Encyclopedia of Survey Research Methods

Total Survey Error (TSE)

Contributors: Marek Fuchs
Editors: Paul J. Lavrakas
Book Title: Encyclopedia of Survey Research Methods
Chapter Title: "Total Survey Error (TSE)"
Pub. Date: 2008
Access Date: October 15, 2013
Publishing Company: Sage Publications, Inc.
City: Thousand Oaks
Print ISBN: 9781412918084
Online ISBN: 9781412963947
DOI: http://dx.doi.org/10.4135/9781412963947.n585
Print pages: 897-903

http://dx.doi.org/10.4135/9781412963947.n585

Total survey error (TSE) is a term that is used to refer to all sources of bias (systematic error) and variance (random error) that may affect the validity (accuracy) of survey data. Total error in surveys can be conceptualized and categorized in many ways. One traditional approach is dividing total error into sources of sampling error and sources of nonsampling error. Another categorization is dividing it between coverage error, sampling error, nonresponse error, and measurement error. A more modern approach is to group various sources of error into the classes of *representation* and *measurement*. This entry provides a big picture perspective on all of the major types of error that occur in surveys and thus comprise total survey error.

Unfortunately, there is no such thing as a survey without error. Nevertheless, survey methodologists and survey practitioners aim for the most accurate surveys that can be conducted given the finite budget **[p. 897 ↓]** available to fund them. The quality of a survey statistic such as a mean, a percentage, or a correlation coefficient is assessed by multiple criteria: the timeliness of reporting, the relevance of the findings, the credibility of researchers and results, and the accuracy of the estimates—just to mention a few. Among those criteria the accuracy of the estimate is not necessarily the most important one. However, the accuracy is a dimension of the overall survey quality for which survey methodology offers a wide range of guidelines and instructions. Also, standard measures for the magnitude of the accuracy are available. The accuracy of a survey statistic is determined by its distance to or deviation from the true population parameter. If, for example, a survey aims to determine the average household income in a certain population, any deviation of the sample estimate from the true value— that is, what would have been determined if all members of the target population were asked their income and they all answered accurately—would decrease accuracy.

Representation And Measurement

There are two types of survey error that harm the accuracy of a survey estimate: random error and systematic error. Whereas random errors are assumed to cancel out each other—that is, negative deviations of the measurement from the true value are compensated by an "equal" amount of positive deviations— systematic errors shift the sample estimate systematically away from the true value; for example, because of certain question wording, respondents in a survey may tend to report a higher number of doctor visits than actually occurred in a given reference period. For linear estimates (such as means, percentages, and population totals), an increase in the random error leads to an increase in variance, whereas a rise in any systematic error results in an ascended bias of the survey estimate. The accuracy of a survey estimate is affected by either an increase of the bias or by a rise of the variance.

In a traditional view, the driving factors or sources of those survey errors are differentiated into two groups: sampling error and nonsampling error. Non-sampling error would then be further differentiated into coverage error, nonresponse error, and measurement error—some older textbooks mention processing error as well. However, a more modern theory-driven approach differentiates *observational errors* and *nonobservational errors*. While observational errors are related to the measurement of a particular variable for a particular sample unit, nonobservational errors occur when a net sample is established that is supposed to represent the target population. Following this path, Robert M. Groves and his colleagues have grouped the sources of error into two primary classes: *representation* and *measurement:*

This extension of the traditional total survey error concept provides room for a detailed analysis of the mechanisms by considering several sources of error at the same time, including possible interaction effects between the sources.

Total Survey Error Components Affecting Representation

Coverage Error

For a sample to be drawn, a sampling frame is necessary in order to provide the researcher with access to the members of the population from whom data are to be gathered. The incompleteness of this frame and the possible bias of its composition cause misrepresentations of the population by the sample. If a group is underrepresented in the frame—for example, individuals who own mobile phones as their only telecommunications device are missing from traditional random-digit dialing (RDD) sampling frames because they do not have a landline telephone number—the sociodemo-graphic or substantive characteristics of this group cannot be considered when computing the survey statistic. This causes a lack of accuracy of the survey estimate since some groups might be underrepresented in the frame and subsequently in any sample that is selected from the frame, resulting in coverage bias.

Sampling Error

Once a frame is available, a random sample needs to drawn: for example, a simple random sample, a stratified sample, a cluster sample, or a more complex **[p. 898 ↓]** sample design. Based on this sample, the standard error is computed by taking the square root of the division of the variance in the sample and the number of cases in the sample. The standard error is then used to compute the confidence limits and the margin of error—both are indicators for the precision of the estimate. Accordingly, the magnitude of the sampling error is one key component of the total survey error. It depends heavily on the design of the sample. For a fixed number of sample cases, the standard error usually decreases if stratification is applied. By contrast, a clustered sample is generally characterized by a larger variance which, in turn, raises the sampling error for a particular estimate. However, within a fixed budget, clustering usually increases precision, because the effective sample size can be increased even though the variance suffers from the inflationary design effect (i.e. *deff*) caused by clustering.

Nonresponse Error

Unit nonresponse error is the facet that is the best studied among all bias components within the TSE framework. Since the early days of survey methodology, researchers have been aware of the fact that some portions of the gross sample cannot be reached in the field phase of a survey or are not willing to comply with the survey request for cooperation. Because the responses of those groups may differ considerably from the responses of those members of the gross sample who can be reached and who are willing to cooperate, unit nonresponse is considered a serious source of systematic error that generates nonresponse bias. The literature provides comprehensive theoretical approaches to explain the various stages of respondent cooperation and also findings that can be generalized beyond particular surveys. In part, this is due to the fact that a potential nonresponse bias can be assessed for variables where parameters are available from official statistics (e.g., household income). Compared to other sources of error, this leaves survey researchers in a comfortable situation, as a possible bias can be observed more easily and taken into consideration when survey findings are interpreted.

Adjustment Error

Finally, the net sample needs to be adjusted for design effects introduced by the sample design. If the sample design, for example, would require a disproportional stratified sample, an appropriate weighting procedure would have to be devised to compensate for the unequal selection probabilities when estimating the population parameter. In addition, and as noted earlier in this entry, the net sample may need to be adjusted for possible nonresponse bias. Both procedures require complex computations that take into account information from the gross sample, official statistics, or both. Whereas the first approach may potentially increase the random error of the estimate, the second approach may introduce systematic errors into the sample and thus bias the estimate.

Total Survey Error Components Affecting Measurement

The four sources of error discussed so far were related to the representation of the target population by the weighted net sample. Coverage error, sampling error, nonresponse error, and adjustment error all potentially contributed to the random error or systematic error of the survey estimate. The next three sources of error—specification error, measurement error, and processing error—are concerned with the measurement process.

Specification Error

Most concepts of interest in surveys cannot be observed directly. Instead, the measurement process requires researchers to operationalize and translate the construct into questionnaire items that can be asked by interviewers and answered by respondents. For example, the general public's attitudes about illegal immigration ideally should be decomposed into several questionnaire items about the various dimensions of illegal immigration. Respondents then would be asked to report attitudes on each of these items. The combined score of all items would then be treated as a scaled measurement of the attitudes toward illegal immigration. If an important aspect of this construct were omitted from the scale, then the validity of the operationalization of the construct would be harmed, because the scale would not measure the construct completely and a specification error would occur. This can result in a serious bias, because the estimates based on an incomplete scale often would not mirror the complete true attitudes of the members of the target population on illegal immigration.

[p. 899 ↓]

Measurement Error

Measurement error is a complex component of total survey error. It consists of various elements that individually and jointly may cause systematic survey error as well as random survey error. Accordingly, measurement error potentially contributes to an increase of the estimate's variance as well as to its bias. Measurement error arises from the mode of survey administration, from the questionnaire or survey instrument and from the setting in which the instrument is administered, from the interviewers (if present), and also from the respondents.

Survey Mode

A traditional trichotomy of data collection modes differentiates face-to-face surveys, telephone surveys, and self-administered (mail and Internet) surveys. They differ with respect to (a) the presence or absence of an interviewer—which allows for various degrees of standardization of the measurement process, for different types of motivational support to the respondent, as well as explanation and help for the respondent— and (b) the dominant communicative channel (audiovisual, audio-only, visual-only). In recent years, many new survey modes have evolved with the introduction of modern information and communication technologies. Some of these modes transfer an established methodology into a computer-assisted mode (e.g., the shift from paper-and-pencil personal interviewing [PAPI] to computer-assisted personal interviewing [CAPI] or computer-assisted telephone interviewing [CAT!]), other new modes evolve as a consequence of merging survey modes (e.g., mobile Web surveys that use messenger systems or agents or avatars). Each of these survey modes has its particular strengths and weaknesses for specifie survey topics and survey designs. Whereas a Web survey might increase the variance of an estimate because respondents tend to answer a frequency question superficially compared to a face-to-face interview, the response to a face-to-face version of the very same questions might be prone to a higher degree of systematic social desirability distortion because of the presence of an interviewer, which in turn contributes to measurement bias.

Questionnaire

During the past 25 years, questionnaire design has been seriously developed from an art of asking questions to the science of asking questions. This line of research has demonstrated on innumerable occasions that slight modifications in the wording of a question and/or the response categories, or of the order of the questions and/or response categories, or in the visual design of the whole questionnaire, as well as of single questions, can affect the answers obtained from the respondents. Since the early days of the Cognitive Aspects of Survey Measurement (CASM) movement, numerous research papers and textbooks have contributed to a coherent theoretical approach that helps explain and predict random measurement error and systematic measurement error related to the questionnaire.

Respondent

Also within the CASM framework, a detailed theoretical approach on how respondents consider and answer survey questions has been developed. As a result, the question-answer process has been described psychologically in great detail. Using this framework, several systematic and random respondent errors have been identified related to what may happen when respondents answer survey questions. For example, satisficing behavior—as opposed to optimizing response behavior—as well as mood effects have been demonstrated to occur by methodological research.

Interviewer

Finally, it has been demonstrated that the personal and social characteristics of interviewers, if they are present, as well as their task-related and non-task-related behaviors may have a considerable influence on the answers obtained from the respondents. Accordingly, a great deal has been learned in the past 30 years about how to train and monitor interviewers to reduce the likelihood that their behavior will negatively impact respondents' answers. However, it should be recognized that it is impossible to eliminate all of the effects of individual respondent reactions to the personal and social characteristics of an interviewer, as interviewer-administered surveys require a personal encounter of respondents and interviewers.

Processing Error

In addition to lack of specification validity and to measurement error, the errors that may occur when editing and processing survey responses obtained **[p. 900 ↓]** from the respondents are part of the TSE framework. Poor handwriting with open-ended questions, the treatment of answers that were initially not codable, and the classification of occupations are just a few examples of possible errors that may occur at the data-editing stage of a survey. Also, scanning paper forms using OCR (optical character recognition) technology or keying the answers from a paper questionnaire into a database are prone to errors. In addition, crucial responses may need to be imputed as a result of item nonresponse (i.e. missing data), and this is susceptible to random error and to systematic error. Accordingly, these survey steps and the errors associated with them might either increase the variance of a variable—which in turn inflates the standard error and the margin of error—or compromise the accuracy of a response because a bias is introduced.

TSE And A Simplified Formula For Mean Square Error

Statistically speaking, TSE is the difference of a sample estimate and the respective parameter in the target population. This difference is measured by the mean square error (MSE), which in turn consists of two components: (a) the squared sum of the bias components plus (b) the sum of the variance components. For the mean square error, one needs to combine both bias and variance from all sources to obtain an estimate of the TSE. However, although most sources of possible error contribute to bias and to variance simultaneously, some error sources are predominantly responsible for an increase of either variance or bias. Thus, a simplified formula for the mean square error is as follows:

$$\mathrm{MSE} = (\mathrm{B}_{spec} + \mathrm{B}_{meas} + \mathrm{B}_{proc} + \mathrm{B}_{cov} + \mathrm{B}_{nr})^2 + \mathrm{VAR}_{meas} + \mathrm{VAR}_{samp} + \mathrm{VAR}_{adj},$$

where the terms have the following meaning:

B_{spec} = Specification bias (reduced construct validity)
B_{meas} = Measurement bias
B_{proc} = Processing bias
B_{cov} = Coverage bias
B_{nr} = Nonresponse bias
VAR_{meas} = Measurement variance
VAR_{samp} = Sampling variance
VAR_{adj} = Adjustment variance

Even though it is easy to estimate sampling variance, as explained in every introductory statistics textbook, it is less than trivial to estimate the other types of variance and especially the biases. Thus, the MSE as a measure for the TSE is often only of heuristic value, because the exact value of a particular variance or bias component cannot be computed reliably.

The MSE offers the opportunity to evaluate survey designs and the estimates computed based on a survey design. Thus, when reporting the results of a survey, end-users of the particular survey data can assess the quality of the estimate not only based on sampling error and the margin of error but also based on other error components. This is especially important because the bias component of the MSE generally is assumed to exceed the size of the variable error. Thus, the sample estimate of the population parameter often departs more from the true value than what is assumed based on the sampling error alone.

Also, the MSE allows an assessment of various survey designs to facilitate the decision of which design likely would produce data of the highest quality in a given time frame and for a fixed amount of money. However, in practice, survey designs are not only evaluated in terms of their MSE. Instead, survey design A may be preferred even though it produces data of lower quality in terms of the MSE compared to survey design B. For example, if the estimated cost for survey design B is considerably higher that design A's costs, the person responsible for the survey may have no choice but to go with survey design A. Thus, the TSE framework also relates to survey costs and requires survey designers to consider the accuracy in relation to cost and the timeliness of reporting.

Ultimately, the researcher's goal is to reduce the TSE by balancing various trade-offs in design decisions. Most of the time, design decisions—like choosing a certain mode of administration or choosing a special interviewer training procedure—affects not only one source of error but rather multiple sources. Thus, each desirable reduction in terms of a particular error source may be accompanied by an undesirable increase of some other error. Therefore, survey designers need to be able to compromise and balance several sources of error simultaneously.

[p. 901 ↓]

Limitations Of The TSE Framework

Even though TSE offers a convincing framework for the accuracy of a survey estimate, it also suffers from a serious drawback. Currently, the effort necessary to compute a reasonable quantitative estimate of the magnitude for a particular error component usually exceeds the available resources. The estimation of the MSE requires multiple repetitions of the survey design, which is usually too costly and also not feasible because the target population does not remain unchanged in between the repetitions. Also, for many survey designs some error components are not accessible because of the field procedures applied or legal constraints (e.g., privacy laws prohibit extensive nonresponse follow-up studies in many countries). Also, it should be noted that for the exact computation of the MSE, the population parameter needs to be readily available. Because this is usually not the case, the MSE is seldom explicitly determined in practice. More often only a few key components are estimated, or a survey design is rated along the various components of bias and variance on a scale from "low" to "high." The decision for a particular survey design then is made on the basis of a detailed computation of some of the error of the components and a rough assessment of the magnitude of some of the other error components. This leaves the researcher, as well as the end-user of a survey statistic, in a situation where a qualitative assessment of the magnitude of the total survey error is the best available assessment.

Strengths And Benefits Of The TSE Framework

Nevertheless, survey research and survey methodology have benefited greatly from the emerging TSE approach. The TSE framework has helped to make researchers more aware of possible errors in their survey statistics and the implications of these likely errors. For example, if the response rate and the size of the net sample are the only noticeable indicators for a given survey, many likely biases remain undetermined. Thus, the TSE framework motivates a systematic reflection on possible impairments of survey quality. In doing so, it stimulates a professional evaluation of ongoing surveys in terms of data quality and provides a common language and terminology for a critical discussion.

In addition, the framework provides a theoretical explanation for the various types of possible errors (variance and bias) and also for the underlying mechanisms (random error vs. systematic error). Also, it names a wide range of possible sources of threats to data quality. Hence the TSE framework suggests a theoretical approach for further developments of the survey methods beyond traditional approaches (ones that are not working well enough). In addition, it provides measurable indicators in order to evaluate the improvements of these new survey methods.

The TSE framework also has provided a basis for heightened interdisciplinary discourse across the boundaries of traditional disciplines. Surveys have been used for a long time in sociology, psychology, economics, and educational research, but until relatively recently, professionals in these disciplines have not been in close communication with each other. Even though it is too early to state a true integration of the field-specific methodologies, one can say that the survey branches of the subject-specific methodologies have merged, or at least are in the process of integration, based on the TSE framework and the methodological advances it has stimulated.

In an international perspective, the integrated concept of a TSE has contributed to the dissemination of "standardized" quality criteria and a set of methods to meet those criteria. International survey endeavors like the Programme for International Student Assessment, the International Social Survey Program, and the European Social Survey would not be feasible if researchers of diverse cultural and disciplinary backgrounds had not begun to interact and cooperate within a common framework. Even though there are still many national specifies in the design and the administrations of a survey, a minimum degree of concordance in the assessment of the data quality is provided by the TSE framework.

From a constructivist perspective, the TSE framework seems to be naive in one of its fundamental assumptions: Is there really something like a true value? Although one could argue that it is a reporting error if a respondent omits a certain portion of his or her income when asked for the monthly gross income of the household (e.g., the portion that comes from child support/alimony), one might also argue that this survey's definition of income also contributes to a social construction of "income." More traceable, surveys contribute to the shape and guise of public opinion when results of attitude surveys are repeatedly **[p. 902 ↓]** reported by the media and thus function as a reference point for the general public while they form their opinions on various public issues. Even from a less fundamental perspective, it remains questionable whether there is a perfect, faultless way of designing and conducting a survey. Accordingly, the true value is rather a chimera that cannot be measured without intervening with instruments and procedures that are, by themselves, selective and incomprehensive in principle.

However, from an analytic point of view, it definitely makes sense to assume fixed and constant true values at a given point in time. And it remains the principal goal of survey methods that they measure and mirror these parameters in the target population. With the aid of the TSE framework, survey researchers and survey practitioners have the instruments at hand to assess, discuss, and improve the quality of the respective estimates.

Marek Fuchs

http://dx.doi.org/10.4135/9781412963947.n585
See also

Further Readings

Biemer, P. P. , & Lyberg, L. E. (2003). Introduction to survey quality . New York: Wiley.

Groves, R. M. (1989). Survey error and survey cost . New York: Wiley.

Groves, R. M. , Fowler, F. J. , Couper, M. P. , Lepkowski, J. M. , Singer, E. , & Tourangeau, R. (2004). Survey methodology . New York: Wiley.

Lessler, J. T. , & Kalsbeek, W. D. (1992). Nonsampling errors in surveys . New York: Wiley.

Weisberg, H. F. (2005). The total survey error approach: A guide to the new science of survey research . Chicago: University of Chicago Press.