
Statistical Policy Working Paper 4

Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistics



Statistical Policy Working Papers are a series of technical documents prepared under the auspices of the Office of Federal Statistical Policy and Standards. These documents are the product of working groups or task forces, as noted in the Preface to each report.

These Statistical Policy Working Papers are published for the purpose of encouraging further discussion of the technical issues and to stimulate policy actions which flow from the technical findings. Readers of Statistical Policy Working Papers are encouraged to communicate directly with the Office of Federal Statistical Policy and Standards with additional views, suggestions, or technical concerns.

**Office of
Federal Statistical
Policy and Standards**

**Joseph W. Duncan
Director**

Statistical Policy Working Paper 4

Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistics

Prepared by

Richard E. Deighton and James R. Poland,
Co-Chairmen, U.S. Postal Service
Joel R. Stubbs, Internal Revenue Service, and
Robert D. Tortora, U.S. Department of Agriculture
Subcommittee on Nonsampling Errors
Federal Committee on Statistical Methodology



U.S. DEPARTMENT OF COMMERCE
Juanita M. Kreps, Secretary
Courtenay M. Slater, Chief Economist
Office of Federal Statistical Policy and Standards
Joseph W. Duncan, Director

Issued: December 1978

Office of Federal Statistical Policy and Standards

Joseph W. Duncan, *Director*

George E. Hall, *Deputy Director, Social Statistics*

Gaylord E. Worden, *Deputy Director, Economic Statistics*

Maria E. Gonzalez, *Chair, Federal Committee on Statistical Methodology*

Preface

This working paper was prepared by a Glossary Task Force comprised of Richard E. Deighton and James R. Poland of the U.S. Postal Service, Joel R. Stubbs, Internal Revenue Service, and Robert D. Tortora, U.S. Department of Agriculture, all of whom are members of the Subcommittee on Nonsampling Errors, Federal Committee on Statistical Methodology. The Subcommittee was chaired by Monroe G. Sirken, National Center for Health Statistics, Department of Health, Education, and Welfare. Members of the full Subcommittee are listed on the following page.

The Subcommittee decided to prepare this glossary not as an exhaustive list of definitions of terms related to nonsampling errors, but rather to demonstrate the semantic problem, that certain terms may be used with different meanings, creating problems of communication among statisticians.

It is hoped that this glossary will enhance our awareness of the use of the same term for different concepts. The definitions included are taken mostly from published literature. The Federal Committee on Statistical Methodology views this as the first step in developing a more systematic approach to defining non-sampling error terms.

Subcommittee on Nonsampling Error

Monroe Sirken (Chair)
National Center for Health Statistics

Barbara Bailar
Bureau of the Census

Camilla Brooks
Bureau of the Census

John Cremeans
Bureau of Economic Analysis

Tore Dalenius (Consultant)
Brown University and
Stockholm University

Richard Deighton
U.S. Postal Service

James Duffet
U.S. Postal Service

Maria Gonzalez (ex officio)
Office of Federal Statistical
Policy and Standards

Tom Herzog
Social Security Administration

Alexander Korn
Bureau of Economic Analysis

Lillian Madow
Bureau of Labor Statistics

D. H. McElhone
Civil Service Commission

Ron Poland
U.S. Postal Service

Morton Raff
Bureau of Labor Statistics

Jack Scharff
Health Care Finance Administration

Frederick Scheuren
Social Security Administration

Otto Schwartz
Internal Revenue Service

Gary Shapiro
Bureau of the Census

Joel Stubbs
Internal Revenue Service

Robert Tortora
U.S. Department of Agriculture

Rolf Wulfsberg
National Center for Education
Statistics

Members of the Federal Committee on Statistical Methodology

(September 1977)

Maria Elena Gonzalez (Chair)
Office of Federal Statistical Policy and Standards
(Commerce)

Barbara A. Bailar
Bureau of the Census (Commerce)

Norman D. Beller
Economics, Statistics and Cooperatives Service
(Agriculture)

Barbara A. Boyes
Bureau of Labor Statistics (Labor)

Edwin J. Coleman
Bureau of Economic Analysis (Commerce)

John E. Cremeans
Bureau of Economic Analysis (Commerce)

Marie D. Eldridge
National Center for Education Statistics (HEW)

Fred Frishman
Internal Revenue Service (Treasury)

Thomas B. Jabine
Social Security Administration (HEW)

Charles D. Jones
Bureau of the Census (Commerce)

Alfred D. McKeon
Bureau of Labor Statistics (Labor)

Harold Nisselson
Bureau of the Census (Commerce)

Monroe G. Sirken
National Center for Health Statistics (HEW)

Wray Smith
Office of the Assistant Secretary for Planning and
Evaluation (HEW)

Authors' Preface

The glossary has been prepared for the Federal Committee on Statistical Methodology by the Subcommittee on Nonsampling Errors. During subcommittee meetings it became obvious that the terminology on nonsampling errors left something to be desired. Consequently, the subcommittee decided to develop this glossary. A search of current literature for acceptable definitions highlighted the prevailing problem instead of providing a solution to it.

Prior to examining the contents of this glossary, it is important to understand what the purpose of the glossary is and what it is not. The purpose of the glossary is to highlight a semantic problem represented by the fact that:

The same term is sometimes used with different meanings; and

The same phenomenon is sometimes called by more than one term.

The glossary contributes towards the task of developing a standardized terminology. However, it is beyond the scope of the present Subcommittee on Nonsampling Errors to pursue a task of this magnitude.

This glossary is not intended to be used as a dictionary for nonsampling error terms. The authors did not conduct an exhaustive search of the literature, nor did they attempt to select or specify a preferred definition for any term. The reference to a subcommittee document (O1 in the bibliography) is included solely for the convenience of the reader and does not exist as a separate document.

Terms in the glossary are listed alphabetically. The bibliography at the end contains the references which are used to obtain definitions for the terms. Codes for the references are alpha-numeric (i.e., A1, A2, A3, etc. designate the first, second, and third references associated with the first letter of the author's name). For each term defined in the glossary, one or more references are given. For example, on Page 1, the term "ACCURACY" is found in three references, viz., B2, K1, S1.

If a definition is taken from a textbook, the page number of the referenced text is included at the end of the definition. (See BIAS, CONSTANT—Reference H2, Page 17.) Quotation marks are used to identify when a definition was copied verbatim from a reference. The absence of quotation marks implies that the definition was paraphrased or that the definition was taken out of context. When a definition was taken out of context, an editors' note was often added to make the reader aware that the definition can be applied to a more general subject. The editors' note is enclosed in brackets [] (e.g., see BOUNDED RECALL). A "Comment" indicating similar terms which appear in the glossary has been added at the end of each relevant term.

Members of the Glossary Task Force wish to express their sincere appreciation to Professor Tore Dalenius for the criticism, suggestion, and reference material which he provided during the preparation of this glossary. Also, the members of the task force wish to thank all other members of the Subcommittee on Nonsampling Errors for their comments and suggestions on the numerous drafts of the glossary.

Table of Contents

Preface	ii
Authors' Preface	iv
List of terms included in the Glossary of Nonsampling Error Terms	1
Accuracy	1
Allocation	1
Audit	1
Bias	1
Bias, Constant	1
Bounded Recall	1
Bounding	1
Classification Errors	1
Classification Error Rate	2
Coding	2
Coding Error	2
Compiling Error	2
Complete Coverage	2
Completeness Rate	2
Conditioning Effect	2
Content Error	2
Correct Value	2
Coverage Error	2
Defined Goal	2
Definitional Errors	2
Editing	3
Editing Change	3
Error	3
Error Planting	3
Equal Complete Coverage	3
Expected Value	3
Follow-up	3
Frame	4
Gross Difference	4
Ideal Goal	4
Imputation	4
Incomplete Coverage	4
Index of Inconsistency	4
Interviewer Bias	4
Interviewer Error	4
Interviewer Variance	4
Item Nonresponse	4
Limits of Error	4
Mean Square Error	4
Measurement Error	5
Memory Error	5
Net Difference	5
Noncontact	5
Noncoverage	5
Noninterview	5
Noninterview Adjustment	5
Nonobservation Error	5

Nonrespondent	5
Nonresponse	5
Nonresponse Rate	5
Nonsampling Error	6
Not at Home	6
Observational Error	6
Overcoverage	6
Post-audit	6
Precision	6
Preferred Technique	7
Process Control	7
Quality Check	7
Quality Control	7
Recall	7
Recall Errors	7
Recall Losses	7
Recall Period	7
Record Check	7
Refusal Rate	7
Relevance	7
Relevance Error	7
Reliability	8
Response Bias	8
Response Deviation	8
Response Error	8
Response Rate	8
Response Variance	8
Response Variance, Correlated	8
Response Variance, Uncorrelated (Simple)	9
Rotation Bias	9
Sample Design	9
Sample Verification	9
Sampled Population	9
Sampling Bias	9
Sampling Error (of Estimator)	9
Sampling Variance	10
Specification	10
Specification Errors	10
Standard Error of Estimate	10
Statistical Audit (Control)	10
Survey Design	10
Survey Value	10
Systematic Error	10
Tabulation Errors	11
Target Population	11
Telescoping	11
Temporarily Absent	11
True Value	11
Unbounded Recall	11
Undercoverage	11
Validity	11
Variance, Interaction Term	11
Verification, Dependent	11
Verification, Independent	12
Working Technique	12
Bibliography	12

Glossary of Nonsampling Error Terms

ACCURACY

B2 "The quality of a survey result that is measured by the difference between the survey figure and the value being estimated. The true value is seldom known, although it can be approximated in some instances." p. 48.

K1 "Accuracy in the general statistical sense denotes the closeness of computations or estimates to the exact or true values. In a more specialized sense the word also occurs as meaning (a) in relation to an estimator, unbiasedness; (b) in relation to the reciprocal of the standard error, the precision (q.v.). Neither usage can be recommended."

S1 "Closeness to the true value."

Comment—See CORRECT VALUE, SURVEY VALUE, and TRUE VALUE.

ALLOCATION

O1 The process of assigning values to units in the nonresponse group of a survey according to the characteristics that have been observed for the response group or by any other imputation procedure.

AUDIT

O1 The process of applying more extensive methods of measurement to a subsample during the scheduled conduct of a survey in order to determine the effect of nonsampling errors.

Comment—See POST-AUDIT.

BIAS

B2 "The difference between the expected value of an estimator and the value that would be obtained from all the population elements with no corresponding errors of measurement being made. This true value is what we are trying to estimate." p. 48.

H2 The difference between the expected value of the estimator and the true value being estimated. Whenever the bias is 0, the estimator is said to be unbiased. p. 17.

K1 "Generally, an effect which deprives a statistical result of representativeness by systematically distorting it, as distinct from a random error

which may distort on any one occasion but balances out on the average."

Comment—See SYSTEMATIC ERROR.

BIAS, CONSTANT

C1 That component of the total bias in a survey estimator that affects all of the units alike. p. 389.

BOUNDED RECALL

D4 "An interview where the respondent is reminded of what he reported in an earlier interview and is then asked only to report on any new events that occurred subsequent to the bounding interview."

N1 A method of interview that is designed to prevent shifting in time of expenditures reported by respondents.

[Editor's Note—Definition was given for expenditures but may apply to other characteristics.]

S2 "Bounded recall procedures involve a series of interviews with the same panel of respondents. At the beginning of the bounded interview, which is the second or later interview, the respondent is told about the expenditures reported during the previous interview, and is then asked about additional expenditures made since then. The interviewer also checks the new expenditures reported with previous expenditures to make sure that no duplication has occurred." p. 83.

[Editor's Note—Definition was given for expenditures but may apply to other characteristics.]

BOUNDING

B2 "Prevention of erroneous shifts of the timing of events by having the enumerator or respondent supply at the start of the interview (or in a mail survey) a record of events reported in the previous interview." p. 48.

CLASSIFICATION ERRORS

H2 Errors caused by conceptual problems and misinterpretations in the application of classification systems to survey data. p. 84.

CLASSIFICATION ERROR RATE

- O1** The proportion of responses that have been incorrectly classified in a survey.

CODING

- W1** "Coding is a technical procedure for converting verbal information into numbers or other symbols which can be more easily counted and tabulated." p. 234.

CODING ERROR

- B3** Errors that occur during the coding of sample data.
- O1** The assignment of an incorrect code to a survey response.

COMPILING ERRORS

- H2** Errors introduced in operations on the original observations such as editing, coding, punching, tabulating and transcribing. p. 84.

COMPLETE COVERAGE

- M2** "A survey (or census) should be called complete if virtually all of the units in the population under study are covered." p. 54.
- Comment—See **COVERAGE ERROR** and **INCOMPLETE COVERAGE**.

COMPLETENESS RATE

- H6** The completeness rate is the percentage of interviews in which the required information is given by the respondent. This rate reflects in part interviewer effectiveness in the interview. However, it is not independent of the response rate. A low response rate may imply that the respondents interviewed are more likely to be cooperative than is the case with a high response rate. In a sense, the interviewer with a low response rate can be thought of as disposing of his uncooperative sample members at the door and interviewing only the relatively cooperative ones, therefore obtaining a higher completeness rate. p. 13.

CONDITIONING EFFECT

- B2** "The effect on responses resulting from the previous collection of data from the same respondents in recurring surveys." p. 48.

CONTENT ERROR

- B2** "Errors of observation or objective measurement, of recording, of imputation, or of other processing which results in associating a wrong value of the characteristic with a specified unit. (Coverage errors are excluded from this definition.)" p. 48.

CORRECT VALUE

- C1** The value obtained for a unit that is measured without error. p. 374.
- Comment — See **ACCURACY**, **SURVEY VALUE**, and **TRUE VALUE**.

COVERAGE ERROR

- B2** "The error in an estimate that results from (1) failure to include in the frame all units belonging to the defined population; failure to include specified units in the conduct of the survey (undercoverage), and (2) inclusion of some units erroneously either because of a defective frame or because of inclusion of unspecified units or inclusion of specified units more than once, in the actual survey (overcoverage)." p. 48.
- Comment—See **NONCOVERAGE**, **OVERCOVERAGE** and **UNDERCOVERAGE**.

DEFINED GOAL

- B2** "The approximation to the true value that would be obtained if the survey were carried out using the specified frame, the method of measurement for the specified characteristic and the method of summarizing the measurements as required in the survey plan." p. 48.
- H4** "Specifications actually set forth for the statistical survey, if carried out precisely and rigorously, would yield the defined goals."
- Comment—See **EXPECTED VALUE**, **IDEAL GOAL** and **SPECIFICATION**.

DEFINITIONAL ERRORS

- H2** Errors that occur in surveys whenever the definitions of the characteristics for which data are to be collected are not pertinent to the purposes of the survey or are not clear to the respondents. p. 83.

EDITING

P1 Identifying potential problems is the first objective of the editing process. We also classify as editing, operations performed on the recorded information that are designed to conform it to the desired format or units. Filling a blank on the basis of redundant information on the form is editing. Similarly, when a respondent indicates that he has reported in pounds whereas reporting in tons was requested, we would regard the conversion of his figure to the specified unit as an editing correction. We are also inclined to classify as editing the supplying of missing totals where the component detail has been reported.

W1 "Editing is a preliminary step in which the responses are inspected, corrected and sometimes precoded according to a fixed set of rules." p. 234.

Comment—See **EDITING CHANGE** and **IMPUTATION**.

EDITING CHANGE

P1 A code that is inserted on a form as a result of an editing process. For example, where a woman is coded as the "wife" of the head of the household and the field for marital status is blank, the code for "married" may be inserted as an editing change in this case.

Comment—See **EDITING** and **IMPUTATION**.

ERROR

K1 "In general, a mistake or error in the colloquial sense. There may, for example, be a gross error or avoidable mistake; an error of reference, when data concerning one phenomenon are attributed to another; copying errors; an error of interpretation.

"In a more limited sense the word 'error' is used in statistics to denote the difference between an occurring value and its 'true' or 'expected' value. There is here no imputation of mistake on the part of a human agent; the deviation is a chance effect. In this sense we have, for example, errors of observation (q.v.), errors

in equations (q.v.), errors of the first and second kinds (q.v.) in testing hypotheses, and the error band (q.v.) surrounding an estimate; and also the normal curve of errors itself."

ERROR PLANTING

D1 A method of control where a set of errors is introduced into the material being subjected to control. If the control were perfect, all the planted errors would be detected. In practice, only a fraction is detected. This fraction may obviously be used as a measure of the performance of the control operation. p. 153.

EQUAL COMPLETE COVERAGE

D3 "The equal complete coverage is by definition the result that would be obtained from investigation of all the sampling units in the frame (segments of area, business establishments, accounts, manufactured articles) by the same field-workers or inspectors, using the same definitions and procedures, and exercising the same care as they exercised on the sample, and at about the same period of time. The concept of the equal complete coverage is fundamental to the use of samples. The adjective equal signifies that the same methods must be used for the equal complete coverage as for the sample. Every sample is a selected portion of the sampling units in the frame; hence **A SAMPLE IS A SELECTED PORTION OF RESULTS OF THE EQUAL COMPLETE COVERAGE.**"

Comment—See **COVERAGE ERROR**.

EXPECTED VALUE

H4 The hypothetical averages from the conceived replicates of the survey all conducted under the same essential conditions.

Comment—See **DEFINED GOAL**, **IDEAL GOAL**, and **SPECIFICATION**.

FOLLOW-UP

D4 "A procedure whereby those members of a selected sample for whom a response is not obtained by one data collection strategy (e.g., telephone or mail) are contacted by the same or another data collection strategy in order to in-

crease response rate. It can also be used to designate repeated surveys among a panel of respondents."

FRAME

- C1** A list of the units which make up the population. p. 7.
- K2** "Physical lists and procedures that can account for all the sampling units without the physical effort of actually listing them." p. 53.
- U1** The frame consists of previously available descriptions of the objects or material related to the physical field in the form of maps, lists, directories, etc., from which sampling units may be constructed and a set of sampling units selected; and also information on communications, transport, etc., which may be of value in improving the design for the choice of sampling units, and in the formation of strata, etc. p. 7.
Comment—See **SAMPLED POPULATION** and **TARGET POPULATION**.

GROSS DIFFERENCE

- B4** The number of cases that are classified differently in the initial survey or census and its replication. p. 2.

IDEAL GOAL

- H4** The set of statistics that would have been produced had all of the requirements been precisely defined and rigorously met constitutes the ideal goal of the statistical survey.

Comment—See **DEFINED GOAL**, **EXPECTED VALUE**, and **SPECIFICATION**.

IMPUTATION

- P1** The process of developing estimates for missing or inconsistent data in a survey. Data obtained from other units in the survey are usually used in developing the estimate.

Example: An editing test classifies an age as wrong when a man is reported as 6 years of age, and also as head of the household, with a wife age 35 and a child age 10. A more rational figure than the 6 is supplied by some procedure such as using the same age difference between

husband and wife as appeared in the preceding household of similar type.

Comment—See **EDITING** and **EDITING CHANGE**.

INCOMPLETE COVERAGE

- M2** A survey (or census) should be called incomplete if a substantial number of the units in the population under study are arbitrarily excluded. p. 54.

Comment—See **COMPLETE COVERAGE** and **COVERAGE ERROR**.

INDEX OF INCONSISTENCY

- B1** The proportion of the total variance of a characteristic that is accounted for by the response variance.

INTERVIEWER BIAS

- K1** Bias in the responses which is the direct result of the action of the interviewer.

INTERVIEWER ERROR

- O1** Errors in the responses obtained in a survey that are due to actions of the interviewer.

INTERVIEWER VARIANCE

- B2** "That component of the nonsampling variance which is due to the different ways in which different interviewers elicit or record responses." p. 48.

ITEM NONRESPONSE

- B5** "The type of nonresponse in which some questions, but not all, are answered for a particular unit." p. 914.

- O1** The type of nonresponse in which a question is missed for an interviewed unit.

LIMITS OF ERROR

- D3** The limits of error are the maximum overestimate and the maximum underestimate from the combination of the sampling and the nonsampling errors.

MEAN SQUARE ERROR

- K1** "The second moment of a set of observations about some arbitrary origin. If that origin is the

mean of the observations, the mean-square deviation is equivalent to the variance (q.v.)."

S1 "Mean value over trials of the square of the response error. It may be expressed as the sum of variances, covariances and the square of the response bias."

O1 The variance of the estimate plus the bias squared.

MEASUREMENT ERROR

B2 "(1) As applied to individual units of analysis, measurement error means the difference between the observed or imputed value and the true value. (2) As applied to an estimate, measurement error means the difference between the estimate and the true value, thus including all sampling as well as nonsampling errors." p. 48.

MEMORY ERROR

M2 Errors associated with the recall of answers to questions about the past. p. 45.

NET DIFFERENCE

B4 The net difference of a tabulated figure for a given class is the difference between the total for the class obtained in the reinterview (or appropriate records) and the original survey. p. 3.

NONCONTACT

O1 A type of nonresponse in which the interviewer has not been able to contact the respondent.

NONCOVERAGE

C1 "Failure to locate or to visit some units in the sample." p. 360.

K2 "Refers to the negative error of failure to include elements that would properly belong in the sample." p. 529.

Comment — See **COVERAGE ERROR**, **OVERCOVERAGE**, and **UNDERCOVERAGE**.

NONINTERVIEW

O1 The type of nonresponse in which no information is available from occupied sample units for such reasons as: not at home, refusals, incapacity and lost schedules.

NONINTERVIEW ADJUSTMENT

O1 A method of adjusting the weights for interviewed units in a survey to the extent needed to account for occupied sample units for which no information was obtained.

NONOBSERVATION ERROR

K2 "Failure to obtain data from parts of the survey population which results from two sources: non-coverage and nonresponse." p. 527.

NONRESPONDENT

M2 Those persons in a sample from whom information has not been obtained. p. 172.

NONRESPONSE

B2 "The failure to elicit responses for units of analysis in a population or sample because of various reasons such as absence from home, failure to return questionnaires, refusals, omission of one or more entries in a form, vacant houses, etc." p. 50.

C1 "We shall use the term nonresponse to refer to the failure to measure some of the units in the selected sample." p. 355.

K1 "In sample surveys, the failure to obtain information from a designated individual for any reason (death, absence, refusal to reply) is often called a nonresponse and the proportion of such individuals of the sample aimed at is called the nonresponse rate. It would be better, however, to call this a 'failure' rate or a 'non-achievement' rate and to confine 'nonresponse' to those cases where the individual concerned is contacted but refuses to reply or is unable to do so for reasons such as deafness or illness."

K2 "Nonresponse refers to many sources of failure to obtain observations (responses, measurements) on some elements selected and designated for the sample." p. 532.

NONRESPONSE RATE

D4 "The complement of response rate. The numerator is those eligible respondents selected in a sample for whom information is not obtained because of refusals, not found at home, unavailable by reason of illness, incompetence, lan-

guage difficulty, etc. The denominator is the total number of eligible respondents initially selected for the sample." p. 46.

NONSAMPLING ERROR

B2 "The error in an estimate arising at any stage in a survey from such sources as varying interpretation of questions by enumerators, unwillingness or inability of respondents to give correct answers, nonresponse, improper coverage, and other sources exclusive of sampling error. This definition includes all components of the Mean Square Error (MSE) except sampling variance." p. 50.

K1 "An error in sample estimates which cannot be attributed to sampling fluctuations. Such errors may arise from many different sources such as defects in the frame, faulty demarcation of sample-units, defects in the selection of sample-units, mistakes in the collection of data (due to personal variations or misunderstandings or bias or negligence or dishonesty on the part of the investigator or of the interviewee), mistakes at the stage of the processing of the data, etc.

"The term 'response error' is sometimes used for mistakes in the collection of data and would not, strictly speaking, cover errors due to nonresponse. The use of the word 'bias' in the place of error, e.g. 'response bias' is not uncommon. The term 'ascertainment error' (Mahalanobis) is preferable as it would include errors due to nonresponse and also cases of collection of data by methods other than interviewing, e.g. direct physical observation of fields for crop estimates."

Comment—See **OBSERVATIONAL ERROR** and **RESPONSE ERROR**.

NOT AT HOME

C1 "Persons who reside at home but are temporarily away from the house." p. 360.

OBSERVATIONAL ERROR

K1 "This term ought to mean an error of observation but sometimes occurs as meaning a response error."

K2 "Errors which are caused by obtaining and re-

coding observations incorrectly." p. 520.

Comment—See **NONSAMPLING ERROR** and **RESPONSE ERROR**.

OVERCOVERAGE

K2 "Positive errors which occur due to the inclusion in the sample of elements that do not belong there." p. 529.

Comment—See **COVERAGE ERROR**, **NONCOVERAGE**, and **UNDERCOVERAGE**.

POST-AUDIT

O1 The process of applying more extensive methods of measurement to a subsample after the scheduled conduct of a survey in order to determine the effect of nonsampling errors.

Comment—See **AUDIT**.

PRECISION

B2 "The quality of a sample result that is measured by the difference between the sample result and the result which would be obtained if a complete count were taken using the same survey procedures. Same as reliability. Usually defined by stating the sampling error." p. 50.

C1 Refers to the size of deviations from the mean obtained by repeated application of the sampling procedure. p. 16.

H2 "The difference between a sample result and the result of a complete count taken under the same conditions . . . or the reliability." p. 10.

H3 A measure of how close the set of possible sample estimates for a particular sample design may be expected to come to the value being estimated. p. 7.

K1 "In exact usage precision is distinguished from accuracy. The latter refers to closeness of an observation to the quantity intended to be observed. Precision is a quality associated with a class of measurements and refers to the way in which repeated observations conform to themselves; and in a somewhat narrower sense refers to the dispersion of the observations, or some measure of it, whether or not the mean value around which the dispersion is measured approximates to the 'true' value. In general the

precision of an estimator varies with the square root of the number of observations upon which it is based."

PREFERRED TECHNIQUE

D3 "Any result, whatever it be, is the result of applying some set of operations. Although there is no true value, we do have the liberty to define and to accept a specified set of operations as preferred, and the results thereof as a master standard (so-called by Harold D. Dodge). Thus, there may be, by agreement of the experts in the subject-matter for any desired property of the material, a preferred survey-technique."

Comment—See **WORKING TECHNIQUE**.

PROCESS CONTROL

B2 A statistical quality control technique where frequent small samples are taken and evaluated to control clerical operations. p. 8.

QUALITY CHECK

M2 "An intensive study of a small sample (relative to the size of the survey) where every effort is made to attain the highest level of accuracy possible." p. 396.

QUALITY CONTROL

B2 "Observation and procedure used in any operation of a survey in order to prevent or reduce the effect of nonsampling errors." p. 50.

K1 "A method of controlling the quality of a manufactured product which is produced in large numbers. It aims at tracing and eliminating systematic variations in quality, or reducing them to an acceptable level, leaving the remaining variation to chance. The process is then said to be statistically under control."

RECALL

N1 A method of obtaining information by means of an interview in which the respondent is required to remember past events. A common application is the recall of consumer expenditures.

Comment—See **BOUNDED RECALL** and **UNBOUNDED RECALL**.

RECALL ERRORS

H2 "Many questions in surveys refer to happenings or conditions in the past, and there is a problem in both remembering the event and of associating it with the correct time period." p. 84.

RECALL LOSSES

N1 Omissions of expenditures due to forgetting of items.

[Editor's Note—Definition was given for expenditures but may apply to other characteristics.]

RECALL PERIOD

N1 Refers to the period of time for which the respondent's report of expenditures is to be utilized.

[Editor's Note—Definition was given for expenditures but may apply to other characteristics.]

RECORD CHECK

B2 "A study in which data on individual units obtained by one method of data collection are checked against data for the same units from available records obtained by a different method of data collection (for example, comparison of ages reported in census with information from birth certificates)." p. 50.

REFUSAL RATE

K1 "In the sampling of human populations, the proportions of individuals who, though successfully contacted, refuse to give the information sought. The proportion is usually (and preferably) calculated by dividing the number of refusals by the total number of the sample which it was originally desired to achieve."

RELEVANCE

H4 "Standards of relevance are concerned with the difference between the ideal goal of a survey and the statistics called for by the survey specifications."

RELEVANCE ERROR

O1 The difference between the ideal goal of a sur-

vey and the statistics called for by the survey specifications.

Comment—See RELEVANCE.

RELIABILITY

M1 The confidence that can be assigned to a conclusion of a probabilistic nature.

[Editor's Note—Translation taken from Crespo, see Reference S1.]

M2 "The extent that repeat measurements made by a scale or test under constant conditions will give the same result (assuming no change in the basic characteristic—e.g. attitude—being measured)." p. 353.

S1 "The degree of confidence in terms of probability associated with conclusions based on a random experiment."

Comment—See VALIDITY.

RESPONSE BIAS

B3 The difference between the average of the averages of the responses over a large number of independent repetitions of the census and the unknown average that could be measured if the census were accomplished under ideal conditions and without error. p. 1.

S1 "Difference between average reported value over trials and true values. It is a combined bias as algebraic sum of all bias terms representing diverse source of biases."

RESPONSE DEVIATION

B3 "The difference between the response recorded for a person on a particular trial and the average of the responses over all trials for the same person." p. 2.

S1 "Difference between individual reported value and the average over hypothetical trials under the same general conditions."

RESPONSE ERROR

B2 "That part of the nonsampling error which is due to the failure or the respondent to report the correct value (respondent error) or the interviewer to record the value correctly (interviewer error). It includes both the consistent

response biases and the variable errors of response which tend to balance out." p. 50.

S1 "Difference between reported and true value."

Comment—See NONSAMPLING ERROR and OBSERVATIONAL ERROR.

RESPONSE RATE

D4 "The percentage of an eligible sample for whom information is obtained. For an interview survey the numerator of the formula is the number of interviews. The denominator is the total sample size minus non-eligible respondents; that is, minus those not meeting the criteria for a potential respondent as defined for that particular study."

H6 "The percentage of times an interviewer obtains interviews at sample addresses where contacts are made, i. e.,

$\frac{\text{Number of contacts}}{\text{Number of interviews}}$. Since a contact must be

either an interview or a refusal, the response rate is also equal to $1 - \frac{\text{Number of refusals.}}{\text{Number of contacts}}$

p. 13.

W1 The response rate is the proportion of the eligible respondents in the sample who were successfully interviewed. For example, the denominator may be the total number of occupied dwellings, and the numerator may be the number of completed interviews. p. 294.

RESPONSE VARIANCE

B2 "That part of the response error which tends to balance out over repeated trials or over a large number of interviewers." p. 50.

B3 "The variance among the trial means over a large number of trials." p. 2.

D4 "The response variance of a survey estimator is the sum of the simple response variance and the correlated response variance."

RESPONSE VARIANCE, CORRELATED

D4 "The correlated response variance is the contribution to the total variance arising from non-zero correlations (in the sense of the distribution of measurement errors) between the response of sample units."

H1 The contribution to the total response variance from the correlations among response deviations.

RESPONSE VARIANCE, UNCORRELATED (SIMPLE)

D4 "The sample response variance contribution to the total variance arises from the variability of each survey response about its own expected value. In terms of a simple random sampling design, the simple response variance is the population mean of the variances of each population unit."

H1 "The variance of the individual response deviations over all possible trials."

H5 "The basic trial-to-trial variability in response, averaged over the elements in the population." p. 116.

S1 "Variance of the reported value over trials."

ROTATION BIAS

O1 A type of bias that occurs in panel surveys which consist of repeated interviews on the same units. Although these surveys are designated so that the estimates of a characteristic are expected to be nearly the same for each panel in the survey, this expectation has not been realized. For example, an estimate from a panel that is in the survey for the first time may differ significantly from estimates from the panels that have been in the survey longer.

Z1 "The downward tendency in the value of the characteristics reported if the observation of the same units is continued over a longer period of time. For example, it was found in expenditure surveys that the average expenditure per item per person is usually higher in the first week of the survey than in the second or the third." p. 203.

SAMPLE DESIGN

H3 A procedure that consists of a sampling plan and method of investigation. p. 7.

K1 "The usage is not uniform as regards the precise meaning of this and similar terms like 'sample plan,' 'survey design,' 'sample plan' or

'sampling design.' These cover one or more parts constituting the entire planning of a (sample) survey inclusive of processing, etc. The term 'sampling plan' may be restricted to mean all steps taken in selecting the sample; the term 'sample design' may cover in addition the method of estimation; and 'survey design' may cover also other aspects of the survey, e.g. choice and training of interviewers, tabulation plans, etc. 'Sample design' is sometimes used in a clearly defined sense, with reference to a given frame, as the set of rules or specifications for the drawing of a sample in an unequivocal manner."

Comment—See SURVEY DESIGN.

SAMPLE VERIFICATION

H2 A quality control procedure for keeping certain clerical errors at a satisfactory level. p. 618.

SAMPLED POPULATION

C1 "The population to be sampled." p. 6.

Comment—See FRAME and TARGET POPULATION.

SAMPLING BIAS

B2 "That part of the difference between the expected value of the sample estimator and the true value of the characteristic which results from the sampling procedure, the estimating procedure, or their combination." p. 50.

SAMPLING ERROR (OF ESTIMATOR)

B2 "That part of the error of an estimator which is due to the fact that the estimator is obtained from a sample rather than a 100 percent enumeration using the same procedures. The sampling error has an expected frequency distribution for repeated samples, and the sampling error is described by stating a multiple of the standard deviation of this distribution." p. 50.

K1 "That part of the difference between a population value and an estimator thereof, derived from a random sample, which is due to the fact that only a sample of values is observed; as distinct from errors due to imperfect selection, bias in response or estimation, errors of obser-

vation and recording, etc. The totality of sampling errors in all possible samples of the same size generates the sampling distribution of the statistic which is being used to estimate the parent value."

SAMPLING VARIANCE

D4 "The sampling variance is that contribution to the total variance arising from the random selection of a sample, rather than a complete enumeration, from the population." p. 45.

H1 The component of the total variance of the survey that represents the contribution due to sampling.

SPECIFICATION

H4 Detailed description of the collection, compilation and presentation of the survey data.

Comment—See **DEFINED GOAL**, **EXPECTED VALUE** and **IDEAL GOAL**.

SPECIFICATION ERRORS

M3 Errors at the planning stage because (i) data specification is inadequate and inconsistent with respect to the objectives of the survey, (ii) omission or duplication of units, incomplete units or faulty enumeration methods and (iii) inaccurate or inappropriate methods of interview. p. 451.

STANDARD ERROR OF ESTIMATE

B2 "This term refers to the sampling error, calculated as the square root of the variance of the estimator." p. 50.

K1 "An expression for the standard deviation of the observed values about a regression line, i.e. an estimator of the variation likely to be encountered in making predictions from the regression equation. For example, in simple linear regression of y on x the standard error of estimate of y is given by $\sigma_y (1 - r^2)^{1/2}$ where σ_y^2 is the variance of y and r is the correlation between y and x ."

STATISTICAL AUDIT (CONTROL)

D2 A procedure to detect the existence of errors that are made in carrying out the fieldwork, the

interviewing, the coding, the computations, and other work. p. 71.

SURVEY DESIGN

H3 "By the survey design will be meant the sample design together with the questionnaire and the method of obtaining the information from the sample, or, more generally, the method of measurement. Thus, the survey design includes the plans for all the parts of the survey except the statement of the objectives. It includes:

- (a) The questionnaire,
- (b) Decision on method of observation or interview,
- (c) Sample design,
- (d) Choice and training of interviewers,
- (e) Assignments of interviewers,
- (f) Decisions on treatment of noninterviews,
- (g) Estimation equations,
- (h) Processing of questionnaires,
- (i) Preparation of tables,
- (j) Studies of precision and accuracy of information,

as well as instructions and methods followed for carrying through these operations." p. 8.

Comment—See **SAMPLE DESIGN**.

SURVEY VALUE

B5 "A value obtained in a complete survey which is intended to be the 'true' value, but which may not be the same because the 'true' data cannot be collected, the population cannot be defined exactly, or there are uncontrollable biases in the process of collecting and assembling the data. For example, age may be poorly reported if someone other than the person involved responds; sometimes the person himself does not know his age." p. 913.

Comment—See **ACCURACY**, **CORRECT VALUE** and **TRUE VALUE**.

SYSTEMATIC ERROR

K1 "As opposed to a random error, an error which is in some sense biased, that is to say, has a distribution with mean (or some equally acceptable measure of location) not at zero."

Comment—See **BIAS**.

TABULATION ERRORS

M3 Errors occurring during the tabulation stage of survey procedures. p. 451.

TARGET POPULATION

C1 "The population about which information is wanted." p. 6.

Comment—See **FRAME** and **SAMPLED POPULATION**.

TELESCOPING

R1 The tendency of the respondent to allocate an event to a period other than the reference period (also called border bias). p. 211.

S2 "A telescoping error occurs when the respondent misremembers the duration of an event. While one might imagine that errors would be randomly distributed around the true duration, the errors are primarily in the direction of remembering an event as having occurred more recently than it did. This is due to the respondent's wish to perform the task required of him. When in doubt, the respondent prefers to give too much information rather than too little." p. 69.

TEMPORARILY ABSENT

O1 A sampling unit for which a respondent cannot be contacted during the survey period.

TRUE VALUE

B5 "An idealized concept of a quantity which is to be measured; in some cases it can be achieved, but in others there is disagreement as to the definition of the quantity. Illustrations are the number of persons who are 'unemployed,' and the 'dollar value of farm sales.'

"In most surveys an approximation to the 'true' value is used, defined in such a way that one would expect to be able to measure it provided there were sufficient time, money, knowledge of techniques, etc., and no errors in the reporting, collection, and processing of the data." p. 913.

D2 A population value determined by a specified set of operations that one preferred, and the results thereof as a master standard. p. 62.

S1 That result which would be obtained with perfect measuring instruments and without committing any error of any type both in collecting the primary data and in carrying out mathematical operations.

Comment—See **ACCURACY**, **CORRECT VALUE** and **SURVEY VALUE**.

UNBOUNDED RECALL

N1 "Ordinary type of recall, where respondents are asked for expenditures made since a given date and no control is exercised over the possibility that respondents may erroneously shift some of their expenditures reports into or out of the recall period."

[Editor's Note—Definition was given for expenditures but may apply to other characteristics.]

Comment—See **BOUNDED RECALL** and **RECALL**.

UNDERCOVERAGE

B2 "The error in an estimate that results from failure to include in the frame all units belonging to the defined population." p. 48.

O1 A type of nonsampling error that results from either failure to include all appropriate sampling units in the frame or failure to include some of the units that are already on the frame.

Comment—See **COVERAGE ERROR**, **OVERCOVERAGE** and **NONCOVERAGE**.

VALIDITY

D4 "A valid measure is one that measures what it claims to and not something else. Validity is a continuous concept so most measures fall between total validity and total nonvalidity. A totally valid measure is one without bias."

VARIANCE, INTERACTION TERM

D4 "The interaction contribution to the total variance of estimate is that component arising from a non-zero covariance between measurement error and sampling error."

VERIFICATION, DEPENDENT

B6 A method of verifying coding quality in which

high level clerks review the work of production coders and determine whether or not the codes assigned are correct.

VERIFICATION, INDEPENDENT

- B6** A method of verifying coding quality in which two or more independent codings of items are conducted for an identical sample of persons and then the coding results are matched.

WORKING TECHNIQUE

- D3** "Unfortunately, it often happens that the preferred technique, usable on a laboratory-scale, is too expensive to apply in a full-scale survey, or it may be objectionable otherwise. Experts in the subject-matter must then supply also a

working technique. Thus, the preferred technique by which to define a person's age might be to compute the difference in time between today and the date shown on his birth-certificate. But some people don't have birth-certificates at all, and few people have them handy. Moreover, some people would not be happy with an interviewer who asked for birth-certificates. The Passport Division can ask for birth-certificates, but interviewers may only ask the person how old he is, and record the result. This would be the working technique by which to measure age."

Comment—See PREFERRED TECHNIQUE.

Bibliography

- B1.** Bershad, Max A. (1964), *The Index of Inconsistency for L-fold Classification System*, Unpublished Bureau of the Census Memorandum, Washington, D.C.
- B2.** Bureau of the Census, *Course on Nonsampling Errors, Lectures 1-9*, International Statistics Program Center, Washington, D.C.
- B3.** Bureau of the Census (1963), *Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: Background, Procedures, and Forms*, Series ER60, No. 1, Washington, D.C.
- B4.** Bureau of the Census (1968), *The Current Population Survey Reinterview Program, January 1961 through December 1966*, Technical Paper No. 19, Washington, D.C.
- B5.** Bureau of the Census (1976), *The Statistical Abstract of the United States*, Washington, D.C.
- B6.** Bureau of the Census (1965), *United States Census of Population and Housing 1960: Quality Control of Preparatory Operations, Microfilming and Coding*, Washington, D.C.
- C1.** Cochran, W. G. (1963), *Sampling Techniques*, J. Wiley and Sons, Inc. New York.
- D1.** Dalenius, Tore (1974), *Ends and Means of Total Survey Design*, University of Stockholm, Stockholm.
- D2.** Deming, W. Edwards (1960), *Sample Design in Business Research*, J. Wiley and Sons, Inc., New York.
- D3.** Deming, W. Edwards (1960), *Uncertainties in Statistical Data, and Their Relation to the Design and Management of Statistical Surveys and Experiments*, Bulletin of the International Statistical Institute, Tokyo.
- D.4** Department of Health, Education and Welfare (1975), *Advances in Health Survey Research Methods*, Research Proceedings Series, DHEW Publication No. (HRS) 77-3154, Washington, D.C.
- H1.** Hansen, M. H., Hurwitz, W. N., and Bershad, M. A. (1960), *Measurement Errors in Censuses and Surveys*, Bulletin of the International Statistical Institute, Tokyo.

- H2. Hansen, M. H., Hurwitz, W. N., and Madow, W. G. (1953) *Sample Survey Methods and Theory, Volume I., Methods and Applications*, J. Wiley and Sons, Inc., New York.
- H3. Hansen, M. H., Hurwitz, W. N., and Madow, W. G. (1953), *Sample Survey Methods and Theory, Volume II, Theory*, J. Wiley and Sons, Inc., New York.
- H4. Hansen, M. H., Hurwitz, W. N., and Pritzker, L. (1967), *Standardization of Procedures for the Evaluation of Data: Measurement Errors and Statistical Standards in the Bureau of the Census*, Bulletin of the International Statistical Institute, Sydney.
- H5. Hansen, M. H., Hurwitz, W. N., and Pritzker, L. (1964), *The Estimation and Interpretation of Gross Differences and the Simple Response Variance*, Contributions to Statistics, Statistical Publishing Society, Calcutta.
- H6. Hauck, Mathew and Steinkamp, Stanley (1964), *Survey Reliability and Interviewer Competence*, Studies in Consumer Savings, No. 4, Bureau of Economic and Business Research, University of Illinois, Urbana.
- K1. Kendall, M. G. and Buckland, W. R. (1960), *A Dictionary of Statistical Terms*, Oliver and Boyd, Edinburgh.
- K2. Kish, Leslie (1965), *Survey Sampling*, J. Wiley and Sons, Inc., New York.
- M1. Morice, E. (1968), *Dictionnaire de Statistique*, Paris.
- M2. Moser, C. A. and Kalton, G. (1971), *Survey Methods in Social Investigation*, Second Edition, Basic Books, Inc., New York.
- M3. Murthy, M. (1967), *Sampling Theory and Methods*, Statistical Publishing Society, Calcutta.
- N1. Neter, J. and Waksberg, J. (1965), *Response Errors in Collection of Expenditures Data by Household Interviews: An Experimental Study*, Technical Paper No. 11, Bureau of the Census, Washington, D.C.
- O1. OMB Subcommittee on Nonsampling Errors (1976), Washington, D.C.
- P1. Pritzker, L., Ogus, J., and Hansen, M. H. (1965), *Computer Editing Methods—Some Applications and Results*, Bulletin of the International Statistical Institute, Belgrade.
- R1. Raj, D. (1972), *The Design of Sample Surveys*, McGraw-Hill Book Company, New York.
- S1. Sanchez-Crespo, J. L. (1975), *Notes on the Accuracy, Precision and Reliability of Statistical Data*, Bulletin of the International Statistical Institute, Warsaw.
- S2. Sudman, Seymour and Bradburn, Norman (1974), *Response Effects in Surveys*, Aldine Publishing Company, Chicago.
- U1. United Nations (1964), *Recommendations for the Preparation of Sample Survey Reports (Provisional Issue)*, Series C, No. 1, Rev. 2., New York.
- W1. Warwick, Donald P. and Lininger, Charles A. (1975), *The Sample Survey: Theory and Practice*, McGraw-Hill Book Company, New York.
- Z1. Zarkovich, S. S. (1969), *Quality of Statistical Data*, Food and Agriculture Organization of the United Nations, Rome.