

EVALUATION IN THE 1971 CENSUS OF CANADA: OVERVIEW AND SELECTED FINDINGS

J. A. Norland, J. P. Litven and T. M. Cottrell

Statistics Canada, Ottawa, Ontario, Canada

Résumé—Il s'agit ici d'une présentation des conclusions d'une étude d'évaluation des données du recensement de 1971. Les trois principaux composants du programme d'évaluation sont: contrôle de qualité, études préparées au préalable et études "sérendipitiques". Le contrôle de qualité débute par la collection des données et il est incorporé à chaque phase de son traitement. Il y avait 53 études d'évaluation au préalable au programme de 1971. L'étude de l'entassement d'âge révèle que la préférence digitale est très faible au Canada. De plus, il y a une tendance déclinante dans l'indice de l'entassement d'âge durant des années. Les études "serendipitiques" ont produit un précieux aperçu de l'origine des erreurs. Ceci se révèle dans les études entreprises afin de résoudre les problèmes suscités par les inconsistances dans les données publiées.

Abstract — In this paper the findings of the recent evaluation studies of 1971 census data are presented. Quality control checks, preplanned studies and "serendipitous" studies are the three major components of the evaluation programme. Quality control checks begin with the collection of data and are incorporated into each stage of its processing. There were 53 preplanned evaluation studies in the 1971 programme. The study of age-heaping shows that digital preference in Canada is very small. Moreover, there is a declining trend in the index of age-heaping over the years. Serendipitous studies have given valuable insight on the source of errors. This is illustrated in the studies undertaken to resolve the inconsistencies in the published data.

I Introduction: Objectives of This Paper

1.1 The Scope of Census Evaluation

Census evaluation is an integral part of modern census-taking. In lieu of the old-fashioned approach which led "*most census officials to be smug and defensive about the official figures*" (Shryock and Siegel, 1973, p. 103), one now finds a candid and intensive effort to measure and report publicly the degree of error to which official figures may be subject. Indeed, the U.N. recommendations concerning the 1970s' national censuses (U.N., 1967, paragraphs 82 and 84) state explicitly:

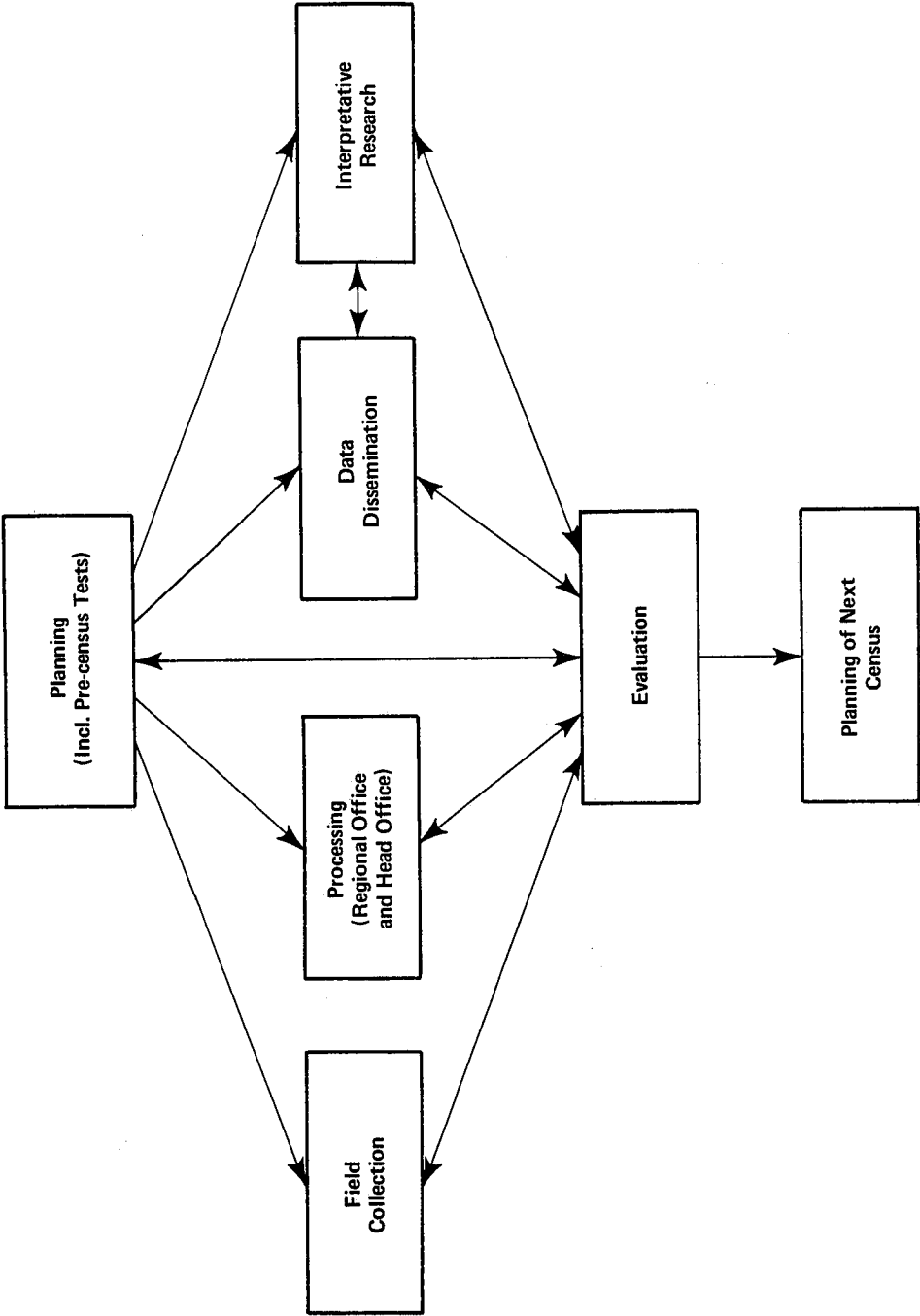
"82. Good census practice requires a careful consideration and evaluation of the completeness and accuracy of the census results.

"84. The publication of census results should include an estimate of coverage error, i.e., the amounts of over- and under-enumeration, together with a full indication of the methods used for evaluating the completeness of the data. Similarly, it is desirable to provide, so far as possible, an evaluation of the quality of the information on each topic, and of the effects of the editing procedures adopted."

The U.N. recommendation is based on the narrow definition of census evaluation, viz., the determination of accuracy limits for the data released by the census authorities. However, the scope and objectives of census evaluation may be conceived in a broader sense which encompasses also:

- (i) Continuous evaluation during the compilation and processing stages (e.g., by means of a quality control system) to establish standards of data acceptability.
- (ii) Exploration of means to produce "adjusted" or "corrected" data, parallel to the raw figures (as a typical example one may cite the computation of a graduated age distribution and the publication of both the raw and the graduated series).

FIGURE 1. RELATIONSHIP BETWEEN EVALUATION AND OTHER PHASES OF THE CENSUS



- (iii) Analysis of error determinants in order to propose means of minimizing the impact of error in future censuses. These evaluation studies include: (a) comparisons among alternative techniques, methods and systems (especially within the framework of pre-census tests) to identify those which are the most efficient; and (b) a retrospective examination of unforeseen obstacles, with the objective of overcoming them in the future.

According to this definition, census evaluation is associated with an assessment of both the census operations and the census data, and, as seen in Figure 1, it "interacts" with all the phases of a census. The discussion of the 1971 Census evaluation programme in this paper is based on the broad definition of "evaluation," as described above.

1.2 Census Evaluation in Statistics Canada: Policy and Programmes

The approach adopted by Statistics Canada (formerly Dominion Bureau of Statistics) vis-à-vis census evaluation may be assessed, inter alia, by examining (i) the pertinent policy statements and (ii) the ongoing evaluation programmes.

With regard to the first point, we note that in his general statement on research policy, the Director General of the Content and Analysis Branch/Census Field, Statistics Canada, identified evaluation as a major area of research with which the organization is concerned in the context of improving the quality, efficiency and usefulness of census data (Bond, 1974; the other major areas are anticipatory, developmental and interpretative research). It is also recognized that while research in all the aforementioned areas may be undertaken by various organizations (government departments, academic staffs, etc.), the responsibility for evaluative research rests primarily with Statistics Canada.

Four considerations substantiate this statement. First, the organization which is charged with compiling and disseminating a body of data is naturally responsible for informing users about the accuracy and limitations of the data, as well as for preparing "adjusted" data if necessary. Second, the quality control system alluded to under Section 1.1(i) is an integral part of processing the census data, and has to be implemented by Statistics Canada staff. Third, the more rigorous methods of data evaluation, viz., micro-matches (which concern checks of individual records, as opposed to macro-matches which concern aggregates) require access to confidential census returns which are available only to staff members. Furthermore, data evaluation often requires consultations and discussions with those who were involved either in decision-making or in the actual work at the various stages of the census-taking operation; it is extremely difficult for non-staff members to engage in extensive consultations and discussions of this kind. Finally, we noted under Section 1.1(iii) that evaluation studies provide the basis for identifying efficient census techniques, methods and systems (when alternatives are compared), as well as for suggesting means to improve future censuses. These suggestions have a particularly strong impact if they come from Statistics Canada staff members who either plan the census or are in direct contact with the census planners.

Being aware of these considerations, Statistics Canada has continuously expanded the census evaluation programme in terms of both scope and resources. The 1971 Census programme, for example, comprised over 50 projects in the preplanned category alone (see Section 2.3, and a selected list of projects in the Appendix). This compares with the six projects associated with the 1966 Census. Yet, there seems to be little awareness of the evaluation programme among demographers and statisticians outside of Statistics Canada. The prime reason probably stems from the fact that most of the information concerning the evaluation programme is reported by the census staff in papers and memoranda designed for "internal" dissemination. Notable exceptions include a chapter in the 1971 Census Administrative Report (Canada, forthcoming), an unpublished paper by Brackstone (1973)

TABLE 1. MAJOR CATEGORIES OF ERROR: DESCRIPTION, MINIMIZATION, MEASUREMENT AND CORRECTIVE STEPS

Type of Error	Description/Characterisation	Minimization	Measurement	Correction/Adjustment
Coverage	Omission or multiple enumeration of individuals or complete households. Particularly "vulnerable" are: infants and young children, transients, students, the aged, persons with more than one home, "de jure persons" temporarily abroad, inhabitants of remote areas.	Choosing an appropriate reference date for the census; designing rigorous checks for all stages of the census, e.g., assuring adequate mapping, clear instructions to canvassers, follow-ups, etc.	The best known techniques include: reverse record checks; demographic techniques (includes comparison of two successive censuses and cohort reconstruction), comparisons with parallel data (macro or micro matches); re-enumeration on a sample basis; internal consistency checks (includes an examination of the relevant time series).	Application of "multipliers" to arrive at better population estimates.
Response	Deliberate or un-intentional misreporting by respondent, and mis-interpretation or mis-recording by enumerator. Particularly recurring are the following response errors of age: reporting of age 0 as age 1; "preference" of terminal digits 0, 5, 2, and 8; overstatement around age of legal majority, age 65 and extreme old age; understatement among adults, especially women.	Careful recruitment and training of field personnel. Preparation of suitable questionnaires and enumeration techniques on the basis of research and pre-census tests. Thus: inclusion of questions with which respondents are able to cope; cautious wording; specifically: date of birth in lieu of age; self-enumeration with adequate instructions to respondents; assistance service (by phone or otherwise).	All the techniques listed under "coverage" apply if conducted separately for specific population groups, e.g., on an age-sex specific basis. Additionally: measures of age heaping (e.g. Myers, Bachl, Whipple, U.N. Secretariat method, Carrier and Ramachandran).	Implementation of efficient edit and imputation systems. Application of age-sex specific "multipliers". Graduation (data smoothing, interpolation, etc.).

Processing	<p>Loss of, or unjustified change in, the data entering the processing stage. Coding errors at the field-coding stage.</p> <p>Introduction of errors in the data capture stage; special example: incorrect "reading" by FOSDIC. Deficiencies of the edit/imputation system.</p> <p>Incorrect assignments for "unknown"; incorrect edits for conflicting responses.</p>	<p>Implementation of strict quality control standards, upgrading of equipment, especially for EDP. Also, reduction of processing volume (via sampling), permits the allocation of more resources for quality control.</p>	<p>Special samples, usually incorporated in the processing stage as quality control procedures or evaluation studies. The "unknown" category of each variable warrants special attention.</p>	<p>Corrective measures (e.g., post-edit fixes) based on the results of quality control and/or evaluation studies.</p>
Sampling	<p>Inherent in the application of sampling. (Applies to variables included in the sample questionnaires of censuses.)</p>	<p>Determination of optimal sample size and sampling technique.</p>	<p>Determination of sampling error by using sampling theory.</p>	<p>Providing confidence intervals for all census counts.</p>

Note: Certain measures which have been suggested and applied in the literature account for the combined effect of several types of error. These measures include the RMSE which is computed for the 1971 Census tables (e.g., Canada, 1973, "Introduction") and indices such as the response variance, index of net shift, index of stability and index of consistency (Spiegelman, 1970, Section 3.4).

and a few references in a recent paper by Fellegi (1973). In an attempt to contribute towards a wider dissemination of this information, the present paper has two specific objectives:

- (i) To provide a general review of selected components of the 1971 evaluation programme (Sections 2.1, 2.3 and 2.5);
- (ii) To present selected findings from recent evaluation studies, in order to illustrate the types and the estimated magnitude of the errors which affect the 1971 Census data. Special emphasis is laid on those studies which are related to the *basic demographic data*, i.e., the data on age, sex and marital status (Section 2.2, 2.4 and 2.6).

1.3 Some Pertinent Background Information

The remainder of this Introduction provides two general background notes pertinent to the discussion in Sections 2.1 to 2.6.

First, a note concerning the classification of census errors. Statistical data from national censuses are generally subject to four categories of errors coverage, response, processing and (if applicable) sampling. The first three are sometimes referred to collectively as "non-sampling errors." The recent literature dealing with the minimization, measurement, and correction of these errors is quite extensive. Table 1 attempts to summarize some of the major points which are particularly relevant to the 1971 Census evaluation programme, and which are referred to in Sections 2.1 to 2.6. However, the table should be considered merely as a guide; it does not attempt to be exhaustive or to provide detailed explanations (for extensive discussions the reader is referred, *inter alia*, to: Barclay, 1958, Ch. 3; Shryock and Siegel, 1973, Ch. 3, 4, 7, 8, and 22; Spiegelman, 1970, Ch. 3; and Wolfenden, 1954, Ch. 2, 4 and 5).

The second note concerns the 1971 Census methodology. Virtually all the steps of the census operation — from the preparation of maps, through the canvassing proper, to the final processing — have bearing on the issue of census evaluation. Clearly, a comprehensive review of these steps is beyond the scope of this paper. Suffice it to list here some of the new features of the 1971 Census, i.e., features which were not incorporated in previous censuses, and to introduce the basic terminology (a detailed account is given in the forthcoming Administrative Report of the 1971 Census; brief accounts may be found in: Beynon et al., 1970 and 1971; Canada, 1971d; Haas and Martin, 1973; MacIntosh, 1971; and Porter, 1971).

Self-enumeration. This method was introduced in the 1971 Census primarily to reduce response errors stemming from canvasser's bias and communication difficulties between canvasser and respondent; also, self-enumeration provides respondents with more time to consult documents prior to completing the census questionnaire. (Self-enumeration covered about 97 per cent of the population, the exclusions being in the remote areas which were enumerated by canvassers.)

Date of birth item. This item was introduced in 1971 because international experience had shown that an item on date of birth yields more accurate data than does the corresponding item on age. Furthermore, in 1971 the respondent was requested to both write in and code his date of birth. This feature permitted the evaluation programme to examine the accuracy of self-coded responses.

Sampling. Sampling has been applied in previous Canadian censuses but in 1971 it incorporated several new aspects pertinent to the evaluation programme. For example, the 1971 sampling applied a raking-ratio weighting system which had to be evaluated; it added to the total variance a component of sampling error which had to be computed; it reduced the cost and time required for processing the census data and thus permitted more resources to be used for the evaluation programme (for further details, see: Barckstone, 1971; Dodds, 1971; and Freedman, 1973). Another pertinent point is the fact that information on the basic

demographic variables (age, sex, marital status, mother tongue) was compiled from the entire population but were stored on two separate files: the "2A file" accounting for all the households, and the "2B file" accounting for a sample of one-third of the households. Due to sampling error and the independent processing of each file, the possibility arose of the two files yielding different distributions for a given variable (say, age); this point had to be dealt with carefully by the evaluation programme.

FOSDIC. In processing the 1971 data, the major device for data capture was the *Film Optical Sensing Device for Input into Computers (FOSDIC)*. From the viewpoint of evaluation, FOSDIC eliminated processing errors due to keypunching of cards, since the FOSDIC input consists of microfilmed census questionnaires. However, FOSDIC also required extensive checks and evaluations of its operation. (See also paper by Kaplan et al., 1973.)

The STATPAK retrieval system. This system was designed to serve users who required unpublished data, by providing a direct and efficient retrieval system. The assistance of STATPAK to the evaluation programme was immeasurable, since virtually any distribution could be obtained promptly from the census files.

The evaluation programme. A small-scale evaluation programme was incorporated in previous censuses, but in 1971 the programme assumed a new character by virtue of its unprecedented diversity and scope. Thus, in addition to the input provided by experts in statistical methodology, the programme also benefitted from the contributions of numerous subject-matter analysts. Further details concerning the scope of the evaluation programme are given in the course of the detailed discussion in Section 2.

II *The Evaluation Programme of the 1971 Census*

This section refers to three components of the 1971 evaluation programme:

- (i) Quality control procedures which were incorporated in the various census operations. Among other things, these quality controls determined the acceptability standards associated with those processing errors that could be anticipated.
- (ii) Preplanned evaluation studies which were designed to examine and measure the major foreseeable determinants of coverage, response, processing and sampling error.
- (iii) "Serendipitous" studies, i.e., studies which were initiated as a consequence of the discovery of inconsistencies in the data released by the census authorities.

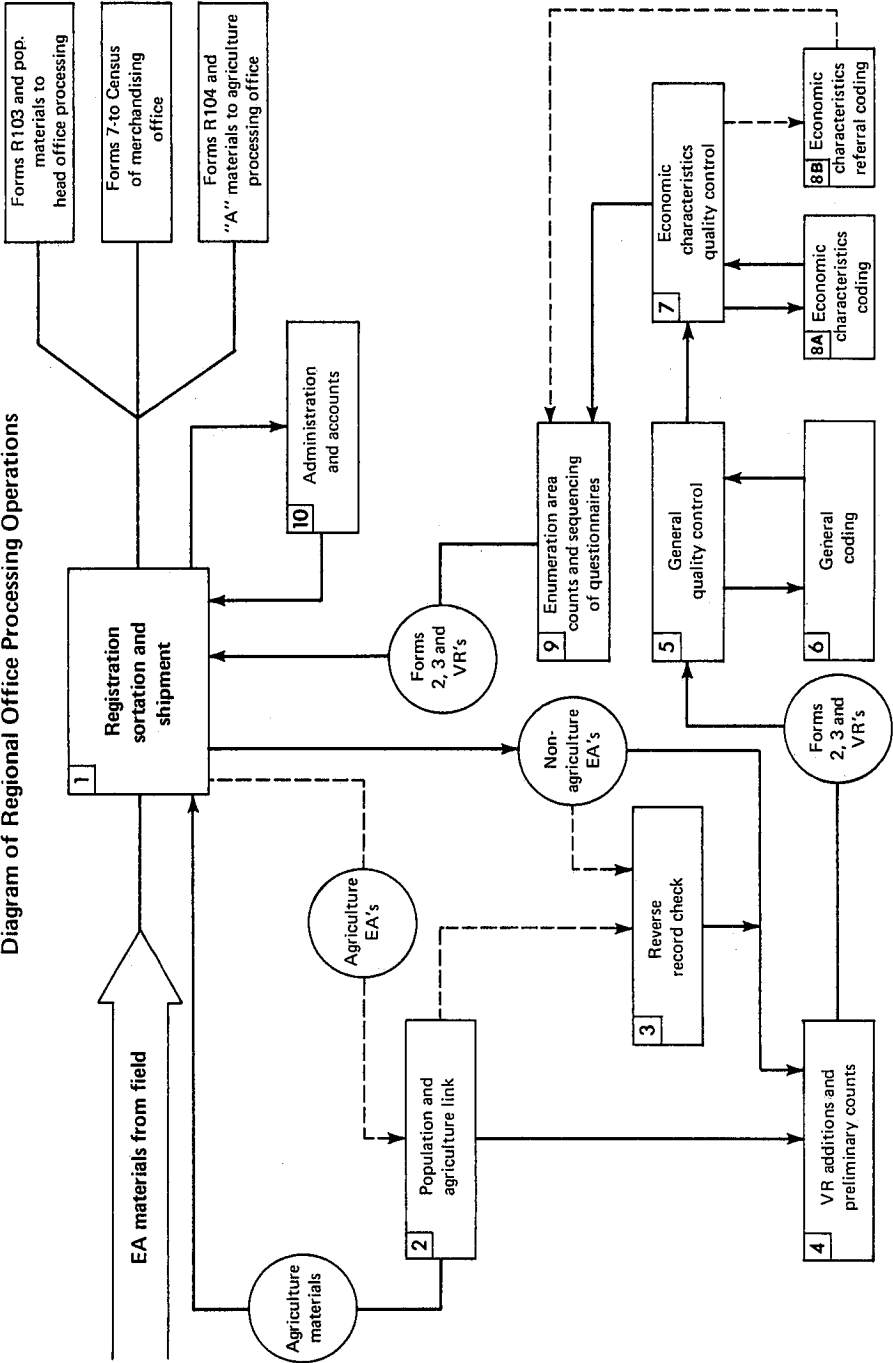
With regard to each of these three components, the subsequent discussions begin with a brief and general overview and proceed to report a specific example in detail.

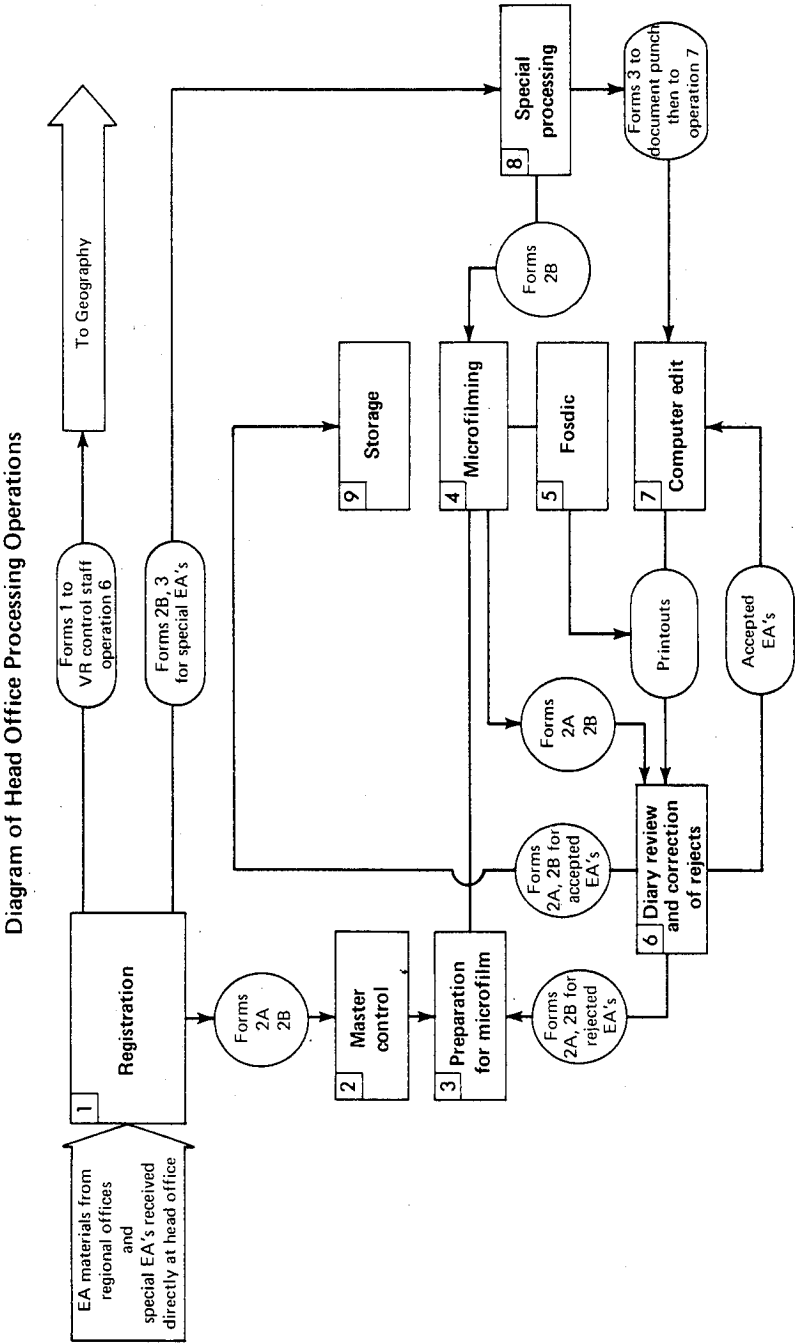
2.1 *Quality Control: An Overview*

Quality control procedures were incorporated in all the steps of the 1971 Census. The present section cites selected examples concerning the pre-operational stage as well as the stages of field collection of data, the Regional Office Processing (ROP) and the Head Office Processing (HOP).

With the introduction of FOSDIC, the standard of the printed census questionnaires became a crucial issue from the viewpoint of reducing processing error. For this reason, the preparation of materials for the 1971 enumeration was stringently quality-controlled. For instance, the census questionnaires were sampled in both single sheet and booklet formats to ensure that they were suitable for FOSDIC reading. Test decks of collated questionnaires were also sampled and put through the microfilm cameras to ensure there would be no mechanical feed problems. The pencils to be included in the census package underwent special testing on a random sampling basis as did the contents of the "stuffed" questionnaire envelopes.

FIGURE 2. REGIONAL AND HEAD OFFICE PROCESSING
Diagram of Regional Office Processing Operations





In the field collection of data, one of the very first instructions to the 1971 enumerator concerned her/his responsibility for checking and correcting the map given to her/him. This check is crucial for minimizing coverage error. Similarly, the enumerator was instructed to manually edit the census questionnaires and to follow them up according to predetermined procedures. For example, each questionnaire was to undergo a quality control check to verify that it included a coded answer to the question on date of birth (Canada, 1971a). Prior to having the material on a given enumeration area sent for processing, it was subject to a rigorous quality control check (on a sample basis) conducted by a specially trained officer.

Several operations of the regional processing stage were concerned with quality control (see upper panel of Figure 2, and detailed discussion in Canada, 1971c). One of the more important of these consisted of clerical staff "matching" each questionnaire with the entries in the listing book ("Visitation Record" or VR) which was compiled by the enumerator during the field collection stage. The objectives of this "match" were to ensure that the questionnaires represented complete coverage, and to provide preliminary census counts for an early release.

The coding process was subject to rigorous quality control, consisting of independent verification on a sample basis, re-coding of rejected lots and retraining or re-assigning inadequate coders.

Referring to HOP, one may distinguish between two groups of quality control procedures. The first consisted of technically-oriented checks, e.g., examination of the quality of the microfilms which fed the data to the FOSDIC machines, or verification that the computer processing operations — such as imputation of sex in cases of "no response" — performed as specified. The second group of checks was subject matter-oriented, designed to verify that the processed data were free of inconsistencies.

The technically-oriented checks were incorporated in each of the nine HOP operations (see lower panel of Figure 2, and detailed discussion in Canada, 1971b). At the outset, when the census forms arrived from the Regional Office, they were checked in accordance with a detailed procedure, to verify that all the necessary materials had indeed been received. Subsequent to this quality control operation, the shipment was incorporated in the master control system and (generally) processing of the various census questionnaires which had been received, began.

In most cases the next step involved preparation for microfilming, a process which included various edits of a quality control nature. Examples include re-marking of substandard coding entries, repairing damaged questionnaires and a check to verify that at least one basic characteristic was coded for each person listed in the census questionnaire (the existence of at least one coded entry was essential for the processing of a given individual; if necessary, the sex item was imputed in the course of this check). During the microfilming stage, rigid quality control standards were enforced, before filming and after the films were developed; these checks were designed to ensure that the films were suitable for the next stage, i.e., FOSDIC reading (as a point of interest, 115 of the 15,575 films were rejected due to poor developing).

FOSDIC reading was subject to checks throughout each day's processing, to ensure a high quality output. The checks consisted of feeding "known" data in microfilm format to the machine and of comparing the resulting printouts with verified standard printouts. Subsequent to the FOSDIC stage, a new battery of checks were performed as part of the Diary Review (see Freedman 1970a, 1970b). The objective of this phase was to detect and correct errors due to the FOSDIC system, faulty microfilming or questionnaires which were either missing or defective (i.e., dirty or poorly marked). As part of this process, control counts based on the Visitation Record were compared with the preliminary computer counts for population, households and dwelling units in each enumeration area. Additionally, the

non-response rate for each census question was compared with predetermined tolerance limits of non-response. A special check was performed on all questionnaires reporting ages over 100; as outlined in Section 2.2, however, this check proved insufficient. After acceptable processing was achieved for all the enumeration areas within a given geographic area (e.g., municipality), preliminary computer counts were produced and compared with the counts which were compiled independently from the Visitation Records of the same area. Once again, areas which deviated from the predetermined tolerance limits were investigated and corrective action was taken as necessary.

Like the preceding operations, the edit/imputation stage which followed was also accompanied by the technically-oriented quality control checks. Additionally, two checks of the subject-matter type were introduced, viz., "Diary II" and checks of basic frequency distributions. For a given variable with n classes i (say, sex with the $n=2$ classes "male" and "female"), Diary II rendered an $n \times n$ matrix showing how many cases were edited from class i to the other $n-1$ classes; this provided a means for detecting deficiencies in the edit system (see Freedman, 1971, 1972; and Viveash, 1971). The frequency-distribution checks enabled the subject matter analysts to report weird distributions and similar inconsistencies, so that one could initiate studies to determine the causes and the possible corrective measures (usually, post-edit fixes). A typical example is discussed in detail in the subsequent section.

The last HOP operation concerned document storage. The successful execution of any "serendipitous" evaluation study (Section 2.5) and of many preplanned studies (Section 2.3) depended heavily on efficient storage of both the census questionnaires and their microfilms. Hence, even this operation, which at first glance may seem unrelated to the evaluation programme, is actually a vital component thereof.

2.2 *Quality Control: The Centenarians*

When scrutinized as part of the subject-matter checks (see Section 2.1), the number of centenarians according to the tentative census file (referring to that version of the census file which was subject to "post-edit fixes," i.e., "fixes," or changes, which were introduced after the major edit and imputation stage had been completed) appeared to be considerably higher than expected: 2,655 as compared with 365 in 1966 and 274 in 1961. To examine this issue further, the records of all the individuals whose age, as stored on the tentative census file, exceeded 100 were compared with the corresponding records on the microfilmed questionnaires. The comparison yielded the following results:

- (i) In 290 cases (11 per cent) age 100+ was imputed as a consequence of non-response. Generally, in the event of non-response for age, an assigned age was determined by adding 25 to the age of child. In those cases where age of child was 75+ and parent's age was not given, the standard imputation created a centenarian.
- (ii) In 1301 cases (49 per cent) a conflict between the coded entry and the written-in entry was detected: for example, date of birth was coded as 1860 and written in as 1880.
- (iii) In 28 cases (1 per cent) a processing error was detected: for example, both the coded and the written-in entries indicated age 85, but the tentative census file indicated 105.

The age of the individuals mentioned in (i) to (iii) was modified via a post-edit fix. However, 1,036 cases (39 per cent) could not be attributed to either processing error or respondent coding error. This large number of centenarians, therefore, could result from response bias, an error factor which warranted a separate study.

2.3 *Preplanned Evaluation Studies: An Overview*

The 1971 evaluation programme comprised 53 preplanned studies (see Appendix for list of selected projects. Subsequent citations of specific evaluation studies refer to this list). The

magnitude of this programme called for a special unit — the Evaluation Section — to co-ordinate, and to monitor the progress made in, the various components. The pertinent studies may be classified broadly into those concerned with data evaluation (i) directly (e.g., the evaluation studies MR-10 and MR-11 on response rate); and (ii) indirectly, i.e., in the course of evaluating other aspects of the census: methods, systems, resource utilization, organization, management, etc. (specific examples include, in particular, the 19 studies under project MR-14 — see Appendix, Footnote 5).

Paramount among the “direct” evaluation studies are the Reverse Record Check (MP-1) and the Mean Square Error Study (MP-2). The first is based on a micro-match between (i) a sample of 27,500 records which were expected to be found among the 1971 returns, and (ii) the actual 1971 records. The “expected records” were derived from the 1966 Census, from records on 1966-71 births and immigrants, and finally from information on persons missed in the 1966 Census. This study revealed an overall under-enumeration of about two per cent, with particularly high rates among young adults of either sex. About half of the missed persons came from households which were enumerated partially, and the balance from those which were missed.

The Mean Square Error (MSE) Study MP-2 was designed to compute measures of reliability of the 1971 data. As defined in this context, the MSE comprises components due to sampling variance (if applicable), response variance and processing variance. (Generally, however, the MSE as defined in the literature comprises components due to bias and variance alike.) Some tables of the MSE have been published in selected census bulletins (e.g., Canada, 1973) and additional tables will be incorporated in the introductions to the various census volumes.

Response bias was the topic of several preplanned evaluation studies. One example, concerning age-heaping, is discussed in detail in the following section.

2.4 Preplanned Evaluation Studies: Age-heaping

The ongoing study of age-heaping in the 1971 Census was initially outlined by Saveland (1971). “Age-heaping” (or “digit preference”) refers to the tendency of some respondents to “prefer” an age which terminates with digits such as 0 or 5, and to “avoid” ages which terminate with digits such as 1 or 9. Of the various indices, the one suggested by Myers is, perhaps, the most commonly applied; its rationale and application have been described, *inter alia*, in Shryock and Siegel (1973, Chapter 8). The following discussions focus on that part of the age-heaping study which is based on Myers’ Index; the findings are cited from Norland and O’Grady (1974/75).

Table 2 presents Myers’ Index by sex for the censuses of 1921-71. Two methodological points should be borne in mind in connection with these data. First, the theoretical range of variation of Myers’ Index extends from zero per cent if no age-heaping exists, to 90 per cent if the age returns of the entire population are concentrated in one given preferred digit. Second, Meyers’ Index may be computed either for an entire age distribution (say, ages 0-89), or for a selected age interval. Analyses have indicated that age-heaping is usually “concentrated” in the approximate age range 20-69, while the younger and older groups involve other types of age-misreporting. Consequently, the indices for the latter age range are usually higher and more meaningful than the indices for 0-89. Both sets of indices, however, are reported in Table 2. (1956 Census data are not shown because ungraduated figures by single years of age are unavailable. 1951 data are shown only for the age range 20-69 because data by single years of age are unavailable for ages 70+.) The major findings that emerge from Table 2 are the following:

- (i) Canada and the provinces have been distinguished throughout the entire period in question by a low degree of age-heaping. Thus, Myers’ Index for the age interval 20-69

TABLE 2. MYERS' INDEX (PER CENT) OF AGE-HEAPING BY SEX,
CANADA AND PROVINCES, 1921-71

GEOGRAPHICAL AREA	M A L E S						
	1921	1931	1941	1951	1961	1966	1971
AGE INTERVAL: 0-89							
CANADA	2.87	1.67	0.92		1.04	1.09	0.68
NEWFOUNDLAND					1.04	0.95	0.88
P.E.I.	3.39	1.94	1.30		0.97	1.23	0.51
NOVA SCOTIA	2.90	1.86	0.96		0.94	0.91	0.59
NEW BRUNSWICK	2.97	1.54	0.95		1.23	1.02	0.59
QUEBEC	2.30	1.50	0.87		1.32	1.24	0.75
ONTARIO	3.13	1.80	0.99		0.91	1.24	0.65
MANITOBA	3.12	1.73	0.99		0.90	0.91	0.75
SASKATCHEWAN	2.71	1.33	0.73		1.02	0.74	0.95
ALBERTA	2.79	1.64	0.96		0.75	0.83	0.63
B.C.	4.00	2.17	1.17		1.31	1.09	0.83
AGE INTERVAL: 20-69							
CANADA	5.01	2.90	1.37	1.28	1.79	2.02	0.71
NEWFOUNDLAND				1.65	2.16	1.51	1.19
P.E.I.	5.64	3.11	1.80	1.11	1.59	1.73	0.76
NOVA SCOTIA	5.27	3.24	1.38	1.20	1.56	1.34	0.52
NEW BRUNSWICK	5.34	2.75	1.10	0.96	1.62	1.34	0.66
QUEBEC	4.44	2.90	1.31	1.27	2.16	2.16	0.82
ONTARIO	5.13	2.95	1.39	1.50	1.67	2.32	0.74
MANITOBA	5.40	2.99	1.63	1.33	1.73	1.73	0.70
SASKATCHEWAN	4.80	2.49	1.19	0.94	1.66	1.31	0.36
ALBERTA	4.64	2.68	1.31	1.05	1.44	1.66	0.57
B.C.	5.85	3.24	1.54	1.46	2.11	1.92	0.98
F E M A L E S							
	1921	1931	1941	1951	1961	1966	1971
AGE INTERVAL: 0-89							
CANADA	3.01	1.97	1.15		1.00	1.22	0.58
NEWFOUNDLAND					1.01	0.99	0.68
P.E.I.	3.07	2.70	1.66		1.17	1.30	0.70
NOVA SCOTIA	3.22	2.07	1.16		0.91	0.93	0.56
NEW BRUNSWICK	3.24	1.98	1.19		0.90	1.00	0.63
QUEBEC	2.45	1.83	1.15		1.35	1.45	0.63
ONTARIO	3.50	2.27	1.37		0.90	1.30	0.59
MANITOBA	3.32	2.14	1.31		0.99	1.08	0.71
SASKATCHEWAN	2.63	1.33	0.85		0.71	0.81	0.92
ALBERTA	2.79	1.61	0.88		0.78	1.06	0.54
B.C.	3.51	2.21	1.21		1.23	1.16	0.51
AGE INTERVAL: 20-69							
CANADA	5.20	3.33	1.68	1.52	1.69	2.11	0.68
NEWFOUNDLAND				1.73	1.58	1.39	0.89
P.E.I.	5.34	4.55	2.13	1.48	1.54	1.38	0.89
NOVA SCOTIA	5.35	3.23	1.65	1.32	1.34	1.36	0.52
NEW BRUNSWICK	5.48	3.28	1.67	1.39	1.50	1.49	0.47
QUEBEC	4.49	3.23	1.81	1.59	2.09	2.38	0.65
ONTARIO	5.63	3.55	1.82	1.80	1.60	2.28	0.79
MANITOBA	5.54	3.45	1.85	1.58	1.50	1.82	0.75
SASKATCHEWAN	4.66	2.58	1.28	0.88	1.02	1.11	0.57
ALBERTA	4.86	2.92	1.44	1.21	1.46	1.86	0.64
B.C.	5.59	3.44	1.74	1.56	1.99	2.05	0.87

FOR EXPLANATIONS, SEE TEXT.

varies from 0.36 per cent in Saskatchewan, 1971 (males) to 5.85 in British Columbia, 1921 (males).

- (ii) The overall 1921-71 trend is one of falling indices; one finds, however, inexplicably high indices in 1961 and 1966 for both sexes and for virtually all provinces. A pattern of declining age-heaping has been found in many countries; with regard to the U.S.A., for example, one finds:

Census Year	1880	1900	1920	1930	1940	1950	1960
Myers' Index	10.4	4.7	4.5	4.3	3.0	2.2	0.8

(Data cited from Shryock and Siegel, 1973, p. 208. Data refer to ages 23+ and to both sexes together.)

- (iii) For a given census, Myers' Index for males is usually lower than that for females (with the inexplicable exception of the 1961 Census). This finding, too, conforms to the pattern prevalent in international data.
- (iv) With the possible exception of the 1921 data, the degree of age-heaping in Canadian censuses has been so low that meaningful inter-provincial differentiation is virtually precluded. This applies in particular to the 1971 data for which the range of Myers' Index extends from 0.36 per cent (Saskatchewan, males) to 1.19 per cent (Newfoundland, males).
- (v) Comparing the indices for the age interval 0-89 with those for the age interval 20-69, one finds the latter to be systematically higher than the former. This finding conforms to patterns noted in other studies (e.g., Bachi, 1954), and is a consequence of the aforementioned fact that age-heaping is "concentrated" in the approximate age interval 20-69.
- (vi) With regard to individual digits, detailed data reveal that the general pattern in Canada, 1921-71, conforms to international findings: 0 and 5 emerge as the most "preferred" digits, and 1, 3 and 9 as the most "disliked."

It will be noted, in this connection, that since the 1971 age data were compiled from a question on date of birth (Section 1.3), a series of indices is being computed currently to check for heaping at "preferred years," e.g., 1900, 1910, etc.

2.5 "Serendipitous" Evaluation Studies: An Overview

Compared with the quality control checks and the preplanned evaluation studies, documentation on the "serendipitous" studies is rare. The major reason stems from the very nature of these studies. As a rule, "serendipitous" studies are prompted by a specific problem which a researcher encounters in the course of carrying out a given research project; the results are rarely published separately. One study for which documentation is available, however, is reported in detail in the next section. This study emerged in the course of preparing interpretative analyses on nuptiality and marital status. The inconsistencies in the published data necessitated an evaluative study, as a first step; the fundamental nature of the problems which surfaced justified the immediate release of the findings in a series of internal reports (Norland, 1974/75).

The dearth of data on "serendipitous" studies precludes the presentation of a comprehensive overview. As a substitute, the following paragraphs provide brief notes on two additional and typical "serendipitous" studies.

In the course of preparing the introductory bulletin for the 1971 Census volume on families, a question arose concerning the different numbers of *census families* (see Dictionary of the 1971 Census terms, Canada, 1972), which are reported in various official tables. For

example, the officially published national count of census families renders 5,076,085 according to the 2B file, and 5,070,685 according to the 2A file referred to in Sect. 1.3 (a difference of 5,400, or 0.11 per cent). With the objective of explaining the source of these differences, a small scale "serendipitous" study was undertaken to prepare and publish a technical users' guide (Wargon, 1975).

With regard to the study on age at first marriage (subsequently referred to as "age at marriage") it will be recalled, firstly, that the 1971 Census asked for date of birth and date of first marriage. However, the coded entries for month of birth were only two: "January to May," and "June to December." In computing age at marriage, the following algorithm was applied:

- (i) Age at marriage = (year of marriage) - (year of birth);
- (ii) If month of birth is June to December, *and* if month of marriage is January to May, deduct 1 from age at marriage of (i).

Example 1: Birth — June, 1930; marriage — July, 1950; age at marriage = 20, in accordance with (i).

Example 2: Birth — June, 1930; marriage — January, 1950; age at marriage = 19, in accordance with (ii).

This algorithm, however, will generate an error if both month of birth and month of marriage fall in the same block (January to May or June to December), *and* if month of marriage precedes month of birth.

Example 3: Birth — February, 1930; marriage — January, 1950; age at marriage = 20, in accordance with (i); actual age at marriage = 19.

In a study related to fertility data, Lavis (forthcoming) documented this point in detail and proceeded to show that a similar error, albeit of a smaller magnitude, was also introduced in the 1961 data. Means for correcting the 1971 data on age at marriage, some of which were published in Canada (1974), are currently being examined. (A simple solution, for example, is the publication of data on *duration of marriage* by age at census time, in lieu of data on *age at marriage* by age at census time.)

2.6 "Serendipitous" Evaluation Studies: The Marital Status Data

As noted previously, the "serendipitous" study on marital status was initiated because the published data were deemed inconsistent; specifically, the figures from both the 2A and the 2B files for the "young" (say 15-34) widowed and divorced persons in 1971 seemed incompatible with the 1951-66 time series (See Tables 3-6 which are based on officially published figures). As an illustration, let us examine the 2A data on widowed males (Table 3).

In the censuses of 1951-66, the number of widowed males, 20-24, varied in the range 180-300; in 1966, for example, the number was 291. In 1971, however, the corresponding number jumped to 1,340, representing a 360 per cent increase over 1966 (Table 3, Panels A and B). Also noteworthy are the 1966-71 "jumps" for cohorts. Thus, in 1971 there were 555 widowed males aged 25-29 per 100 widowed males aged 20-24 in 1966; the comparable ratios for the same age groups in 1951-56, 1956-61 and 1961-66 varied between 250 and 350 (Table 3, Panel C). Similar unusual "jumps" may also be discerned for other age groups of the widowed male population (Table 3) as well as for the widowed female population (Table 4), the divorced male population (Table 5) and the divorced female population (Table 6).

These findings cannot be explained solely in terms of demographic trends. Such factors as the 1966-71 increase in the population of young adults and the 1968 change in divorce laws may, perhaps, explain these trends partially, but the main factor is undoubtedly associated with (non-sampling) error. *A priori*, one may point to the possibilities of (i) coverage error,

TABLE 3. DISTRIBUTION OF WIDOWED MALES BY AGE GROUP, CANADA, 1951-71

Age group	Census year				
	1951	1956	1961	1966	1971
<u>A. Absolute numbers</u>					
15-19	15	39	88	160	1,365
20-24	197	186	233	291	1,340
25-29	845	684	561	572	1,615
30-34	1,564	1,408	1,303	1,032	1,800
35-39	2,936	2,516	2,337	2,405	2,640
<u>B. Per 100 in preceding census</u>					
15-19	—	853
20-24	—	94	125	125	460
25-29	—	81	82	102	282
30-34	—	90	93	79	174
35-39	—	86	93	103	110
<u>C. Per 100 of given cohort in preceding census</u>					
15-19	—	—	—	—	—
20-24	—	838
25-29	—	347	302	245	555
30-34	—	167	191	184	315
35-39	—	161	166	185	256

See notes at end of Table 6.

TABLE 4. DISTRIBUTION OF WIDOWED FEMALES BY AGE GROUP, CANADA, 1951-71

Age group	Census year				
	1951	1956	1961	1966	1971
<u>A. Absolute numbers</u>					
15-19	83	99	262	167	1,540
20-24	823	858	931	1,031	2,405
25-29	3,007	2,429	2,371	2,477	3,970
30-34	6,489	5,690	5,036	4,891	5,480
35-39	10,305	10,586	10,106	11,283	9,670
<u>B. Per 100 in preceding census</u>					
15-19	—	..	265	64	922
20-24	—	104	109	111	233
25-29	—	81	98	104	160
30-34	—	88	89	97	112
35-39	—	103	95	112	86
<u>C. Per 100 of given cohort in preceding census</u>					
15-19	—	—	—	—	—
20-24	—	..	940	394	1,440
25-29	—	295	276	266	385
30-34	—	189	207	206	221
35-39	—	163	178	224	198

See notes at end of Table 6.

TABLE 5. DISTRIBUTION OF DIVORCED MALES BY AGE GROUP, CANADA, 1951-71

Age group	Census year				
	1951	1956	1961	1966	1971
<u>A. Absolute numbers</u>					
15-19	1	11	22	21	405
20-24	148	174	283	373	2,110
25-29	791	822	1,272	1,528	6,875
30-34	1,347	1,405	2,168	2,304	8,480
35-39	1,749	1,853	2,795	3,085	9,480
<u>B. Per 100 in preceding census</u>					
15-19	—
20-24	—	118	163	132	566
25-29	—	104	155	120	450
30-34	—	104	154	106	368
35-39	—	106	151	110	307
<u>C. Per 100 of given cohort in preceding census</u>					
15-19	—	—	—	—	—
20-24	—
25-29	—	555	731	540	1,843
30-34	—	178	264	181	555
35-39	—	138	199	142	411

See notes at end of Table 6.

TABLE 6. DISTRIBUTION OF DIVORCED FEMALES BY AGE GROUP, CANADA, 1951-71

Age group	Census year				
	1951	1956	1961	1966	1971
<u>A. Absolute numbers</u>					
15-19	20	47	58	76	565
20-24	584	667	926	1,468	5,165
25-29	2,044	1,953	2,452	3,665	12,085
30-34	3,046	3,042	3,627	4,524	12,890
35-39	3,513	3,628	4,538	5,402	13,030
<u>B. Per 100 in preceding census</u>					
15-19	—
20-24	—	114	139	159	352
25-29	—	96	126	149	330
30-34	—	100	119	125	285
35-39	—	103	125	119	242
<u>C. Per 100 of given cohort in preceding census</u>					
15-19	—	—	—	—	—
20-24	—
25-29	—	334	368	396	823
30-34	—	149	186	185	352
35-39	—	119	149	149	288

— = Not applicable.

.. = Computations involve cell frequencies of less than 99 cases and are therefore deleted.

e.g., over-enumeration of "young" widowed and divorced persons in 1971, or a consistent under-enumeration of these persons in 1951-66; (ii) response error, e.g., failure of "young" widowed and divorced respondents to report correct age and/or marital status; and (iii) processing error, e.g., incorrect FOSDIC readings, deficient edits/imputations, and faulty post-edit fixes.

To examine this issue in detail, a three-phase micro-match comparison was devised. The objective of the first (or "pilot") phase, was (i) to identify, for a given age group, all the households in which a "young" widowed or divorced person was recorded according to the 2A file (these are subsequently termed "pertinent households"); (ii) to extract from the 2A file the complete records of all persons in these pertinent households; (iii) to compare these records with the corresponding records from the "unedited" 2A file, i.e., the tentative file which was created subsequent to the FOSDIC stage — this comparison would permit an assessment of the impact of edits/imputations and post-edit fixes; (iv) to compare the "unedited" 2A records with the microfilmed 2A questionnaires, and (to a limited degree) with the questionnaires proper — these comparisons would permit an assessment of the impact of processing errors resulting from the "initial EDP" on the one hand, and response errors on the other hand. ("Initial EDP" refers to Box No. 5 in the lower panel of Figure 2, which is labelled "FOSDIC." Actually, this stage includes: FOSDIC sensing, decoding, purges of duplicate records, and record consolidation. Involved are FOSDIC hardware and software alike.)

While the "pilot" phase of the study concentrated on "suspicious" population groups (i.e., "young" widowed and divorced persons), the second phase was designed to examine a random sample of all households. The objective here was to explore the pervasiveness of the types of error which the "pilot" study was to identify as the main determinants of the above-mentioned inconsistencies in the 1971 data on marital status. The procedures of the second phase were outlined parallel to those of the "pilot" phase, i.e., steps (ii) to (iv), as described above, were to be repeated for a sample of all households exactly as they were implemented in the "pilot" phase for the pertinent households.

The third phase of the study was designed to repeat the "pilot" phase (and, if the budget permitted, the second phase as well) with reference to the 2B file.

At the time of writing, the "pilot" phase, restricted to the province of B.C. (367 households), is being completed, the second, implemented, and the third, planned. As noted above, the findings from the "pilot" phase are being reported in a series of Advance Reports (Norland, 1974/75) a summary of which is given in the following paragraphs.

- (i) Coverage error, response error, edits/imputations and post-edit fixes were identified as relatively insignificant contributors to the inconsistencies in the 1971 data on marital status.
- (ii) "Initial EDP," on the other hand, appeared to be the major determinant of these inconsistencies. Specifically, on the basis of an examination of the microfilmed questionnaires it was found that "initial EDP" either assigned incorrect codes for age and marital status or failed to read the coded entries. The reason for these deficiencies will be explored further by examining a sample of questionnaires (rather than microfilmed returns).
- (iii) An important byproduct of the "pilot" phase concerns recommendations aimed at improving the data of future censuses. For example, checking for response error in the age variable, we found a recurring pattern of a sixty-year gap between the coded and the written-in entry. This resulted from the fact that on the 1971 Census questionnaire, the coding columns for decade of birth were arranged thus:

o	186-	o	192-
o	187-	o	193-
o	188-	o	194-
o	189-	o	195-
o	190-	o	196-
o	191-	o	197-

By coding in the correct line but in the wrong column (i.e., by crossing the separating line) an individual is "rejuvenated" by 60 years. A suggested remedy is to arrange the columns thus:

o	186-	192-	o
o	187-	193-	o
o	188-	194-	o
o	189-	195-	o
o	190-	196-	o
o	191-	197-	o

It should be emphasized at this point that while the 1971 data on selected age-sex-marital status groups, such as those which are reported in Tables 3 to 6, indicate inconsistencies, the validity of the 1971 data as a whole is neither disputed nor questioned. As the study quoted here progresses, means will be suggested for adjusting the "deviant" cells of the age-sex-marital status cross-classification. In the interim, users may be advised to exercise caution, as elaborated in the next section.

III Summary and Conclusion

Reviewing the evaluation programmes for the censuses of 1961-76, one observes a clear pattern of evolutionary expansion. It is the authors' opinion, however, that the 1971 programme represents the most significant breakthrough. Thus, the 1971 Census was the first to incorporate a comprehensive programme, backed by a budget (over \$1,000,000 for the preplanned studies alone) and many researchers who committed themselves to the programme. The preplanned studies represent one component of the 1971 programme; the numerous quality control checks on the one hand, and the "serendipitous" studies on the other, represent two other components of this programme.

With regard to each of the three components mentioned, the foregoing discussions attempted to provide (i) a general overview and (ii) a detailed description concerning one specific example. These were intended to demonstrate, respectively, the scope of the programme as a whole, and selected findings from individual studies within the programme. At the same time, the discussions were aimed at emphasizing that *census data, published as well as unpublished, should be used and interpreted with caution*. This, of course, also applies to data from previous Canadian censuses. However, since the 1971 Census is the first to incorporate a comprehensive evaluation programme, little documentation is available on errors in previous censuses. Consequently, it may appear that the 1971 data are particularly deficient; there is no evidence to substantiate such an impression. In this context, the general information provided in the introductory bulletins of the various census volumes serves a vital function but may often prove insufficient. These bulletins cannot, of course, account for findings of "serendipitous" studies which may begin years after the bulletins are released. Nor can they provide detailed methodological discussions. What, then, are the best channels for obtaining up-to-date information pertinent to the census evaluation?

For researchers outside of Statistics Canada the authors suggest three major sources (Statistics Canada researchers have direct access to "internal" reports which provide the best answer to the question). The first is contained in papers which Statistics Canada researchers prepare for presentation at conferences or for publication in the professional literature. Examples include the papers by Brackstone (1973), Fellegi (1973) and Kaplan et al. (1973).

With respect to several issues, these papers constitute the best available source of written documentation, but, to be sure, this source also suffers from several deficiencies (e.g., the papers concerned are usually unofficial, the responsibility for the contents resting solely with the authors). The second source is the forthcoming Administrative Report of the 1971 Census. The great advantage of this document is the official approval associated with it, but one should not overlook the disadvantages of this source, e.g., the fact that it deals with the evaluation issue in one relatively brief chapter. Hence, the third and best way of obtaining updated information on the census evaluation is through communication with Statistics Canada researchers who are engaged in the evaluation programme. Contact may be facilitated through the User Inquiry Services, Statistics Canada, Census Field, No. 8 Temporary Building, Ottawa, Canada, K1S 5A4. Alternative (and more efficient) methods are being considered in the framework of the 1976 Census.

Acknowledgment

The authors are grateful to the many staff members of Statistics Canada who contributed to this paper by providing information, documents or commentary. Special thanks are due to D. Bond, G. Brackstone, D. R. Bradley, I. P. Fellegi, T. Gray, A. Kempster, E. Lewis, K. P. O'Grady, E. T. Pryor, J. K. Scott, L. O. Stone and D. Wilson.

Disclaimer

The responsibility for the data and opinions contained in this paper rests entirely with the authors.

References

- Bachi, R. 1954. Measurement of the Tendency to Round-off Age Returns. *Bulletin de L'Institut International de Statistique* (Proceedings of the 28th Session, Rome, September 6-12, 1953), 34(3): 129-138.
- Barclay, G. W. 1958. *Techniques of Population Analysis*. New York: John Wiley and Sons.
- Beynon, T. G., G. B. Joshi and F. K. Pierre-Pierre. 1970. Some Methodological Aspects of the 1971 Census of Canada. *Canadian Journal of Economics* 3: 95-110.
- . 1971. Plans for the 1971 Census of Canada. *Demography* 8: 271-286.
- Bond, D. E. 1974. A Statement on Research Policy Within Census. Statement disseminated among Census Field professional staff members December 5. Ottawa: Census Field, Statistics Canada.
- Brackstone, G. J. 1971. The Census Weighting Procedures. Population and Housing Research Memorandum No. PH-Gen-9. Ottawa: Census Division, Dominion of Statistics.
- . 1973. The Identification and Measurement of Error in the Canadian Census. Paper presented at the Conference of the International Association of Survey Statisticians, Vienna, August.
- Canada. Dominion Bureau of Statistics. 1969/71. Planning and Procedural Memoranda, Office Processing Series. Issues 71-OP-1 to 71-OP-16. Ottawa: Census Division, Dominion Bureau of Statistics.
- . 1971a. 1971 Census of Canada: Procedures Manual/Self Enumeration Areas. Manual No. 41. Ottawa: Dominion Bureau of Statistics.
- . 1971b. 1971 Census of Canada: Head Office Processing Manual. Manual No. H167. Ottawa: Dominion Bureau of Statistics.
- . 1971c. 1971 Census of Canada: Regional Processing Manual. Manual No. R123. Ottawa: Dominion Bureau of Statistics.
- . 1971d. Plans for the 1971 Census. *Canada Yearbook*, 1970/71, pp. 207-210. Ottawa: Information Canada.
- Canada. Statistics Canada. 1972. *Dictionary of the 1971 Census Terms*. Ottawa: Statistics Canada.

- _____. 1973. 1971 Census of Canada, Population: Sex Ratios. Volume I, Part 2, Bulletin 1.2-2. Ottawa: Information Canada.
- _____. 1974. 1971 Census of Canada, Population: Age at First Marriage. Volume I, Part 5, Bulletin 1.5-10. Ottawa: Information Canada.
- _____. Forthcoming. 1971 Census of Canada, Administrative Report. Volume VI, Part 1.
- Dodds, D. J. 1971. Sampling in the Self-enumeration Areas of the 1971 Census. Population and Housing Research Memorandum No. PH-Gen-8. Ottawa: Census, Statistics Canada.
- Eastman, D. A. 1970. Computer Edits Production Flow. Planning and Procedural Memorandum No. 71-CP-5. Ottawa: Census Division, Dominion Bureau of Statistics.
- Fellegi, I. P. 1973. The Evaluation of the Accuracy of Survey Results: Some Canadian Experience. *International Statistical Review* 41: 1-14.
- Freedman, H. A. 1970a. The 1971 Census Data Capture System: Report of the Diary Task Force. Planning and Procedural Memorandum No. 71-CP-4. Ottawa: Census Division, Dominion Bureau of Statistics.
- _____. 1970b. Final Report of the Diary I Task Force. Planning and Procedural Memorandum No. 71-CP-6. Ottawa: Census Division, Dominion Bureau of Statistics.
- _____. 1971. Report of the Diary II Task Force. Planning and Procedural Memorandum No. 71-CP-7. Ottawa: Census Division, Dominion Bureau of Statistics.
- _____. 1972. Diary II: Five per cent EA Sample of Initial and Edited Records. Planning and Procedural Memorandum No. 71-SM-1 (Revised). Ottawa: Census Division, Statistics Canada.
- _____. 1973. On Differences Between 1971 Census 100 per cent and Sample Figures. Population and Housing Research Memorandum No. PH-Meth-2. Ottawa: Census Division, Statistics Canada.
- Haas, B. and A. Martin. 1973. How Canada Takes a Census. Interim Paper No. 73-1. Ottawa: Statistics Canada.
- Kaplan, D. L. et al. 1973. The US and Canadian Experience with FOSDIC. Paper presented at the International Population Conference, Liège, September.
- Lavis, D. R. Forthcoming. Canadian Census Data on Age at First Marriage. Paper proposed for publication in a Census Field "internal series." Ottawa: Characteristics Division, Census Field, Statistics Canada.
- MacIntosh, D. A. 1971. Census Methodology. Population and Housing Research Memorandum No. PH-Gen-6. Ottawa: Census Division, Statistics Canada.
- Norland, J. A. and K. O'Grady. 1974/75. Evaluation of the 1971 Census Data on Age-sex: Advance Reports (No. 1, 5.11.74; No. 2, 20.11.74; No. 3, 15.12.74; No. 4, 28.1.74). Ottawa: Characteristics Division, Census Field, Statistics Canada.
- Norland, J. A. 1974/75. Evaluation of the 1971 Census Data on Marital Status: Advanced Reports (No. 1, 4.11.74; No. 2, 10.2.75; No. 3, 6.3.75; No. 4, 22.3.75; No. 5, 17.4.75). Ottawa: Characteristics Division, Census Field, Statistics Canada.
- Porter, W. D. 1971. The 1971 Census — An Overview. *Canadian Statistical Review* 46(4): 4-5, 114-116.
- Saveland, W. 1971. 1971 Census Evaluation Project DAR-2: Measurement of Age-heaping. Planning and Procedural Memorandum No. 71-E-15. Ottawa: Census Division, Dominion Bureau of Statistics.
- Shryock, H. S. and J. S. Siegel. 1973. The Methods and Materials of Demography. Second Printing (revised). U. S. Bureau of the Census. Washington, D. C.: U. S. Government Printing Office.
- Spiegelman, M. 1970. Introduction to Demography. Revised edition (third printing). Cambridge, Massachusetts: Harvard University Press.
- United Nations. 1967. Principles and Recommendations for the 1970 Population Censuses. Catalogue 67.XVII.3, Series M/44. New York: United Nations.
- Viveash, D. 1971. Examination Procedures for Diary II Output. Planning and Procedural Memorandum No. 71-SM-2. Ottawa: Census Division, Statistics Canada.

J. A. Norland, J. P. Litven and T. M. Cottrell

Wargon, S. T. 1975. The 1971 Census Housing, Household and Family Data: Some Technical Aspects and Considerations. Characteristics Division Research Memorandum No. 71-HOU-6. Ottawa: Census Field, Statistics Canada.

Wolfenden, H. H. 1954. Population Statistics and Their Compilation. Revised edition. Chicago: University of Chicago Press.

Received May, 1975; Revised August, 1975.

APPENDIX: LIST OF SELECTED PREPLANNED EVALUATION STUDIES
OF THE 1971 CENSUS

Project No.	Title/Description	Investigator	Documentation	Status
1. CA-1	Evaluation of the Pre-data Processing Stages of the 1971 Census	T.M. Cottrell J. Riddle T. McGurn	P & P/71-E-42 RM/71-E-16	Completed
2. CA-2	Evaluation of the Data Processing Stages of the 1971 Census	G. Berthiaume S. Bhargava C. Kneen	"Internal Report"	Completed
3. DAR-2	Measurement of Age-heaping	W. Saveland J.A. Norland	P & P/71-E-15 "Internal Report"	Ongoing
4. DC-3	Evaluation of 1971 Census Coding of Demographic and Social Characteristics	D.R. Bradley F. Lafontaine	P & P/71-E-36 RM/71-E-13 (rev.)	Completed
5. DC-4	Analysis of the Use of the "Miscellaneous" Group in the Coding of Vocational Courses	G. Tetlock	P & P/71-E-39	Ongoing
6. DC-6	Evaluation of 1971 Census Reporting of Data on Migration: Micro-match with Applications to the Citizenship Branch for Proof of Canadian Citizenship by Canadian Born and Foreign Born			Postponed
7. DC-7	Evaluation of 1971 Census Reporting on Demographic and Social Characteristics: Micro-match of By-product Data from the Reverse Record Check	L. Rouillard J.F. Gosselin	P & P/71-E-30 "Internal Report"	Completed
8. EC-1	Evaluation of 1971 Census Reporting of Occupation and Place of Employment	A. Kempster C. Montigny	P & P/71-E-47	Ongoing
9. EC-4	Evaluation of 1971 Census Income Statistics: Macro-comparison with Income Tax Statistics	R.S. Samlalsingh A. Rashid	P & P/71-E-37 "Internal Report"	Completed
10. FH-1	Evaluation of 1971 Census Reporting of Government Subsidized Dwelling Units: Micro-match with CHMC Lists	G.E. Priest	P & P/71-E-12 RM/71-E-11	Completed
11. FH-2	Evaluation of 1971 Census Reporting of Type of Dwelling: Micro-match with 1966 and 1961 Censuses	G.E. Priest D. Blanchet	P & P/71-E-10 (rev.) RM/71-E-10 (P & H/PH-Hou-3) (P & H/PH-Hou-5)	Completed
12. FH-4	Evaluation of 1971 Census Reporting of "Selling Value" of Owner-occupied Dwelling Units: Micro-match with Comparison Sales Catalogue Provided by Members of the Canadian Real Estate Association	G.E. Priest D. Alford M. Bailey	P & P/71-E-38 RM/71-E-9	Completed

Project No.	Title/Description	Investigator	Documentation	Status
13. FO-1	Evaluation of Coverage of Households and Persons: LFS - Census Match	C.G. Gilfillan J.F. Gosselin	P & P/71-E-9 "Internal Report"	Completed
14. G-4	Evaluation of 1971 Census Delineations of MA's and Urbanized Areas	F. Ricour-Singh		Ongoing
15. G-6	Evaluation of the 1971 Census Unincorporated Places Population Figures	M. Dehoux	P & P/71-E-34	Completed
16. MP-1	1971 Reverse Record Check	G. Brackstone J.F. Gosselin	P & P/71-E-6 RM/71-E-12 RM/71-E-23 "Internal Report"	Completed
17. MP-2	Mean Square Error Study	D. Dodds C. Hill	P & P/71-G-7	Ongoing
18. MP-3	Agriculture Quality Control (AQC) and Post-censal Survey	D. Holt	P & P/71-E-3 SC Cat. 96-701 SC Cat. 96-735	Completed
19. MP-4	Work Measurement Programme	E.J. Selwood	P & P/71-E-7 "Internal Report"	Completed
20. MP-5	Evaluation of Quality Control in 1971 Census	G.J.C. Hole	P & P/71-E-33 "Internal Report"	Completed
21. MP-6	1971 Census Observer Programme	F. Pierre-Pierre (See Footnote 3)	P & P/71-E-21	Completed
22. MR-2	Estimation of Sampling Bias	H. Aschenkamp D. Dodds G. Brackstone	P & P/71-E-44 "Internal Report"	Completed
23. MR-3	Evaluation of Change of Address on Coverage	H.V. Aschenkamp G. Brackstone	P & P/71-E-8 RM/71-E-14	Completed
24. MR-4	Mail Response Rate Study	K. Devon G. Brackstone	P & P/71-E-40 RM/71-E-18	Completed
25. MR-5	Estimation of Coding Variance	D. Dodds H. Arora	P & P/71-E-18	Ongoing
26. MR-8	Evaluation of Observation of 1970 U.S. Census	(See Footnote 4)		Completed

27.	MR-10	Evaluation of Response Rates of Collective Dwellings	(H.R. Arora (F. Pierre-pierre ((P & P/71-E-29 (P & P/71-E-46 ((P & P/71-E-5 (P & P/71-E-29 (P & P/71-E-46 (Completed
28.	MR-11	Evaluation of Response Rate of Private Dwellings	(G. Brackstone ((RM/71-E-20 (RM/71-E-21 (Completed
29.		Legitimate Non-response Study	J. Riddle J. McGurn	RM/71-E-15	Completed
30.	MR-12	Evaluation of Postal Check in Improving Coverage	L. Woelfle	P & P/71-E-41 RM/71-E-22	Completed
31.	MR-13	Evaluation of the Use of Temporary Residents Form	H. Arora G. Brackstone	P & P/71-E-32 RM/71-E-17	Completed
32.	MR-14	Administrative Support Systems	T.M. Cottrell (See Footnote 5)	RM/71-E-19	Completed
33.	OP-1	Evaluation of the Ramifications of Removing Office Coding from the Short Questionnaire 2A	T.M. Cottrell	P & P/71-E-43 "Internal Report"	Completed
34.	SJ-1	Evaluation of 1971 Census Responses to Selected Socio-economic Questions: Micro-match with Documents of the May/71 Household Facilities and Regular LFS	H. Singh H. Monck	P & P/71-E-19 P & P/71-E-19 (A) P & P/71-E-19 (A) rev.	Postponed
35.	SJ-2	Evaluation of Selected 1971 Census Statistics: Macro-comparison with Statistics of May/71 Household Facilities and Regular LFS June/71	H. Singh R.S. Samalsingh G. Tetlock A. Kempster	P & P/71-E-28 (A) rev. P & P/71-E-45 P & H/PH-EC-4 "Internal Report"	Completed

Notes: 1. "Investigator" - determined from printed documentation and interviews.

1. "Documentation" — P & H refers to the series "Population and Housing Research Memoranda"; P & P refers to the series "Planning and Procedural Memoranda"; RM refers to the series "Results Memoranda"; SC Cat. ... refers to an official Statistics Canada Publication.
2. Except as noted in footnote 4, the serial number of all RMs is preceded by "CON"; for example, 71-E-10 should be referred to as CON 71-E-10.
3. Reports related to Project MP-6 were released in a series of eight RMs: CON 71-E-1 through CON 71-E-8.
4. Reports related to Project MR-8 were released in a series of 29 RMs: US-E-1 through US-E-25 and US-EA-1 through US-EA-4.
5. Reports related to Project MR-14 were released in a series of 19 "Internal reports" as follows:

Authors	Title	Publication Date
1. Charlevoix, M.	Head Office Processing Procedures	July 14, 1972
2. Charlevoix, M.	Head Office Processing Training	July 14, 1972
3. Greenway, N.	Regional Office Processing Procedures	July 14, 1972
4. Greenway, N.	Regional Office Processing Training	July 14, 1972
5. Hall, M.E.	Administrative Support Services	November, 1972
6. Hall, M.E.	Scheduling, Planning and Controlling	January 31, 1972
7. Hicks, J.D.	Telephone Assistance Services (TAS)	March 24, 1972
8. Hicks, J.D.	Training	January 31, 1972
9. Jackman, D.	Geographic Programme for the 1971 Census	April 6, 1973
10. Laplante, J.M.	Regional Office processing Organization	December 29, 1972

11. Leboeuf, R.	1971 Census Account Forms	February 1, 1973
12. Leclair, R.J.W.	Enumeration of Indian Reserves	July 31, 1971
13. McCannon, M.	Field Collection Personnel: their Recruitment, Selection and Remuneration	March 3, 1972
14. McCannon, M.	Major Components of the Field Operations	March 10, 1972
15. McCannon, M.	Space Provided for the 1971 Census Programme	January 31, 1972
16. Martin A. & Hirsh, M.	Field Information Reporting System (F.I.R.S.T.)	January 31, 1972
17. Robillard, P.	Head Office Processing Personnel: their Recruitment, Selection and Remuneration	July 14, 1972
18. Robillard, P.	Regional Office Processing Personnel: their Recruitment, Selection and Remuneration	July 14, 1972
19. Trudel, R.P.	Logistics	February 15, 1972