by one. For example, the number 8499 rounded to thousands would be 8 and the number 8500 rounded to thousands would be 9.

9. ADDITIONAL INFORMATION

Additional information about this survey can be obtained from the individuals listed below. Data from the survey is available through published reports, special request tabulations, and this micro data file. The microdata file is available from the Housing, Family and Social Statistics Division of Statistics Canada at a cost of $1500.00. Tabulations can be obtained at a cost that will reflect the resources required to produce the tabulation.

Sample Selection Procedures, Weighting and Estimation
Dave Paton
Social Survey Methods Division
(613) 951-1467

Subject Matter, Data Collection and Data Processing
Edward Praught
Housing, Family and Social Statistics Division
(613) 951-9180

APPENDIX A

Guidelines for Working with Cycle Eight Normalized Record Structure Files and its Multiple Weighting Factors

STRUCTURE OF FILE

The normal analysis file structure of one fixed length record per respondent does not represent the most efficient data model for the analysis file for Cycle 8 of the General Social Survey. This is because the amount of data collected for a respondent depends on the number of crime and accident incident reports that have been completed.

Following from Cycle 3, it was decided to use a normalized record structure data model. This leads to a file which has more than one record for some respondents. The number of basic records per respondent is determined by the highest number of incident records (reports), either accidents or criminal victimizations, that the respondent experienced. Each basic record contains personal level information (from the Personal Risk Questionnaire GSS 8-2), space for Accident Report information (from the GSS 8-2F) and space for Crime Incident Report information (from the GSS 8-2G).

Each respondent is identified by a unique CASE_ID number that is repeated (as part of the person level information) on each of the basic records needed to contain all the reports for that respondent. There are 11,960 records on the analysis file, but only 10,385 different CASE_ID numbers, representing data from 10,385 separate respondents sampled.

Four flags have been used to describe the basic record type(s) for a given respondent (CASE_ID):

    FLAG_PER: Adult (GSS 8-2) flag
    FLAG_CHD: Child (GSS 8-2) flag
    FLAG_ACC: Accident Report (GSS 8-2F) flag
    FLAG_CRI: Crime Incident Report (GSS 8-2G) flag

Each of the flag fields has two possible "states": 0 being OFF and 1 being ON. The definitions for ON and OFF are given below.

- In the case of the person flag FLAG_PER, the ON state indicates the first occurrence of the person level information while the OFF state indicates that this is a subsequent occurrence.

       10,385 records have FLAG_PER = 1
        1,575 records have FLAG_PER = 0

Note: The person level information is contained on ALL records (FLAG_PER = 1 or 0).

This information is provided so that tables may be produced for the Accident (or Crime
Incident) report data broken down by, say, province (DVPROV) and age (DVAGEGR1) of
the respondent.

- For FLAG_CHD, the ON state indicates the randomly selceted child of the respondent less
than 15, living in the household. The OFF state indicates that the respondent did not have
any children meeting these criteria.

2,946 records have FLAG_CHD = 1
7,439 records have FLAG_CHD = 0

- For FLAG_ACC the ON state indicates valid data for the Accident Report component of
the record, while the OFF state indicates only default values are to be found in that section
of the record.

2,353 records have FLAG_ACC = 1
9,607 records have FLAG_ACC = 0

Note: Records with FLAG_ACC = 0 have only default data for that component of the record.

- FLAG_CRI is similar to FLAG_ACC, covering Crime Incident Report data.

3,740 records have FLAG_CRI = 1
8,220 records have FLAG_CRI = 0

Note: Records with FLAG_CRI = 0 have only default data for that component of the record.

"SELECT" will be used throughout the documentation to indicate a computer operation.
The operation is defined as creating a sub-file containing only those records which match on
the SELECT criteria. To create, as an example, a working file of all Accident reports simply
SELECT all records with FLAG_ACC = 1.

There are seven possible basic record types, based on combinations of the values of the three
flags (to simplify discussion, FLAG_CHD will not be considered as it is equal to one only when
FLAG_PER = 1). Below is a table giving each possible record type and the number of records
on the file with that type:

Basic Record Types (FLAG Combinations)

FLAG_PER	FLAC_ACC	FLAG_CRI	Number of records for each basic record type combination
1	0	0	6,728
1	1	0	1,185
1	0	1	1,715
1	1	1	757
0	0	1	1,164
0	1	0	307
0	1	1	104

11,960

Here are some examples of the record type combinations present for certain CASE_IDs on the
actual file.

Examples of Basic Records for Respondents

CASE_ID	FLAG_PER	FLAG_ACC	FLAG_CRI	Number of basic records for this CASE_ID (respondent)
12020003				
	1			
	0			
	0			
	1			
12020012				
12020012				
	1			
	0			
	0			
	0			
	1			
	1			
	2			
12020086				
12020086				
12020086				
	1			
	0			
	0			
	1			
	1			
	0			
	1			
	1			

1

3

12020122
12020122
12020122
1
0
0
1
1
1
0
0
0

3

CASE_ID 12020003 is the typical case of a respondent reporting zero Accident (GSS8-2F)
reports and zero Crime Incident (GSS 8-2G) Reports.

CASE_ID 12020012 is a respondent having two Crime Incident Reports (GSS 8-2G) and zero
Accident Reports (GSS 8-2F).

CASE_ID 12020086 has two Accident reports (GSS 8-2F) and three Crime Incident reports
(GSS8-2G).

CASE_ID 12020122 has three Accident reports (GSS 8-2F) and zero Crime Incident reports (GSS
8-2G).

CASE_ID 's 12020003, 12020012, 12020086, and 12020122 (as will all groups) have only one
basic record with FLAG_PER = 1.

NOTES:
1) These cases are given as examples only, other combinations of numbers of Accident reports
(GSS8-2F) and numbers of Crime Incident reports (GSS 8-2G) for a respondent exist.

2) All possible basic record types (combinations of FLAG_PER, FLAG_ACC and FLAG_CRI)
are given in the examples.

3) Each REPORT (Accident (GSS 8-2F) or Crime (GSS 8-2G)) may represent a number of

INCIDENTS.

DESCRIPTION OF WEIGHTING FACTORS

There are 4 weights on the file which can be used to produce several types of estimates.

1. Adult Weight: WGHT_PER

This is the weighting factor used to obtain an estimate of Canadians aged 15 and over having
various demographic characteristics.

SELECT: Records with FLAG_PER = 1

10,385

WGHT_PER = 21,644,240
= an estimate of the number of persons 15 years of age and older in the population.

Household Weight: WGHT_PER/DVELLIGC

No household weight is provided on the file, however, an approximation of a household
weight can be made by dividing the person weight by the total number of household members
15 years of age and older (DVELLIGC).

SELECT: Records with FLAG_PER = 1

10,3856

(WGHT_PER/DVELLIGC) = 10,897,374
= an estimate of the number of households in the target population (Note: using the uncapped value of DVELLIGC, the estimated number of households is 10,883,983)

2. Child Weight: WGHT_CHD

This is the weighting factor used to obtain an estimate of children 14 years of age and younger
having various demographic and accident characteristics.

SELECT: Records with FLAG_CHD = 1 (NOTE: FLAG_PER will always be equal 1 when FLAG_CHD equals 1)

2,946

WGHT_CHD = 5,773,277
= an estimate of the number of children 14 years of age and younger in the population.

3. Accident Weight: WGHT_ACC

This is the weight assigned to the Accident Report (GSS 8-2F). The total number of accidents is calculated by accumulating the accident incident weight (WGHT_ACC) multiplied by the number of incidents the report represents (DVNUMACC) over all accident reports.

SELECT: Records with FLAG_ACC = 1

2,353

(WGHT_ACC * DVNUMACC) = 4,855,856
= an estimate of the total number of accident incidents reported

The total number of accidents can also be determined by accumulating WGHT_PER * B7 over all respondents who reported at least one accident incident:

SELECT: Records with FLAG_PER = 1 and 0 < B7 < 999

1,942

(WGHT_PER * B7) = 4,855,856

The estimated number of people who reported at least one accident is determined by
accumulating WGHT_PER for those respondents with 0 < B7 < 999.

SELECT: Records with FLAG_PER = 1 and 0 < B7 < 999

1,942

WGHT_PER = 3,931,425
= an estimate of the total population with at least one accident incident

Thus 4,855,856 accident incidents were reported by 3,931,425 people.

Accident Series Weight

In the incident estimations above, series incident reports were assumed to represent the actual
number of incidents (DVNUMACC) reported by the respondent. In some cases the value of
DVNUMACC is high, contributing a large number of accidents to the total. Interpretation of
estimates made using the actual series value should be made with caution.

However, if each series incident is only counted as one then an underestimate of the total number
of incidents is very likely. One alternative to this dilemma is to count series incidents as a
maximum of three (one of the necessary conditions before an incident could be classified
as a series incident by the interviewer). The effect of these adjustments can be seen below.

Counting each Accident Report as 1 incident:

SELECT: Records with FLAG_ACC = 1
SET "Series Factor" = 1 for all Accident Reports

2,353

    (WGHT_ACC * "Series Factor") = 4,746,824

Counting Accident Reports representing a series as 3 incidents:

    SELECT: Records with FLAG_ACC = 1
    SET "Series Factor" = 1 for all DVNUMACC = 1
    SET "Series Factor" = 2 for all DVNUMACC = 2
    SET "Series Factor" = 3 for all DVNUMACC >= 3

2,353

    (WGHT_ACC * "Series Factor") = 4,823,008

These estimates can be compared with the estimate of 4,855,856 produced above when the
actual number of incidents is used.

4. Crime Weight: WGHT_CRI

This is the weight assigned to the Crime Incident Report (GSS 8-2G). This weight has been
adjusted for the number of people that could have theoretically reported this crime depending on
whether the victimization was personal or household related.

The following are classified as personal victimizations:

    sexual assault
    assault
    robbery or attempted robbery
    theft or attempted theft of personal property.

The following are classified as household victimizations:

    break and enter or attempted break and enter
    theft or attempted theft of a motor vehicle or part
    theft or attempted theft of household property
    vandalism.

The total number of crime incidents reported is calculated by accumulating the Crime weight
(WGHT_CRI) multiplied by the number of incidents that the report represents (DVNUMINC):

SELECT: Records with FLAG_CRI = 1

3,740

     (WGHT_CRI * DVNUMINC) = 7,727,843
                  = an estimate of the total number of crime incidents It is very important that the number of crime incidents is estimated using WGHT_CRI and

DVNUMINC. Unlike accident incidents which can be estimated using WGHT_ACC and DVNUMACC or WGHT_PER and B7 the number of crime incidents can be estimated only by using WGHT_CRI and DVNUMINC. Using WGHT_PER and C12 results in an overestimate of the number of crimes.

The magnitude of the above noted adjustments can be seen by comparing the incorrect estimate with the correct one:

SELECT: Records with FLAG_PER = 1 and 0 < C12 < 999

2,472

(WGHT_PER * C12) = 10,527,756

The incorrect value of 10,527,756 is 36% higher then the correct estimate of the total number of crimes (7,727,843). The difference is due to the adjustments made to the crime report weight.

The estimated number of people who reported at least one crime is determined by accumulating WGHT_PER for all respondents with at least one crime:

SELECT: Records with FLAG_PER = 1 and 0 < C12 < 999

2,472

WGHT_PER = 5,086,785
= an estimate of the population with at least one crime incident

Thus 5,086,785 people had 7,727,843 incidents.

Crime Series Weight

In the incident estimations above, series incident reports were assumed to represent the actual number of incidents, DVNUMINC, reported by the respondent. The value of DVNUMINC is very high in some cases and thus these cases can contribute a large number of incidents to the total. Interpretation of estimates made using the actual series value should be made with caution.

At the other extreme, if series incidents are only counted as one an underestimate of the total number of incidents is very likely. One alternative to this dilemma is to count series incidents as three incidents (by definition a series is three or more crime incidents of the same type which the respondent cannot distinguish). The effect of these adjustments can be seen below.

SELECT: Records with FLAG_CRI = 1
"Series Factor" = 1 for all Crime reports

3,740

(WGHT_CRI * "Series Factor") = 5,077,128

SELECT: Records with FLAG_CRI = 1
SET "Series Factor" = 1 for all Crime reports with DVNUMINC = 1
SET "Series Factor" = 2 for all Crime reports with DVNUMINC = 2
SET "Series Factor" = 3 for all Crime reports with DVNUMINC >= 3

3,740

(WGHT_CRI * "Series Factor") = 5,556,911

These estimates can be compared with the estimate of 7,727,843 produced above when the
actual number of incidents is used.

METHODS OF ESTIMATION AND INTERPRETATION OF ESTIMATES

I: When estimates of the Number of Persons are desired:

SELECT: Records with FLAG_PER = 1
Weight to be used: WGHT_PER

Examples & Interpretation:

(i) Fifty-seven percent of adult Canadians (12.4 million) think their neighbourhood has
a lower amount of crime in comparison to other areas in Canada (A3=3).

(ii) Eighteen percent (3.9 million) of the adult population (DVAGEGR1 >= 01)
were involved in at least one accident in 1992/93 (0<B7<999).

(iii) Twenty-four percent of adult Canadians (5.1 million) were victims of one or more
crimes in 1992/93 (0 < C12 < 999). Thirty-seven percent (1.4 million) of those aged
15-24 (01 <= DVAGEGR1 <= 03) were victims of one or more crimes in 1992/93.

(iv) Assault was the most serious victimization experienced by 450,833 youth aged 15-24
(DVMSVIC=104 and 01 <= DVAGEGR1 <= 03).

Cautions and Restrictions:

When making estimates of numbers of people, one must use person-level variables,

i.e. variables that describe the person rather than the incident. These person level variables may be derived from the incident reports, for the person (for example, DVMSVIC). When doing so, care must be taken in identifying what incident the analyst wishes to describe - the most serious incident, the most recent incident, etc. The person weight must be used when dealing with person level variables.

II: When estimates of the number of households are desired:

SELECT: Records with FLAG_PER =1
Weight to be used : (WGHT_PER/DVELLIGC)

Example & Interpretation:

(i) Nearly twenty-three percent of households experienced at least one victimization (0 <= C12 <= 999) in 1992/93 (i.e. either a personal victimization of an adult household member or a household victimization).

Cautions & Restrictions:

Relevancy of variable as a variable describing the household. Households must be distinguished from dwelling units which may contain one or more households.

III: When estimates of the number of children are desired:

SELECT: Records with FLAG_CHD=1
Weight to be used: WGHT_CHD

Example and Interpretation:

(i) Ten percent of children 14 years of age and younger (559,167) had at least one accident in 1992/93 (B15 = 1). Fifty-eight percent (321,700) of children who had an accident were male (DVCHDSEX = 1).

IV: When estimates of the number of accident incidents are desired:

SELECT: Records with FLAG_ACC = 1
Weight to be used : (WGHT_ACC * "Series Factor")

Note: The choice of "Series Factor" will be based on DVNUMACC and will depend on the number of incidents that the analyst wishes each series report to represent. See Description of Weighting Factors.

Examples & Interpretation:

(i) 4.8 million accidents were reported by 3.9 million Canadians (Note that 3.9 million is the estimate of the number of people reporting accidents, from Section I above).

Of the 4.8 million accidents reported, 1.6 million incidents (32%) were reported by those aged 15-24 (01 <= DVAGEGR1 <= 03).

(ii) The majority of these accidents (1.3 million, 27%) occurred in the summer months (F2 = 06 (June), 07 (July), or 08 (August)).

(iii) There were a total 684 thousand falls (ACC_SCRQ = 'B5A') in 1992/93.

Cautions and Restrictions:

(1) When making estimates of numbers of incidents, one must use incident-level variables, i.e. variables that describe the incident rather than the person. The incident weight must be used when working with incident level variables. This may necessitate deriving incident level variables from person level characteristics (eg. age of respondent reporting incident as in (i) above). No adjustment to person-level variables is necessary in these instances. As was shown earlier, this is not the case for the reverse situation.

(2) Comparisons of incident estimates with estimates based on other sources must acknowledge the potentially significant differences in definitions.

V: Estimate the number of people with certain types of accident characteristics

SELECT: Records with FLAG_ACC = 1
Weight to be used: WGHT_PER

ERROR:
Using those records with FLAG_ACC = 1 means that you are working with variables that are associated with accident incidents. The GSS 8-2 variables are only included on the records with FLAG_ACC = 1 and FLAG_PER = 0 to facilitate the derivation of new incident level variables. WGHT_PER is appropriate for use only with respondent level variables.

Restrictions:

Estimates made this way have no meaningful interpretation. To combine data items from the Personal Risk Questionnaire (GSS 8-2) and Accident Incident Form (GSS 8-2F) parts of the questionnaire the analyst must either:

(1) derive a person level variable from the incident reports for the respondent (e.g. total number of major activity days loss from all accidents in 1992/93, total out-of-pocket expenses from all accidents in 1992/93, number of accidents occurring in the home), then use the new variable with GSS 8-2 variables and WGHT_PER, or

(2) derive an incident level variable from the Personal Risk Questionnaire (GSS 8-2) variables of the respondent (e.g. age of person reporting incident, household size of person reporting incident, type of accident of most concern to person reporting incident), then using the new variable with Accident Incident variables and WGHT_ACC.

Examples of (1) can be found in Section I and of (2) in Section IV.

VI: When estimates of the number of Crime Incidents are desired:

SELECT: Records with FLAG_CRI = 1
Weight to be used : (WGHT_CRI * "Series Factor")

Note: the choice of "Series Factor" will be based on DVNUMINC and will depend on the number of incidents that the analyst wishes each series report to represent, see Description of Weighting Factors.

Examples & Interpretation:

(i) The 5.1 million Canadians, an estimate of a number of persons, therefore estimated as in I above, victimized in 1992/93 reported 5.6 million victimizations that year (an estimate of a number of incidents i.e. FLAG_CRI = 1). Of these victimizations 1.7 million incidents (30%) were reported by the 15-24 age group (01 <= DVAGEGR1 <= 03 from the Personal Risk Questionnaire).

(ii) The majority of these victimizations (1.5 million, 27%) occurred in the summer months (G2 = 06 (June), 07 (July) or 08 (August)).

(iii) There were a total of 1.4 million assaults (DVMSCRIM = 304) last year.

Cautions and Restrictions:

(1) When making estimates of numbers of incidents, one must use incident-level variables, i.e. variables that describe the incident rather than the person. The incident weight must be used when working with incident level variables. This may necessitate deriving incident level variables from person level characteristics (eg. age of respondent reporting incident as in (i) above). No adjustment to person-level variables is necessary in these instances. As was shown earlier, this is not the case for the reverse situation.

(2) Comparisons of incident estimates with estimates based on Department of Justice or police records must acknowledge whether they were reported or not and the significant differences in definitions.

VII: Estimate the Number of People with Certain Crime Incident Characteristics

SELECT: Records with FLAG_CRI = 1
Weight to be used : WGHT_PER

ERROR:
Using those records with FLAG_CRI = 1 means that you are working with variables that are associated with crime incidents. The GSS 8-2 variables are only included on the records with FLAG_CRI = 1 and FLAG_PER = 0 to facilitate the derivation of new incident level variables. WGHT_PER is appropriate for use only with respondent level variables.

Restrictions:

Estimates made this way have no meaningful interpretation. To combine data items from the Personal Risk Questionnaire (GSS 8-2) and the Crime Incident Report (GSS 8-2G) parts of the questionnaire the analyst must either:

(1) derive a person level variable from the incident reports for the respondent (e.g. total value of items stolen in year, number of times assaulted, number of victimizations while away from home), then use the new variable with

Personal Risk Screening variables (i.e. person level variables) and WGHT_PER,
or

(2) derive an incident level variable from the Personal Risk Questionnaire
variables of the respondent (e.g. age of person reporting incident, household
size of person reporting incident, type of crime of most concern to person
reporting incident), then using the new variable with Crime Incident Report
variables (GSS 8-2G) and WGHT_CRI.

Examples of (1) can be found in Section I and of (2) in Section VI.

APPENDIX B

Approximate Variance Tables

APPROXIMATE VARIANCE TABLES

By using the Approximate Variance Tables and the following rules users should be able to determine approximate coefficients of variation for aggregates (totals), percentages, ratios, differences between totals and differences between ratios.

There is one table for each type of estimate produced:
1) Estimates using the Adult Weight (WGHT_PER),
2) Estimates using the Child Weight (WGHT_CHD),
3) Estimates using the Accident Weight (WGHT_ACC) and
4) Estimates using the Crime Weight (WGHT_CRI).

Each table is for Canada level estimates. Corresponding tables for each of provinces, the Atlantic region (Newfoundland, Prince Edward Island, Nova Scotia, and New Brunswick), and the Prairie region (Manitoba, Saskatchewan, and Alberta) are available upon request. The table below contains cutoff values, or minimum releaseable estimates for Canada, provincial and regional estimates. These values are the population estimates that give a coefficient of variation of 33.3% or less. Population estimates of this size or larger are releaseable provided they also contain the minimum number of contributors, as given in Section 8.1.

Users should ensure that the Approximate Variance Table or cutoff table used corresponds with the type of estimate being considered (Chapter 7 and Appendix A contain information on the correct production of these types of estimates). The tables for estimates using the Accident Weight (WGHT_ACC) and the Crime Weight (WGHT_CRI) are only valid for estimates produced using a series factor of 3.

GENERAL SOCIAL SURVEY CYCLE 8
MINIMUM RELEASEABLE POPULATION/INCIDENT ESTIMATES (000s)

Geographic Area	WGHT_PER	WGHT_CHD	WGHT_ACC*	WGHT_CRI*
Canada	30	30	45	60
Atlantic Region	10	9	21	18
Newfoundland	9	8	19	18
Prince Edward Island	5	3	9	11
Nova Scotia	11	9	24	18
New Brunswick	10	9	23	25
Quebec	30	35		

50
60

Ontario
40
45
60
65

Prairie Region
15
16
30
22

Manitoba
12
13
21
20

Saskatchewan
10
11
18
16

Alberta
18
17
35
25

British Columbia
22
23
40
40

* Minimum releaseable estimates given here for WGHT_ACC and WGHT_CRI are calculated for a series factor of 3. See Appendix A titled Guidelines for working with Cycle 8 Normalized Files and Its Multiple Weighting Factors. This section gives a detailed description of how to work with Crime Incident Report data and Accident Report data and how to interpret series report information.

Rules for Obtaining Approximate Variances

The following rules should enable the user to determine the approximate coefficients of variation from the Approximate Variance Tables for estimates of the number, proportion or percentage of the surveyed population possessing a certain characteristic and for ratios and differences between estimates.

As noted in Section 8.1, all estimates should contain at least the minimum number of contributors in order to be released, regardless of the Approximate Coefficient of Variation.

Rule 1: Estimates of Numbers Possessing a Characteristic (Aggregates)

The coefficient of variation (cv) depends only on the size of the estimated aggregate itself.

In the appropriate Approximate Variance Table, locate the estimated aggregate in the left-most column of the table (headed "Numerator of Percentage") and follow the asterisks across to the first figure encountered. This figure is the estimated coefficient of variation.

Example 1:

A user estimates that in Canada 2,118,710 females aged 15 years and over feel very unsafe walking alone in their area after dark (question A6). How does the user determine the approximate coefficient of variation for this estimate?

Refer to the approximate variance table for Canada level estimates produced using the adult weight (WGHT_PER). The estimated aggregate does not appear in the left-most column (the 'Numerator of Percentage' column), so it is necessary to use the closest figure, namely 2,000,000. The coefficient of variation for an estimated aggregate is found by referring to the first non-asterisk entry for that row, in this case 3.7%. This cv falls within the range of cv's for 'Unqualified' estimates (i.e. 0.0% - 16.5%, pg. 22) allowing the estimate to be released without restriction.

Rule 2: Estimates of Percentages or Proportions Possessing a Characteristic

The coefficient of variation of an estimated percentage or proportion depends on both the size of the percentage or proportion and the size of the total upon which the percentage is

based. Estimated percentages or proportions are relatively more reliable than the corresponding estimates of the numerators of the percentages, particularly if the percentages are 50 percent or more. (Note that in the tables the cv's decline in value reading from left to right).

When the percentage or proportion is based upon the total population of the geographic area covered by the table, the cv of the percentage or proportion is the same as the cv of the numerator of the percentage. In this case, Rule 1 can be used.

When the percentage or proportion is based upon a subset of the total population (e.g., those in a particular age-sex group), reference should be made to the percentage (across the top of the table) and to the numerator of the percentage or proportion (down the left side of the table). The intersection of the appropriate row and column gives the coefficient of variation.

Example 2:

A user estimates that in Canada 19.2% of females aged 15 years and over feel very unsafe walking alone in their area after dark (question A6). This is the expression of the estimate obtained in Example 1 as a percentage of all females aged 15 years and over in Canada.
How does the user determine the approximate coefficient of variation for this estimate?

Refer to the approximate variance table for Canada level estimates produced using the adult weight (WGHT_PER). Because the estimate is a percentage which is based on a subset of the population covered by the table, it is necessary to use both the percentage (19.2%) and the numerator portion of the percentage (2,118,710) to determine the approximate coefficient of variation. Since the numerator does not appear in the left-most column (the 'Numerator of Percentage'column), it is necessary to use the figure closest to it, namely 2,000,000. Similarly, the percentage estimate does not appear among the column headings, so it is necessary to use the figure closest to it, namely 20.0%. The figure at the intersection of the row and column selected, namely 3.5%, is the coefficient of variation. This cv falls within the range of cv's for 'Unqualified' estimates (i.e. 0.0% - 16.5%, pg. 22) allowing the estimate to be released without restriction.

Rule 3: Ratios

In the case where the numerator is a subset of the denominator, the ratio should be converted to a percentage and Rule 2 applied. This would apply, for example, to the case where the denominator is the number of males and the numerator is the number of males who feel that crime has increased in their neighborhood during the past five years.

In the case where the numerator is not a subset of the denominator, the coefficient of variation of the ratio of two estimates is approximately equal to the square root of the sum of squares of each coefficient of variation considered separately. That is, the standard deviation of a ratio

R = X / Y
is
sd(R) = R * (cv(X)2 + cv(Y)2)1/2

The coefficient of variation of R is approximately:

cv(R) = sd(R) / R
= (cv(X)2 + cv(Y)2)1/2

This formula will tend to overstate the error if X and Y are positively correlated and understate the error if X and Y are negatively correlated.

Example 3:

A user estimates that in Canada, among females aged 15 years and over, 2,118,710 feel very unsafe walking alone in their neighborhood after dark (question A6) and among males aged 15 and over 404,796 feel very unsafe walking alone in their neighborhood after dark. The user is interested in the ratio of females who feel very unsafe versus males who feel very unsafe. How does the user determine the approximate coefficient of variation for this ratio estimate?

The numerator of the ratio estimate is 2,118,710 (X). Using Rule 1 (refer to Example 1), the coefficient of variation for this estimate is determined to be 3.7% cv(X). The denominator of the ratio estimate is 404,796 (Y). Again using Rule 1, the coefficient of variation is determined to be 8.7% cv(Y). Using Rule 3, the coefficient of variation of the ratio estimate is

cv = (0.0372 + 0.0872)0.5

= 0.0945

Therefore at the Canada level, the ratio of females 15 years of age and over who feel very unsafe walking alone in their area after dark versus males 15 years of age and over who feel very unsafe walking alone in their neighborhood after dark is 2,118,710/404,796 or 5.2 to 1. The coefficient of variation of this estimate is 9.5%, and so the estimate can be released without restriction.

Rule 4: Differences Between Totals or Percentages

The standard deviation of a difference between two estimates is approximately equal to the square root of the sum of squares of each standard deviation considered separately. That is, the standard deviation of a difference:

d = X - Y
is
sd(d) = ((X * cv(X))2 + (Y * cv(Y))2)1/2

The coefficient of variation of d is approximately:
cv(d) = sd(d) / d

This formula is accurate for the difference between separate and uncorrelated characteristics but is only approximate otherwise.

Example 4:

A user estimates that in Canada, among those 15 years and over, 19.2% (X) of females feel very unsafe walking alone in their area after dark and 3.8% (Y) of males feel very unsafe walking alone. The user is interested in the difference between these two estimates. How does the user determine the approximate coefficient of variation for the estimate of the difference?

From Example 2, the coefficient of variation for the female estimate is 3.5%. The coefficient of variation for the male estimate is 8.6%.

The difference between the estimates is 15.4%. Using Rule 4, the standard deviation of the difference between the estimates is

sd = ((0.192 x 0.035)2 + (0.038 x 0.086)2)0.5

$$= 0.00746$$

and the coefficient of variation is

$$cv = \frac{0.00746}{0.154}$$

$$= 0.048$$

Therefore the coefficient of the difference between the estimates is 4.8% and so the estimate can be released without restriction.

Rule 5: Difference of Ratios

In this case, Rules 3 and 4 are combined. The cv's of the two ratios are first determined using Rule 3, and the cv of their difference is found using Rule 4.

Confidence Limits

Although coefficients of variation are widely used, a more intuitively meaningful measure of sampling error is the confidence interval of an estimate. A confidence interval constitutes a statement on the level of confidence that the true value for the population lies within a specified range of values. For example a 95% confidence interval can be described as follows:

If sampling of the population is repeated indefinitely, each sample leading to a new confidence interval for an estimate, then in 95% of the samples the interval will cover the true population value.

Using the standard error of an estimate, confidence intervals for estimates may be obtained under the assumption that under repeated sampling of the population, the various estimates obtained for a population characteristic are normally distributed about the true population value. Under this assumption, the chances are about 68 out of 100 that the difference between a sample estimate and the true population value would be less than one standard error, about 95 out of 100 that the difference would be less than two standard errors, and about 99 out of 100 that the differences would be less than three standard errors. These different degrees of confidence are referred to as the confidence levels.

Confidence intervals for an estimate, are generally expressed as two numbers, one below the estimate and one above the estimate, as (-k, +k) where k is determined depending upon the level of confidence desired and the sampling error of the estimate.

Confidence intervals for an estimate can be calculated directly from the Sampling Variability Tables by first determining from the appropriate table the coefficient of variation of the estimate, and then using the following formula to convert to a confidence interval CI:

CIx = { - (t)()( ), + (t)()( )}

where is the determined coefficient of variation of
t = 1 if a 68% confidence interval is desired
t = 1.6 if a 90% confidence interval is desired
t = 2 if a 95% confidence interval is desired
t = 3 if a 99% confidence interval is desired

Example 5(a):

An estimated 2,118,710 females aged 15 years of age and over feel very unsafe walking alone in their area after dark. This estimate has an approximate coefficient of variation of 3.7% The 95% confidence interval for this estimate is thus:

{2,118,710 - (2)(2,118,710)(0.037), 2,118,710 + (2)(2,118,710)(0.037)}

{2,118,710 - 156,785, 2,118,710 + 156,785}
{1,961,925, 2,275,495}

With 95% confidence it can be said that between 1,961,925 and 2,275,495 of females aged 15 years and over in Canada, feel very unsafe walking alone in their area after dark.

Example 5(b):

An estimated 19.2% of females aged 15 years and over in Canada feel very unsafe walking alone in their area after dark or 0.192 expressed as a proportion. From Example 2 this estimate has an approximate coefficient of variation of 3.5%. A 95% confidence interval for this estimate (expressed as a proportion) is

CI = {0.192 - (2)(0.192)(0.035), 0.192 + (2)(0.192)(0.035)}
= {0.179,0.205}

With 95% confidence it can be said that between 17.9% and 20.5% of females aged 15 years and over in Canada, feel very unsafe walking alone in their area after dark.

Note: Release guidelines which apply to the estimate also apply to the confidence interval. For example, if the estimate is not releasable, then the confidence interval is not releasable either.

T-test

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The sample estimates can be numbers, averages, percentages, ratios, etc. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical.

Let X1 and X2 be sample estimates for 2 characteristics of interest. Let the standard error of the difference X1 - X2 be

$$\text{If } t = \frac{X1 - X2}{\text{--------}} \text{ is between -2 and 2,}$$

then no conclusion about the difference between the characteristics is justified at the 5% level of significance. If however, this ratio is smaller than -2 or larger than +2, the observed difference is significant at the 5% level (Note: at the 1% level, values of -3 and +3 should be used, etc.).

Example 6:

A user wishes to test at the 5% level of significance the hypothesis that at the Canada level there is no difference between percentage estimates of males and females who feel very unsafe walking alone in their area after dark. From Example 4 the estimate of the standard deviation of the difference between the estimates is 0.00746.

$$\text{Hence} = \frac{0.192 - 0.038}{0.00746}$$

$$= 20.64$$

Since t = 20.64 is greater than 2, there is evidence to reject the hypothesis at the 5% significance level.

APPENDIX C

Overview of the 1993 GSS on Personal Risk, complete with Cycle 8 and Cycle 3 Questionnaires (Working
Paper #7)

PREFACE

This report provides a history of the development of the 1993 General Social Survey (Cycle 8) from questionnaire design to data collection. Specifically, it examines changes and omissions from the 1988 GSS (Cycle 3) questionnaire and the reasons for these changes. In addition, it provides a summary of question comparability with Cycle 3, the 1993 Violence Against Women Survey (Statistics Canada), and the 1989 National Alcohol and Other Drugs Survey (Health and Welfare).

TABLE OF CONTENTS

BACKGROUND

The General Social Survey (GSS), conducted annually, provides policy makers and analysts with national statistical information on the living conditions and well being of Canadians. The GSS has two main objectives: first, to monitor changes in Canadian society over time, and second, to provide information on specific policy issues of a current or emerging nature.

The GSS operates on a five year repeating cycle, each year examining one of five subject areas which form the core content of the survey. Health, time use, personal risk, education and work, and the family are the five core content subjects which are rotated during each five year cycle. Core sections (see Figure 1) provide a data source for monitoring trends in Canadian society.

Focus content of the survey obtains statistical information on specific areas which are of particular interest to federal or provincial departments or other user groups. For example, focus content for Cycles 1, 2 and 3 examined social support, language and victim's services, respectively (see Figure 1).

A third and final component of each General Social Survey consists of classification which provides the means of delineating population groups and is used in the analysis of the core and focus data. Examples of classification variables include age, education and income (see Section E of 1993 questionnaire).

While resources for classification and core are included as part of the standard GSS budget, costs associated with focus content are recovered from sponsors.

The sample size targeted for each General Social Survey is 10,000. However, interested users can purchase additional sample size for specific target groups or provinces to meet special policy information needs. The sample size was increased in Cycles 1, 2, 5 and 6.

Figure 1

Cycle	End of Collection	Core Content	Focus Issue	Additional Sample	Target of Over -sample
1	Oct 1985	Health	Social support	yes	Population age 55 and over
2	Dec 1986	Time use, social mobility	Language	yes	Bilingual areas of Canada*
3	March 1988	Personal risk	Victim services	no	
4	March 1989	Education and work	none	no	
5	April 1990	Family and friends			

none					
yes					
Ontario and population age 65 and over					
6					
Dec 1991					
Health					
Various topics**					
yes					
Population age 65 and over					
7					
Dec 1992					
Time use					
Culture, sport & unpaid work activities					
no					
8					
Dec 1993					
Personal risk					
Alcohol and drug use					
no					

* North and East New Brunswick; Montreal, Quebec; Eastern Townships, Quebec; Outaouais area, Quebec and Ontario; East and North East Ontario

** Flu shots, job benefits, sources of pension/disability income, health status indicators and emotional well-being.

B. Personal Risk Survey Rationale

The eighth cycle of the GSS marks the first repeat of the GSS core subject on Personal Risk related to criminal victimization and accidents (see GSS 1988 - Cycle 3). One important application of data on personal risk is to measure the incidence of accidents and criminal victimization in order to complement the officially reported data for these incidents as derived from administrative sources.

Hospital morbidity data, workers' compensation data, workers' accident and injury data and vital statistics are all potential sources of data relating to accidents in Canada. However, many accidents are not reported by these programs and information can only be obtained directly from accident victims as reported in a general population survey. The data from Cycle 8 of the GSS complement existing statistics and provide insight into the social and demographic characteristics of Canadians who have experienced an accident in the home, at work, during recreation, or on the road during the past 12 months. Periodic collection of information concerning personal risk of accidents provides important information for policy makers, legislators, academics and other persons involved with prevention programs.

Similarly, official crime statistics have long been available from such sources as the police, courts and prisons. However, population surveys like GSS Cycle 8, which focus on the victimization experience, provide an important complement to officially recorded crime rates as they measure both those offenses which are reported to the police and those which are not. This survey also taps the victim's experience of crime, revealing the consequences of crime for victims, the reason victims decide to report offenses to the police, as well as Canadian's perceptions of the level of crime around them. Periodic collection of information concerning personal risk of criminal victimization provides important information for criminal justice planners, policy makers, legislators and medical practitioners.

C. 1993 GENERAL SOCIAL SURVEY

1. GENERAL

The core content for the 1993 GSS, personal risk, is based primarily on the core variables used in the 1988 GSS Cycle 3. The following subjects are common to both:

perception of crime
perception of police
perception of courts

crime prevention precautions
frequency of evening activities
accident screening section
criminal victimization screening section
accident report
crime incident report

As well, measures of fear of crime and childhood injuries were identified as appropriate additions to the core content.

Focus content for Cycle 8 covers alcohol and drug use. In addition to being related to some of the core content, these data will provide an update to information first collected in the 1989 National Alcohol and Other Drugs Survey. This focus content is being sponsored by the Health Promotion Directorate of Health and Welfare Canada. It should be noted that focus content for the 1988 survey, knowledge and perceptions of various criminal justice services (police, courts, assistance programs), is not part of the 1993 survey.

The target population of the 1993 General Social Survey consists of all persons aged 15 and over living in the 10 provinces of Canada, with the exception of full-time residents of institutions. The population will be sampled by random digit dialling (RDD) techniques.

Data for this Cycle will be collected monthly, over the period February 1993 to December 1993 inclusive. Collection will take place from four regional offices - Halifax, Montreal, Winnipeg, and Vancouver. The target sample size for Cycle 8 is 10,000, with an expected response rate of approximately 80%.

The comparable 1988 survey collected data from 9,870 persons during the period January 22nd to February 27th, 1988. Collection took place from Statistics Canada's eight regional offices. A response rate of 82.4% was achieved.

2. COMPUTER ASSISTED TELEPHONE INTERVIEWING (CATI)

Cycle 8 is the first time GSS will collect data using Computer Assisted Telephone Interviewing (CATI). With CATI, the survey questions appear on a computer monitor. The interviewer asks the respondent the questions, then enters the responses into the computer as the interview progresses. Built-in edits and fewer processing steps result in more timely and better quality data. CATI methodology also eliminates the need for paper and pencil questionnaires, as in previous GSS cycles. These forms were produced, however, as reference documents (see Appendix C). In Cycle 8, the CATI system provides the interviewer with four main "components" which can be imagined to represent four paper questionnaires.

QUESTIONNAIRE	AGE GROUP	TITLE
GSS 8-1	All age groups	Survey Control Form
GSS 8-2	Age 15 and over	Personal Risk Questionnaire
GSS 8-2F	Age 15 and over	Accident Report
GSS 8-2G	Age 15 and over	Crime Incident Report

It should be noted that there are slight differences between the appended forms and the CATI version of the survey. Specifically, random selection of the respondent is now done by the computer rather than taken from a pre-printed Selection Grid Label. Furthermore, the CATI version asks the respondent for information regarding the relationship of each household member to the selected respondent, while in previous cycles and in the version appended, relationship to a designated reference person for each household economic family was collected (reference Z9 of GSS 8-1). Other differences involve items which appear on the forms but do not appear on the CATI version. For example, interviewer check items are visible on the questionnaire but exist only as internal edits on the CATI system. Similarly, skip patterns are visible on the questionnaire but exist internally on the CATI system. Lastly, a few questions, such as date of birth and language first spoken in childhood, are asked in a different manner using CATI (e.g. instead of asking date of birth, CATI asks three separate questions - year of birth, month of birth and day of birth).

4. SURVEY CONTROL FORM

The GSS 8-1 is completed for each telephone number generated in the sample. When a private household is contacted, all members of the household are enumerated and basic demographic information (age, sex, and

marital status) is collected for each household member. A respondent, aged 15 years of age or older, is then randomly selected by a CATI algorithm. The relationship of each household member to the selected respondent is determined and the main questionnaire (GSS 8-2) is completed for the respondent. Proxy interviews are not accepted.

5. PERSONAL RISK QUESTIONNAIRE

The main component of the survey (GSS 8-2) collects information on the following topics: the respondent's attitudes towards various components of the judicial system, satisfaction with various aspects of the judicial system, perception of risk with regard to accidents and crime incidents (Section A); information on the kind and number of times the respondent had been involved in an accident (Section B) or a crime incident (Section C) during the past 12 months, information on alcohol and drug use (Section D) and basic background information on the respondent (Section E). Based on pilot test data, the sequence in which the questions were asked for GSS 8 (1993) was altered. The sequence of questions for GSS 8 follows the pattern outlined in Figure 2 (Section A, B, C, F, G, D, E and H) not that outlined by the questionnaire presented in Appendix C.

The 1993 survey content (see Appendix C) was designed to be comparable, where possible, with the 1988 GSS (see Appendix B) and to a lesser extent with other GSS cycles. Appendix A presents a question by question summary of comparability.

The following sections describe the contents of the personal risk questionnaire and indicate changes that have been made in comparison to the 1988 GSS. Figures 2 and 3 illustrate the flow pattern of both questionnaires.

A. PERCEPTIONS, HISTORY AND RISK

This section collects information on a variety of issues dealing with perception of crime and accidents, contact and satisfaction with various aspects of the criminal justice system, crime prevention measures and frequency of evening activities, all of which were repeated from the 1988 GSS. A new set of questions on "fear of crime" has been incorporated into the 1993 survey. This was accomplished by transferring a subset of questions (A9 -A14) on "fear of crime" from the Violence Against Women Survey also scheduled to be conducted in 1993. As well, questions concerning contact with police (A19a - A19e), precautions taken by the respondent to avoid being a victim of crime(s) (A20A, A22a-e) or involved in an accident(s) (A24a-d, A26) were included.

Furthermore, questions regarding the most serious crime the respondent has been a victim of (A21) and satisfaction level with personal safety (A23) were incorporated.

B. ACCIDENT SCREENING

This section collects information on the kind and number of times the respondent has been involved in an accident over the past 12 months. Each time an accident is reported in this section an Accident Report (Section F) is completed. Questions in this section are, for the most part, a repeat of the 1988 GSS questions however, a few changes were made. Instead of referring to the past year (e.g. 1987), the 1993 GSS asks respondents to refer to the previous 12 months. This change was made to simplify the reference period concept for respondents now that data is collected monthly throughout the year.

Furthermore, the definition of "accident" has been broadened from the narrower focus of Cycle 3 to include "required medical attention from a doctor or nurse" as another criteria which will screen a respondent into Section F (Accident Report) of the questionnaire. The definition was broadened as a safety measure to ensure that accidents, which are of significance to the health care system, are not somehow being missed by the Cycle 8 screening criteria. Additionally, extending the definition of what constitutes an accident may act as a trigger for memory and prompt respondents to remember an accident previously forgotten. Respondents having children less then 15 years of age and living in the same household were also asked a module of questions (B11-B25) regarding accidents/injuries suffered by one of their (randomly chosen) children. This was added to compensate for the current information deficiency on childhood injuries.

C. CRIMINAL VICTIMIZATION SCREENING

This section collects information on the type and number of times the respondent has been a victim of crime over the past 12 months. Each time a crime is reported in this section a Crime Incident Report (Section G) is completed. Questions in this section are, for the most part, a repeat of the 1988 GSS questions, however, a few changes have been made. Instead of referring to the past year (e.g. 1987), the 1993 GSS asks respondents to refer to the previous 12 months, again because of monthly data collection. Additionally, two questions (C9 and C10) regarding sexual assault have been taken from the Violence Against Women survey. The intent of these questions is to allow comparison with the Violence Against Women Survey, to provide a baseline for future trend analysis of change over time, and to collect information on men's experience with sexual assault.

D. ALCOHOL AND DRUG USE

The questions in this section were sponsored by Health and Welfare Canada and are designed to monitor changing patterns of alcohol and drug use. The majority of questions dealing with alcohol and drug use are comparable to the National Alcohol and Other Drugs Survey (NADS) conducted by Statistics Canada on behalf of Health and Welfare Canada in 1989. This section examines patterns of alcohol use (e.g. frequency, quantity), places people drink (e.g. bars, weddings), alcohol-related problems (e.g. problems with one's own alcohol use or with other people's alcohol use), and illicit and licit drug use.

E. CLASSIFICATION

Section E provides background characteristics of respondents regardless of whether they have been a victim of crime(s) or involved in an accident(s). Most classification questions have evolved with each cycle of the GSS. For Cycle 8, response lists were enlarged to reduce manual coding of "other specify" responses for questions with response options which were not read to the respondent (e.g. ethnic origin, language first spoken). A question was also added to determine "handedness" (E23) of the respondent, as some literature has implied some association between handedness and accident history. As well, a question concerning activity limitation of the respondent (E24) was included. Furthermore, questions dealing specifically with hours worked per week (E30) and hours usually worked (E34) were added to this section. Finally, questions concerning a spouse's main activity (E40 -E45) were also incorporated.

F. CONTACTS FOR FOLLOW-UP

Beginning with the 1990 GSS on Family and Friends, each GSS cycle includes a final section asking respondents for detailed contact information that would facilitate a follow-up of some respect in the future. Currently, the GSS is considering re-interviewing respondents to link data from other cycles or re-interviewing respondents on the same topic five years later.

A pilot follow-up test was conducted in September 1990 using data from the 1990 Cycle 5 survey on Family and Friends and the 1991 Cycle 6 survey on Health. This test was successful and indicated longitudinal follow-up studies are feasible. Any longitudinal component for the main survey is contingent on funding support.

6. ACCIDENT REPORT

An Accident Report is completed for every accident reported in Section B of the Personal Risk Questionnaire. The Accident Report collects information on characteristics of accidents (e.g. date, time, place), consequences of accidents (e.g. injuries, disability days), causes of accidents and perceptions of preventability. For the most part, the 1993 Accident Report is similar to the 1988 Accident Report with the exception of the following changes: a module of questions dealing with accidents which occurred while participating in sports, exercise or recreational activity has been added (F14-F18); and questions concerning the relationship between accidents and alcohol or drug use have also been incorporated.

7. CRIME INCIDENT REPORT

A Crime Incident Report is completed for every crime reported in Section C of the Personal Risk Questionnaire.

The Crime Incident Report collects information on characteristics of victimization incidents (e.g. date, time, place), consequences of victimization, both physical and financial, and perception of criminal justice services.
A large part of the 1993 Crime Incident Report is similar to the 1988 Crime Incident Report. However a few changes have been made. Questions concerning threats made to the respondent (G17 and G18) have been included. Inquiries as to whether the incident was related to the use of alcohol or drugs by the perpetrator or the victim (excluding victims of sexual assault) have also been added (G24a - G24c). Furthermore, the nature of police action (G57) as well as any other agencies contacted for help (G63) have been incorporated into the Crime Incident Report.

8. OPERATIONAL CHANGES

The 1985 and 1986 GSS cycles collected data in the fall starting in September. In subsequent cycles, (1988, 1989, 1990) collection started mid-January and ended mid-March. However, since some topics in the GSS are subject to seasonality (e.g. evening activities, alcohol use), the 1991, 1992 and 1993 surveys have used a monthly data collection procedure with a fraction of the interviews (approximately 960) being conducted each month.
Due to the introduction of CATI, there was a one month delay in the start-up of the 1993 GSS, mandating a collapsing of data collection into an 11 month period.

In preceding cycles, two methods of random digit dialling were used: Elimination of Non-Working Banks (ENWB) and Waksberg. Since 1991, only ENWB has been employed. Waksberg was eliminated, as ENWB is more operationally efficient and telephone companies are now able to provide lists of working banks for essentially all areas in the ten provinces.

The majority of edits which normally would have been conducted during the editing phase, after the completion of data collection, are now programmed into CATI data collection. Errors and inconsistencies are now detected at the source and are resolved immediately with the respondent's co-operation. For example, the interviewer was told at the beginning of the interview that the respondent is 25 years of age. In section E, question E1, the respondent tells the interviewer her date of birth is 13/08/1966. The CATI system at this point will send an error message indicating that previously the respondent said she was 25 years of age, but the date of birth given is 1966. The CATI system will than ask which answer requires correction, thus allowing the interviewer to immediately clarify the situation with the respondent.

In the past, errors of this nature may have gone unnoticed until the data were transmitted to head office, at which point the edits would have rejected the case because of the inconsistency. The record would then be retrieved manually and the correction made. While manual edits will continue to be performed as part of second stage editing, CATI dramatically reduces the number of manual edits which have to be performed. This is expected to result in improved data quality and increased speed in data processing.

RELATED ACTIVITIES/RECOMMENDED SOURCES

1) General Social Survey on Personal Risk (Cycle 3 - 1988):

Accidents in Canada, Catalogue 11-612E, No. 3.

National Alcohol and Other Drugs Survey: Highlights Report. Health and Welfare Canada, June 1990.

Patterns of Criminal Victimization in Canada, Catalogue 11-612E, No. 2.

Public Use Microdata File Documentation and User's Guide. Personal Risk - 1988. Statistics Canada, 1988.

2) Canadian Centre for Justice Statistics:

General:

Johnson, H. and G. Lazarus. "The Impact of Age on Crime Victimization Rates". Canadian Journal of Criminology. July, 1989.

Johnson, H. and Vince Sacco. "Victimization in Canada: Data from a National Survey" in R. Silverman, J. Teevan and V. Sacco (eds) Crime in Canadian Society (4th ed). Toronto: Butterworths, 1991:92-98.

Uniform Crime Reporting Survey. Canadian Centre for Justice Statistics.

Violence Against Women Survey, Questionnaire Package. Statistics Canada, 1993.

Canadian Urban Victimization Survey:

"Victims of Crime". Canadian Urban Victimization Survey. Bulletin 1, Ministry of Solicitor General, 1983.

"Reported and Unreported Crimes". Canadian Urban Victimization Survey. Bulletin 2, Ministry of Solicitor General, 1984.

"Crime Prevention: Awareness and Practice". Canadian Urban Victimization Survey. Bulletin 3, Ministry of Solicitor General, 1984.

"Female Victims of Crime". Canadian Urban Victimization Survey. Bulletin 4, Ministry of Solicitor General, 1985.

"Cost of Crime to Victims". Canadian Urban Victimization Survey. Bulletin 5, Ministry of Solicitor General, 1985.

"Criminal Victimization of Elderly Canadians". Canadian Urban Victimization Survey.
Bulletin 6, Ministry of Solicitor General, 1986.

"Household Property Crimes". Canadian Urban Victimization Survey. Bulletin 7, Ministry
of Solicitor General, 1986.

"Patterns in Violent Crime". Canadian Urban Victimization Survey. Bulletin 8, Ministry of
Solicitor General, 1987.

"Patterns in Property Crime". Canadian Urban Victimization Survey. Bulletin 9, Ministry
of Solicitor General, 1988.

"Multiple Victimization". Canadian Urban Victimization Survey. Bulletin 10, Ministry of
Solicitor General, 1988.

Juristat:

"Public Perceptions of Crime and the Criminal Justice System". January 1991, Vol. 11,
No. 1. Juristat Service Bulletin. Statistics Canada/Canadian Centre for Justice Statistics.

"Conjugal Violence Against Women". May 1990, Vol. 10 No. 7. Juristat Service Bulletin.
Statistics Canada/Canadian Centre for Justice Statistics.

"Criminal Victimization in Canada: The Findings of a Survey". October 1990, Vol. 10,
No. 16. Juristat Service Bulletin. Statistics Canada/Canadian Centre for Justice Statistics.

"Violent Crime in Canada". October 1990, Vol. 10 No. 15. Juristat Service Bulletin.
Statistics Canada/Canadian Centre for Justice Statistics.

3) Canadian Social Trends:

Johnson, Holly and Vincent Sacco. "Perceptions of the Criminal Justice System". Canadian Social Trends. Winter 1990:8-11. Statistics Canada.

Johnson, Holly. "Wife Abuse". Canadian Social Trends. Spring 1988:17-20. Statistics Canada.

Millar, Wayne and Owen Adams. "Accidents". Canadian Social Trends. Summer 1990:22-24. Statistics Canada.

Sacco, Vincent and Holly Johnson. "Household Property Crime". Canadian Social Trends. Spring 1990:15-17. Statistics Canada.

Sacco, Vincent and Holly Johnson. "Violent Victimization". Canadian Social Trends.
Summer 1990:10-13. Statistics Canada.

PLANS FOR DATA RELEASE:

Output	Release Date	Cost
Questionnaire Package	February 1993	No cost
Working Paper #7: Overview of the 1993 General Social Survey on Personal Risk (GSS-8)	September 1993	No cost
Initial data release	April 1994	No cost
Public Use Microdata File	July/August 1994	$1500.00
Cycle 8 Publication	1995	To Be Determined

CONTACTS

Cycles 1, 3, 6 and 8: (GSS 1985, 1988, 1991, 1993)

Edward Praught, Survey Cycle Manager,
General Social Survey,
Housing, Family and Social Statistics Division,
Statistics Canada,
7-D7, Jean Talon Bldg., Tunney's Pasture,
Ottawa, Ontario
K1A 0T6
Telephone: (613) 951-9180

Cycles 2, 4 and 7: (GSS 1986, 1989 and 1992)

Ghislaine Villeneuve, Survey Cycle Manager,
General Social Survey,
Housing, Family and Social Statistics Division,
Statistics Canada,
7-D7, Jean Talon Bldg., Tunney's Pasture,
Ottawa, Ontario
K1A 0T6
Telephone: (613) 951-4995

Cycle 5: (GSS 1990) and Violence Against Women Survey (1993)

Josephine Stanic, Survey Cycle Manager,
General Social Survey,
Housing, Family and Social Statistics Division,
Statistics Canada,
7-D7, Jean Talon Bldg., Tunney's Pasture,
Ottawa, Ontario
K1A 0T6
Telephone: (613) 951-8644

Cycle 9: (GSS 1994)

Pierre Turcotte, Survey Cycle Manager,
General Social Survey,
Housing, Family and Social Statistics Division,
Statistics Canada,
7-D7, Jean Talon Bldg., Tunney's Pasture,
Ottawa, Ontario
K1A 0T6
Telephone: (613) 951-0878

General Information:

Dr. Douglas Norris,
GSS Program Manager and Assistant Director,
Housing, Family and Social Statistics Division,
Statistics Canada,
7-D7, Jean Talon Bldg., Tunney's Pasture,
Ottawa, Ontario
K1A 0T6
Telephone: (613) 951-2572

GSS FAX: (613) 951-0387

BIBLIOGRAPHY

"A Safer Canada: Year 2000 Injury Control Objectives for Canada". Proceedings of a
National Symposium: May 21-22, 1991. Edmonton, Alberta).

Patterns of Criminal Victimization in Canada, Catalogue 11-612E, No. 2.

Accidents in Canada, Catalogue 11-612E, No. 3.

National Alcohol and Other Drugs Survey: Highlights Report. Health and Welfare Canada, June 1990.

APPENDIX A

QUESTION COMPARABILITY

Appendix A provides a comparison of the 1993 Cycle 8 and 1988 Cycle 3 questionnaires. Minor changes are distinguished by an "*" . For example, the change in reference period from Cycle 3 to Cycle 8 was classified as a minor change. In Cycle 3 the reference period was "1987", in Cycle 8 the reference period is "the past 12 months". This change is relatively minor, in so far as the meaning of the question is not altered by this change in reference period (See, for example, question B1).

Changes classified as major are distinguished by an "X". For example, the definition of what constitutes an accident was broadened from Cycle 3 to Cycle 8 and now includes "required medical attention from a doctor or nurse". This change has been classified as major because the meaning of what constitutes an accident has been changed somewhat. Furthermore, the changes which have occurred to the question regarding education (E16 in Cycle 8) have been classified as major (and marked with an "X"). Changes to this question include adding the qualifier "excluding kindergarten", as well as changing elementary or high school to elementary and high school. In general, any alteration which changes the meaning of the question is considered major and marked with an "X".

1993 General Social Survey
1988 GSS

Item	Topic	Comparison	Question
	PERCEPTIONS, HISTORY AND PERSONAL RISK		
A1	Introduction	*	A1
A2	Introduction	*	A2
A3	Neighbourhood crime compared to other areas in Canada	same	A3
A4	Perception of neighbourhood crime	*	A4
	PERCEPTION OF SAFETY		
A5	Introduction to section		
A6	When walking alone after dark	*	(VAWS - A2)
A7	Frequency of walking alone after		

dark
same
(VAWS - A3)

A8
Interviewer check item
same
(VAWS - A4)

A9
If safer would walking alone after dark increase
same
(VAWS - A5)

A10
Worried using public transportation after dark
*
(VAWS - A6)

A11
Frequency of using public transportation after dark
same
(VAWS - A7)

A12
Interviewer check item
same
(VAWS - A8)

A13
If safer would use of public transportation increase
same
(VAWS - A9)

A14
Worried home alone in the evening
same
(VAWS - A15)

PERCEPTION OF POLICE

A15a
Enforcement of laws
same
A7a

A15b
Prompt response to calls
same
A7b

A15c
Approachable/Easy to talk to
same
A7c

A15d
Crime prevention information
same
A7d

A15e
Ensuring safety of citizens in area

PERCEPTION OF COURTS

A16a
Providing quick justice
same
A8a

A16b
Helping the victim
same
A8b

A16c
Determining guilt or innocence of accused
same
A8c

A16d
Protecting the rights of the accused
same
A8d

A17
Severity of criminal court sentences
same
A9

A18
Previous contact with criminal court
X
A17

POLICE CONTACT IN PAST 12 MONTHS

A19a
For a public information session

A19b
For a traffic violation

A19c
As a victim of crime

A19d
As a witness to crime

A19e
Other police contact

SATISFACTION LEVEL WITH POLICE ACTION

A19a(A)
For a public information session

A19b(A)
For a traffic violation

A19c(A)
As a victim of crime

A19d(A)
As a witness to crime

A19e(A)
For any other reason

PRECAUTIONS TO PROTECT SELF OR PROPERTY FROM CRIME

A20a
Changed routine, activities, or avoided certain places
*
A23a

A20b
Installed new locks
*
A23c

A20c
Installed burglar alarm
*
A23c

A20d
Self defence course
same
A23d

A20e

Changed phone number
same
A23b

A20f
Obtained a dog

A20g
Obtained a gun

PRECAUTIONS TAKEN IN PAST 12 MONTHS TO PROTECT SELF OR PROPERTY FROM CRIME

A20a(A)
Changed routine, activities, or avoided certain places

A20b(A)
Installed new locks

A20c(A)
Installed burglar alarms

A20d(A)
Self defence course

A20e(A)
Changed phone number

A20f(A)
Obtained a dog

A20g(A)
Obtained a gun

A21
Most serious crime experienced by respondent

A21A
Age of respondent when most serious crime occurred

PRECAUTIONS TAKEN BY RESPONDENT TO FEEL SAFER FROM CRIME

A22a
Carry something to defend self or alert others
*
(VAWS - A17)

A22b
Lock doors when alone in car
*
(VAWS - A22)

A22c
Check back seat of car for intruders
*
(VAWS - A23)

A22d
Plan route

A22e
Stay home at night

A23
Level of satisfaction with personal safety

ACCIDENT PREVENTION PRECAUTIONS

A24a
Seat belt

A24b

Wear bicycle helmet

A24c
Safely store medicine and cleaning supplies

A24d
Use safety equipment

A25a
Working smoke detector
same
A24a

A25b
Working fire extinguisher
*
A24b

A25c
First aid supplies
same
A24c

A26
Household member trained in first aid

A27a
Work, attend class, go to meetings, do volunteer work
same
A25a

A27b
Restaurants, movies or theatre
*
A25b

A27c

Bars or pubs
*
A25c

A27d
Sports, exercise, and recreation
same
A25d

A27e
Shop
same
A25e

A27f
Visit relatives or friends in their home
same
A25f

A27g
Other evening activities
same
A25g

ACCIDENT SCREENING QUESTIONS

B1
Accident Definition
X
C1

Location of Accident

B2a
In a car, van, truck or on a motorcycle
*
C2a

B2b
On a recreational vehicle
*
C2b

B2c
While a pedestrian or on a bicycle
*
C2c

B3
At a job or business
*
C3

B4
During sports, exercise, recreation
X
C4

B5a
Fall which resulted in injury
*
C5a

B5b
Suffer burns, smoke inhalation, other fire-related accidents
*
C5b

B5c
Poisoning by substance or liquids
*
C5c

B5d
Cut with a knife, broken glass, other object
*
C5d

B6
Any other accidents which may/may not have resulted in injury
*
C6

B7
Interviewer check item
same
C7

B8
Driver of a motor vehicle
same
C8

B9
Kilometres or miles driven
*
C9

B10a
Hours spent as a driver of a vehicle
same
C10a

B10b
Hours spent as a passenger of vehicle
same
C10b

ACCIDENTS INVOLVING HOUSEHOLD CHILDREN

B11
Interviewer check item

B12
Child with next birthday

B13
Interviewer check item

B14
Injuries/poisonings in past 12 months

B15
Medical/dental attention received in past 12 months

B15b
Number of incidents requiring medical attention

B16
Month of most recent injury

B17
Time of day injury occurred

B18
Location at the time of accident

B19
Activity at the time of accident

B20
Nature of the injury

B21
Part(s) of body injured

B22
Where first taken to have injury examined

B23
Hospital care

B24
Children's hospital

B25
Nights in hospital

B25a
Number of nights in hospital

CRIMINAL VICTIMIZATION SCREENING QUESTIONS

C1
Introduction
*
D1

C2
Deliberate damage of personal or household property

C3a
Theft by force or threat of force
*
D2a

C3b
Break and enter or attempt to B&E
*
D2b

ITEMS STOLEN IN PAST 12 MONTHS

C4a
From things usually kept outside of home
*
D6a

C4b
From work, school, or public place
*
D6b

C4c
From a hotel, vacation home, cottage, car, truck or while travelling
*
D6c

C5
Household member motor vehicle owner
*
D4

C6a
Steal or attempt to steal motor vehicle or parts
same
D5a

C6b
Deliberate damage of motor vehicle
same
D5b

C7
Steal or attempt to steal anything else
*
D7

C8a
Attacked by anyone
*
D3a

C8b
Threatened to hit, attacked or threatened with a weapon
*
D3b

C9
Victim of forced or attempted sexual activity

C10
Touched in any sexual way against respondent's wishes

C11
Any other crimes
*
D9

C12
Interviewer check item
same
D10

C13
Interviewer check item
same
D12

ALCOHOL AND DRUG USE

Drinking

D1
Introduction

D2
Past 12 months
*
NADS - 12

D2a
Past month

D3
Ever drink
same
NADS - 16

D4
Never had a drink
same
NADS - 17

D5
12 month frequency
*
NADS - 19

D6
Number of drinks
*
NADS - 20

D7
Frequency of five or more drinks in past 12 months
same
NADS - 22

D8
Interviewer check item

D9
Frequency of five or more drinks in past month

D10
Highest number of drinks on one occasion
same
NADS - 23

D11
Frequency of driving after two or more drinks
same
NADS - 34

BAR/TAVERN

D12A
12 month frequency
*
NADS - 25f

D12B
Frequency of drinking
same
NADS - 25fB

D12C
Average number of drinks
*

NADS - 25fC

PARTY/SOCIAL GATHERING/WEDDING

D13A
12 month frequency
same
NADS - 26c

D13B
Frequency of drinking
same
NADS - 26cB

D13C
Average number of drinks
*
NADS - 26cC

QUIET EVENING AT HOME

D14A
12 month frequency
*
NADS - 25a

D14B
Frequency of drinking
same
NADS - 25aB

D14C
Average number of drinks
*
NADS - 25aC

HARMFUL EFFECTS OF PERSONAL DRINKING

D15a
On friendships or social life

X
NADS - 36a

D15b
Physical health
X
NADS - 36b

D15c
Outlook on life\happiness
X
NADS - 36c

D15d
Home life\marriage
X
NADS - 36d

D15e
Work, studies, employment
opportunities
X
NADS - 36e

D15f
Financial position
X
NADS -36f

D16a
Amount of drinking for
designated driver

D16b
Amount of drinking for person
who is not the designated
driver

EXPERIENCE WITH OTHER PEOPLE'S DRINKING PROBLEMS

D17a
Insulted/humiliated
same
NADS - 53ab

D17b
Serious arguments/quarrels
same
NADS - 53bb

D17c
Friendships break up
same
NADS - 53cb

D17d
Family problems/marriage difficulties
same
NADS - 53db

D17e
Passenger of a driver who drank too much
same
NADS - 53eb

D17f
Motor vehicle accident
same
NADS - 53fb

D17g
Property vandalized
same
NADS - 53gb

D17h
Pushed, hit or assaulted
same
NADS - 53hb

D17i
Disturbed by loud parties
same
NADS - 53ib

D17j
Financial trouble
same
NADS - 53jb

MEDICATIONS

D18a
Aspirin or similar pain reliever
same
NADS - 53Aa

D18b
Tranquilizers
same
NADS - 53Ab

D18c
Diet pills/stimulants
same
NADS - 58Ac

D18d
Anti-depressants
same
NADS - 58Ad

D18e
Codeine, demerol, morphine
same
NADS - 58Ae

D18f
Sleeping pills

same
NADS - 58Ak

DRUG USE

D19a
Marijuana\hash
*
NADS - 59aB

D19b
Cocaine\crack
*
NADS - 59bB

D19c
LSD, speed, heroin
X
NADS-59cB-eB

D20
Interviewer check item

D21
12 month frequency of
marijuana or hash use
same
NADS - 64

1993 General Social Survey
1988 GSS

Item	Topic	Comparison	Question
	CLASSIFICATION		
E0	Introduction	*	B1
E1	Date of birth	same	B12
E2	Type of dwelling	same	B2
E3	Dwelling owned by household member	*	B4
E4	Postal code	*	B3
E5	More than one telephone	*	B5
E6	All have same telephone number	same	B6

E7
Comments

E8
Interviewer check item

E9
Different telephone numbers
*
B7

E10
Business numbers only
same
B8

E11
Amount of business numbers only
same
B9

E12
Country of birth
*
B10

E12a
Province of birth
same
B10a

E13
Year of immigration
same
B11

E14
Language(s) first spoken
*
B13

E14A

Still understand language(s) first spoken
*
B13

E15
Language most frequently spoken at home
*
B14

E16
Years of elementary\high school
X
B15

E17
High school graduate
same
B16

E18
Further schooling after high school
same
B17

E19
Highest level of education attained
*
B18

E20
Religion
*
B19

E21
Frequency of church attendance
*
B20

E22
State of health

E23
Handedness

E24
Limited in amount or kind of activity

E25
Main activity in past 12 months
*
B22

E26
Full-time or part-time student

E27
Job or self-employed
*
B23

E28
Year of last paid work

E29
Number of weeks employed in past 12 months
same
B24

E30
Hours worked per week

E31
Paid worker\self-employed
same
B26

E32

Paid employees
*
B29

E33
Number of paid employees
same
B30

E34
Hours usually worked

E35
Whom did you work for
*
B31

E36
Type of business, industry or service
same
B32

E37
Type of work
same
B33

E38
Most important activities at job

E39
Personal income in past 12 months
X
B35

SPOUSE'S MAIN ACTIVITY

E40
Interviewer check item

E41
Spouse's main activity in past 12 months

E42
Interviewer check item

E43
Number of other household members who received income in past 12 months

E44
Interviewer check item

E45
Total household income

ACCIDENT REPORT

F1
Identification
same
F1

F2
Month of accident
*
F2

F3
Accident screening

F4
Interviewer check item

F5
Time of day accident occurred
*
F3

F6
Car, van, truck, motorcycle, snowmobile, or ATV involved
same
F7

F7
Pedestrian, on a bicycle or in a motor vehicle
same
F8

F7a
Type of motor vehicle
same
F8a

F8
Other pedestrians, bicycles or motor vehicles involved
same
F9

F8a
What exactly was involved in the accident
same
F9a

F9
Accident occur at place of work
same
F10

F10
Apply for Workers Compensation
same
F11

F11
Interviewer check item
same
F12

F12
Incident occur while participating in sport, exercise or recreational activity
*
F13

F13
Type of sport, exercise or recreational activity
*
F14

F14
Supervised by a trained leader

F15
Length of time respondent had been participating in sport, exercise, or recreational activity

F16
Frequency of participation in sport, exercise or recreational activity

F17
Recommended safety equipment

F18
Would safety equipment have prevented this incident

F19

Location of accident
X
F15

F20
Physical injury a result of incident
*
F16

F21
Type of injury
*
F17

F22
Location of injury
*
F18

F23
Was medical attention received
same
F20

F23a
Overnight stay in hospital
same
F20a

F23b
Number of nights spent in hospital
same
F20b

F24
Medical attention received from a doctor or nurse
same
F21

F25
Accident result in spending a day in bed
X
F22

F25a
Number of days spent in bed because of accident
same
F22a

F26a
Accident related to someone else's alcohol or drug use

F26b
Accident related to own alcohol or drug use

F27
Others injured in accident
same
F23

F28
Number of persons injured
same
F24

F29
Any persons injured less than 15 years old
same
F25

F29a
Number of persons less than 15 years old injured
same
F25a

F30a
Accident mainly caused by carelessness or unsafe activity
same
F26a

F30b
Accident mainly caused by something unpredictable or

unavoidable
same
F26b

F31
Whose carelessness mainly caused the accident
*
F27

F32
Main activity during week of accident
*
F28

F33
Disruption of main activity due to accident
same
F29

F33a
Number of days main activity was disrupted
same
F29a

F34
Extra expenses because of accident
*
F30

F35
Recovery of costs
same
F31

F35a
Recovery of costs through Workers Compensation or Insurance
same
F31a

F36
Out-of-pocket expenses
same

F32

F37
Interviewer check item
same
F33

F38
Interviewer check item
*
F34

F39
Number of accidents remaining
with similar details
*
F35

F40
Interviewer summary
same
F36

F41
Report Status
same
F38

F42
Interviewer check item
*
F39

F43
Interviewer check item
*
F40

CRIME INCIDENT REPORT

G1
Identification
same
G1

G2
Month incident occurred

*
G2

G3
Time of incident
X
G3

G4
Incident occur at place of work
same
G7

G5a
Place of work at time of incident
X
G9

G5b
Location of incident
X
G8

G6
Same dwelling as present dwelling
same
G10

G7
Type of dwelling at time of incident
same
G11

G8
Perpetrator resided with victim
*
G12

G9
Perpetrator let in
same
G13

G10
Person get in or just try to get in
same

G14

G11
Evidence of forced entry
same
G15

G11a
Type of evidence
same
G15a

G12
Interviewer check item
same
G16

G13
Present at time of the incident
same
G17

G14
Weapon present at time of incident
*
G18

G14a
Type of weapon

G15
Attacked in any way during incident
*
G19

G16
Threatened
*
G20

G17
Type of threat

G18
Perception of likelihood of threat being carried out

G19
Type of attack
*
G21

G20
Physical injury as a result of incident

G21
Medical attention at a hospital as a result of incident
same
G22

G21a
Overnight stay in hospital
same
G22a

G21b
Number of nights spent in hospital
same
G22b

G22
Medical attention from doctor or nurse
same
G23

G23
Stay in bed for all or most of a day because of incident
X

G24

G23a
Number of days spent in bed
same
G24a

G24a
Incident related to someone else's drug or alcohol use

G24b
Interviewer check item

G24c
Incident related to own drug or alcohol use

G25
One person involved in committing the act
same
G25

G26
Perpetrator known by victim
same
G26

G27
How well known was the perpetrator
same
G27

G28
Perpetrator's relationship to respondent
*
G28

G29
Interviewer check item
same
G29

G30
Number of persons involved in incident
same
G30

G31
Any persons known
same
G31

G32
How well known were the persons involved
same
G32

G33
Person's relationship to respondent
*
G33

G34
Anyone else harmed or threatened during incident
same
G34

G34a
Number of persons harmed or threatened
same
G34a

G35
Any persons harmed or threatened less than 15 years of age
same
G35

G35a
How many persons less than 15 years of age were harmed or threatened
same
G35a

G36
Interviewer check item

G36A
Stolen household or personal property
same
G36

G37
Items stolen
same
G37

G38
Value of property or cash stolen
same
G38

G39
Any stolen items recovered
same
G39

G39a
All stolen items recovered
same
G39a

G40
Interviewer check item
same
G40

G41
Take or attempt to take personal or household property
same
G41

G42
Type of property attempted to be taken
*
G42

G43
Household or personal property damaged but not taken
same
G43

G44
Type of damage
same
G44

G45
Monetary estimate of damage during incident
same
G45

G46
Damaged items repaired or replaced
same
G46

G47
Will damaged items be repaired or replaced
same
G47

G48
Interviewer check item

G48A
Insurance

G49
Obtain or attempt to obtain compensation through insurance
*
G48a

G49a
Was compensation obtained

*
G48a

G50
Obtain or attempt to obtain compensation through civil or criminal court
*
G48b

G50a
Was compensation obtained
*
G48b

G51
Out-of-pocket expenses
same
G49

G52
Main activity during the week of the incident
*
G50

G53
Disruption of main activity as a result of incident
same
G51

G53a
Number of days main activity was disrupted
same
G51a

G54
Incident brought to the attention of the police
same
G52

G55
Method by which police were made aware of incident
same
G53

REASON FOR REPORTING INCIDENT TO POLICE

G56a
Stop incident/receive protection
*
G54a

G56b
File report to claim insurance/compensation
same
G54b

G56c
Catch and punish offender
same
G54d

G56d
Duty to notify police

G56e
Recommended by third party

G57
Police action

G58
Level of satisfaction with police action

G59
Interviewer check item

REASON FOR NOT REPORTING INCIDENT TO POLICE

G60a
Dealt with another way

G60b
Not important enough
same
G56d

G60c
Fear of revenge by offender
same
G56c

G60d
Insurance would not cover

G60e
Police could not do anything about it

G60f
Police would not help

G60g
Did not want to get involved with police
*
G56f

G60h
Nothing was taken/items were recovered
same
G56a

G60i
Incident was a personal matter/did not concern police
same

G56e

G61
Interviewer check item

G62
Main reason for not contacting police

G63
Help sought from other person or organization
*
G57

G63a
Who was contacted for help

G64
Interviewer check item
same
G60

G65
Interviewer check item
same
G61

G66
Incidents remaining with similar details
*
G62

G67
Summary of incident
same
G63

G68
Report status
same

G65

G69
Interviewer check item
*
G66

G70
Comments

APPENDIX B

APPENDIX C

APPENDIX D

Topical Index to Variables

APPENDIX E

Data Dictionary

APPENDIX F

Record Layout