

## DIRECTIVE NO. 1 <sup>2</sup>

### STANDARDS FOR STATISTICAL SURVEYS

The standards in this Directive apply primarily to statistical surveys. As used here, a "statistical survey" is one in which the primary purpose of the individual returns is to supply data which can be aggregated to provide statistical or numerical information needed on a particular class, segment, activity, or geographic area. Certain of the standards and recommended practices apply also to the compilation of statistics based on data from forms used in applying a Federal law or regulation to an individual or firm (such as tax returns or the financial and operating reports required by regulatory commissions), or applications and registrations, or administrative records. These standards have only limited relevance for administrative surveys designed to collect summary information needed by an agency in administering a specific program (such as summaries or procurements or production status, or fiscal accounting reports from the States or grant-in-aid programs).

Standards cannot be applied uniformly or precisely in all situations. Special considerations may be involved in exploratory, experimental, or methodological surveys or in pilot or preliminary surveys where the primary purpose is to test feasibility or to evaluate alternative approaches or techniques. Sponsors should be prepared, however, to justify any significant departures from these standards.

The standards in this Directive are minimal. The statistical agencies are encouraged to develop their own standards for surveys, which may be more detailed and in some cases more stringent.

---

<sup>2</sup> Directive No. 1 supersedes section 7(a) and Exhibit A of OMB Circular No. A-46 dated May 3, 1974.

#### 1. Purpose of the Survey

Before any other steps are taken in planning a survey there must be a clear understanding of the precise purposes to be served in terms of information to be collected, hypotheses to be tested, or problems to be solved. Consultation with users may be helpful in defining these purposes. It may be necessary to modify the initial goal of measurement of a concept in order to define an objective statistic which can be measured; e.g., "health" is not objectively measurable, but "bed-days" may be. It is helpful in reaching an understanding to prepare a concise statement of the information needed and the purpose or purposes to be served.

#### 2. Relation to Other Surveys or Programs

Before the scope and content of a survey are determined, prior work and current activities in the field should be reviewed. Those planning the survey should satisfy themselves that part or all of the data (or approximations) which are needed are not available from some existing source (published or unpublished) either within or outside the Government or could not more appropriately be obtained by adding questions to an existing survey or by some other agency. If there were a previous survey, it may reveal deficiencies which can be corrected in the new survey.

#### 3. Development of the Survey Plan

Decisions should be made on points noted in the subsections below in terms of the purposes to be served, costs involved, time required, degree of precision needed, and administrative limitations (established practices, location of field offices, etc.). When these decisions have been made it is useful to prepare a detailed

technical description of specifications for the survey for use of analysts and other survey personnel and to facilitate the preparation of a technical appendix or bulletin (if the survey is of sufficient importance) for publication along with survey findings.

**a. Target populations and extent of coverage.**

The target population or universe (i.e., all the people, establishments or other units that the survey designer wants to learn about) should be defined. Within the target population it must be decided whether coverage should be complete or partial. Consideration should be given to best ways of handling small units where these are likely to account for a very small proportion of the sales, consumption, cost or other value to be measured so no essential information will be lost by substantially reducing the coverage of such units.

Sample surveys have many advantages for most purposes over those involving complete enumeration. If the universe to be surveyed is large or geographically widespread and data are not required for small groups or areas, or if the report form or procedure is complex, sampling can be more economical for the agency, with consequent reduction in the reporting burden. Also, with sampling, nonresponse can usually be handled more effectively, the data can be processed more quickly, and in some instances the quality of the response can be improved. More attention can be given to developmental aspects of the survey since less time need be spent processing large numbers of responses.

Reasons for selection of a particular method of sampling should be included in the description of the survey plan.

**b. Sampling.** The first step is to develop and describe a frame, or sets of frames, i.e., the set of materials (lists, directories, records, maps, etc.) to enable the survey designer to deal with the universe. Ideally, the frame should include all the units in the universe. If the undercoverage in the frame is more than trivial a bias may be present in the survey results.

If the purpose of a survey is to develop estimates for a universe based upon a sample of the universe, the selection of the sample should be made in accordance with probability theory—that is, every sampling unit in the frame

should have a known, nonzero probability of being chosen, although the probabilities may be unequal. In such cases, probability sampling is necessary to avoid unknown biases of selection and to permit one to arrive at estimates for the universe with accompanying measures of reliability. However, where the sample is restricted to a small number of respondents or geographic areas, as in pilot studies, pretests, or some types of feasibility studies, a judgment or purposive sample may be indicated.

When the nature of a survey or study indicates the need for probability sampling, the particular method to be used should be administratively feasible and should provide maximum precision for the funds available. The size of the sample and the nature of the sample design (including estimation procedures) will depend not only on available funds, but also on such considerations as the amount of detail required in the statistical estimates, the precision desired for key items, the variability of the data being sampled, availability of other data resources, and the size of the universe. Once the sample units have been selected, no substitutions can be permitted for reasons of convenience or because of nonresponse unless provision is specifically made in the basic survey for acceptable alternatives.

Because of the technical nature of sampling and the many ways of development of estimates from a sample, agencies proposing sample surveys should have available a statistician or a consultant trained in sampling theory with experience in the conduct of surveys. In the case of departments having primary statistical agencies this may be a person on the staff of one of these agencies. A trained statistician is also necessary for contractors or grantees conducting federally sponsored sample surveys.

The description of the sample design should cover such items as the primary units (description and number in universe), the sampling unit used (at each stage if there is multistage or multiphase sampling), the sampling frame (including any exclusion or constraints), and the criteria or stratification and clustering (if the sample is stratified or clustered). It should also include cost estimates, determination of size of sample, variance estimates used in the optimum allocation of the sample, the method of drawing the sample at each stage, the meth-

od of estimation, and the method of estimating sampling variances. The description should indicate methods for controlling and measuring nonsampling error and biases in data collection, processing and tabulating.

c. **Frequency and timing.** It must be determined whether the survey is to be singletime, irregularly recurrent, or periodic. If periodic, consideration should be given to using a rotating sample basis at monthly, quarterly, annual or other intervals. The timing of a survey should be considered in relation to what is known about cyclical, seasonal or other variations over time in the data being studied, so that the survey results may be of maximum use.

d. **Method of collection.** A decision must be made as to whether the data are to be collected by mail, personal interview, telephone, telegraph, or other method, or by some combination of these methods. Nonsampling errors which can be expected in each of the methods should be considered and ways of measuring and minimizing them should influence the selection of the collection method. Experience from past experiments and surveys should be reviewed in making these decisions. If past experience is not helpful a pilot study to test alternatives may be advisable. The results of such a study should be made available to other interested agencies.

e. **Consideration of sampling and nonsampling errors.** If probability sampling is used, sampling error should be estimated, and whenever possible, total mean square error should be constructed in approximate terms. Potential nonsampling errors, including reporting errors, response variance, interviewer and respondent bias, nonresponse, imputation error and errors in processing the data should be carefully considered. This is important in designing the survey, in establishing controls over survey operations, and for the information of users of the data when they are published.

f. **Evaluation of survey data.** It is desirable to provide for the evaluation of survey results. This is especially important for repetitive surveys where the evaluation could provide future improvement of the data. Resources should be provided to carry out such evaluations.

g. **Plans for processing and tabulating the data.** The basic design of the survey plan should

provide procedures (including quality control) for editing, coding, and tabulating the data. These are necessary to ensure uniform processing and to measure quality. Copies of editing instructions, codes, etc. should be retained in the agency's files to answer inquiries and for possible use in future surveys. A decision, which will affect the form design, must be made on the manner in which the data will be processed—i.e., by use of punch cards, computer, optical scanning devices, etc. Tabulation plans, including dummy tables, should be worked out at this stage. This compels the agency to think through what it expects to get from the survey, and can result in changes in the survey design. The method of handling "unknowns" and refusals (e.g., whether distributed or shown separately) should be described. If the agency decides to have the data processing done under contract, the procedures to be followed must be clearly understood by both parties.

h. **Allowance for pretests.** It is desirable to test the feasibility of most new survey plans in advance, and the survey plan should include time and funds for this. A pilot study may be necessary to determine whether the survey is practicable or useful at all. The relative effectiveness and cost of alternative questionnaires, instructions, and operating procedures can be evaluated by means of a small-scale pretest. Pretesting can provide information on the probable refusal rate; the intelligibility to respondents (and interviewers) of the questions and instructions; and problems in compiling the survey results. The pretest is sometimes also used to provide estimates of variance which may aid in developing a more efficient sample design. The tabulations from the pretest should be limited to those necessary to answer the specific questions for which the pretest was designed, and generally should include additional tabulations for research purposes or preliminary findings only to the extent it is essential to allow sufficient time to make any changes indicated by the results of the pretest.

i. **Provision for follow-up.** In most surveys the response to the first attempt to collect the information is insufficient for final estimates, and a plan for follow-up is needed. Usually, if the survey is conducted by mail, one or more mail follow-ups are made to the nonrespond-

ents. If the rate of nonresponse is still unsatisfactory in terms of the nature and purposes of the survey, provision must be made for a more intensive type of follow-up, such as telephone calls or personal visits to a subsample of non-respondents.

j. **Proposed calendar.** Approximate dates for the following steps in the conduct of the survey, as well as the estimated manhours for each step, should be determined in advance. This tests the feasibility of the survey as well as helping to ensure its orderly completion.

Beginning and completion of the preliminary survey design

OMB clearance of pretest

Beginning and completion of one or more pretests

Analysis of the pretest results and modification of the survey design and questionnaire if indicated

OMB clearance of survey

Beginning of field work or mailing of questionnaires

Completion of field work or due date for questionnaires

Beginning and completion of follow-ups

Completion of the editing and coding

Completion of the tabulations

Completion of the preliminary report

Completion of the final report

k. **Cost estimates.** An estimate should be made of the cost of the survey, including costs of personnel, travel, equipment, and supplies, etc. The total should include the costs of pretests and follow-ups; preparing and printing the forms; compilation of the list of respondents; mailing or enumeration; and editing, coding, tabulation, and publication of the data. When appropriate an allowance for research should also be included, as should the estimated share of overhead costs. A more detailed description of costs to be included, together with suggested methods for developing these costs, is included in the "Guide to Estimating Reporting Costs" issued by the Office of Records Management of the National Archives and Records Service, General Services Administration.

#### 4. Questionnaire and Instructions

Even the best designed survey will produce disappointing results if the respondents do not understand the questions. The type and characteristics of the respondents should always be

kept in mind. For example, particular care should be taken to keep questions simple when addressed to persons of limited education. In general, legal or technical terminology should be avoided unless the respondents are lawyers or technicians.

A clear, easy-to-read, and easy-to-complete form is important for a successful survey. If it is to be self-administered, the questions should be as brief and self-evident as possible.

The questions or items on a form should be arranged in logical order, with related items generally grouped together. It is sometimes desirable to guide the thoughts of the respondents from the general to the specific. If the data are to be transcribed or machine processed, this should also be considered in the questionnaire format, the arrangement of items, and in provision for responses.

Separate instructions included with mail surveys are sometimes ignored. If possible the questions should be self-explanatory or brief parenthetical instructions should be associated with each question in order to avoid separate instructions.

The title of the questionnaire or form is generally the first thing the respondent reads and care should be given to its selection. It should be sufficiently informative to be used in lists of forms compiled for publication. Subtitles may be used where necessary to supplement the title or to cite the authority under which the form is issued. Forms and other documents approved by the Office of Management and Budget under the Federal Reports Act must carry an OMB number. Other identification such as an agency form number or control symbol may be desirable.

Instructions for return should be on the questionnaire or form itself in a mail survey, as the form may become separated from the transmittal letter and the return envelope. These instructions should give the address to which the completed form is to be returned, and ordinarily should specify a target date for return. The importance of a response should be emphasized.

Persons planning surveys may find helpful the *Household Survey Manual*<sup>3</sup> issued in 1969 by the Office of Management and Budget. This

<sup>3</sup> The Office of Federal Statistical Policy and Standards now has responsibility for this publication.

Manual presents the major concepts and definitions and some of the questions used by Federal agencies to collect data on personal and family characteristics, education, employment and unemployment, income, and quality of housing.

Guidelines for the design of forms contained in the Records Management Handbook, *Forms Design*, issued by the National Archives and Records Service (NARS) of the General Services Administration in 1960, may be useful. This manual covers the mechanics of form design, including size, margins and spacing, type faces and rules for printing, and forms construction (type of paper, ink, perforating, etc.). The 1960 Handbook has very limited applicability to forms used in automatic data processing. The NARS Handbook, "Specialty Forms," may be useful for forms used in automated systems.

### **5. Preliminary Clearance**

Although it is not necessary for an agency to request clearance until it is ready to collect data or promulgate recordkeeping requirements, many agencies have found it advantageous to take advantage of the proviso in OMB Circular No. A-40 which permits an agency to request "preliminary clearance" after they have formulated their plan but before they have signed a data collection contract. A disapproval of a clearance request can be costly to an agency after a contract has been let. While preliminary clearance does not guarantee a final clearance, it does indicate to the agency that the data collection or recordkeeping requirements will not be disapproved on the basis that the need for the data has not been established.

### **6. Control of Survey Operations**

For a successful operation, it is necessary not only that a good survey design be chosen, but also that the conduct of the survey be faithful to that design. Every phase of survey operation is subject to risks of deviation from intention. Positive procedures must be established to assure that these deviations are kept within tolerable limits. If sampling is used additional controls are needed to ensure that the selection of the sample, the collection of the data, and the tabulation and estimation are carried out as specified. Strenuous efforts should be made to collect data from every unit in the

sample, using follow-ups where necessary. If the data are obtained by field enumeration the interviewer should not be permitted to make substitutions (see 3b, Sampling).

Allowance should be made for sufficient time to train enumerators if the data are collected by personal or telephone interview. For large-scale surveys manuals should be prepared for use of field workers and persons doing the editing and coding as well as carrying out quality-control operations in order to ensure consistency and adherence to the survey design. Provision for adequate supervision must be made. Procedures should be prescribed for progress reporting. Quality control procedures consistent with the nature of the survey should be part of the planning process. This may include validity checks with subsamples of respondents. If a longitudinal survey is conducted, it is especially important to keep detailed records, because of probable attrition in the group studied and in the staff conducting the survey.

### **7. Preparation and Publication of Final Report**

Good statistical practice for the conduct of a survey includes careful presentation of results. Graphic presentations, as well as statistical tables, are often desirable, particularly in reports intended for laymen. Graphs and charts should be clearly labeled with a title, the time period to which the data refer, the geographic area included, and the units used on the vertical and horizontal scales. Conventional types of tables are text tables, summary tables and detailed (or reference) tables. For easy legibility the printed copy should be no smaller than newsprint and preferably larger. Either the title or the subtitle should contain information on "what," "how classified," "where," and "when," for the data in the table.

In preparing the final report attention should be given to Directive No. 2, "Standards for the Publication of Statistics."

### **8. Relations With the Public**

Finally, maintenance of good relations with the public is essential if Federal statistics are to continue to merit public support. Objectivity and integrity in the compilation and presentation of statistics is the surest means of obtaining such support. Particular attention, however, should be given to relations with respond-

ents and users of the statistics.

**a. Respondents.** Good relations with respondents involve obtaining the information needed with the minimum burden, reassuring respondents that their interests are being protected, and in general dealing with them fairly and honestly.

As noted above, the burden on respondents must be considered in designing the survey. In drafting the questionnaire, avoid questions which do not contribute to the purposes of the survey. The reportability of the data from the respondent's knowledge or from his records must be considered and particular attention should be paid to requests for information for prior periods or dates long past. When a new or substantially changed request is made for data which may be difficult to provide, it is sometimes helpful to give advance notice to prospective respondents. Agencies are encouraged to consult with potential respondents or their representatives in developing survey plans and forms.

To the extent possible respondents should be reassured that their interests are being protected. Agencies collecting data for general statistical purposes are usually in a position to assure respondents that the information they supply will be used only for statistical tabulations, and that individual returns will be kept confidential. Respondents to other types of surveys should be informed of the users of the data and the extent of disclosure. Agencies collecting data from business respondents particularly should be aware of the provisions of the Freedom of Information Act (P.L. 89-487) and may need to consult legal counsel on the extent to which confidentiality may be pledged. Agencies collecting data from individuals should be aware of the provisions of the Privacy Act of 1974 (P.L. 93-579).

The Office of Management and Budget requires special justification for surveys coming under the Federal Reports Act which involve questions of a sensitive nature. This justification should include the reasons why the agency considers the questions necessary and the specific uses to be made of the data obtained. While the meaning of "sensitive" will vary, depending on the respondent and the circumstances, questions on sex behavior and attitudes, mental illness and psychological problems, re-

ligious beliefs, and income and assets are generally regarded as sensitive by individuals. Business respondents often regard as sensitive questions on profits and wage and salary scales—particularly executive scales, as well as questions which may involve trade secrets. If an individual or firm is given all relevant information about a project and is completely free to participate or not, invasion of privacy as such may not be a problem, but response may be affected. Where a survey calls for information of a sensitive nature the feasibility of anonymous replies should also be considered.

Care must be taken to avoid giving respondents the impression that they must respond to surveys which are voluntary. For this reason the Office of Management and Budget has prohibited a statement on the form or in the letter of transmittal that this survey is authorized by law in surveys where response is not mandatory. Where response is mandatory this should be indicated and the applicable statute should be cited.

If response is voluntary, cooperation can best be obtained by explaining the purposes for which the data are to be used and by stating clearly and persuasively the need for the data by the Government or the public. Even where an agency has mandatory authority it is better to secure the willing cooperation of the respondents. Where the respondents are business firms or institutions, a commitment to provide a copy of the published report is frequently useful.

**b. Users.** Consultation with users is important to ensure maximum usefulness of the survey. In developing the survey plan users often can make contributions not only on the data to be collected, but also on the timing and frequency of repetitive surveys, and the degree of precision needed. In the case of important statistical series it also is desirable occasionally to obtain user comments on the usefulness of the published data by means of user questionnaires, establishment of advisory committees, or in other ways.

A description of the survey plan, an assessment of the accuracy of the data published, and a statement explaining any limitations should be available to users of the data (see Directive No. 2, "Standards for the Publication of Statistics").