

Interviewer Falsification in Survey Research: Current Best Methods for Prevention, Detection and Repair of Its Effects

1. Preface

Survey researchers have an obligation to truth in data collection and in reporting of survey results. Throughout the design, data collection, and processing protocol of the survey, they must be committed to protecting the integrity of research data. To be effective, the commitment must extend beyond the principal researcher to all survey staff, including interviewers.

If interview data do not reflect the answers or characteristics of the respondent, but rather are the invention of the interviewer, data integrity is directly affected. Hence, the survey researcher must work to prevent and detect falsification of research data by survey interviewers. When falsification is detected, the survey researcher must attempt to repair the damage.

In much of survey research, data collectors or interviewers are part-time, temporary employees with limited tenure in the organization. They generally do not participate in the design of the sampling protocols and the questionnaire, nor do many have prior training in scientific research ethics. Nevertheless, it is vital to data integrity that interviewers strictly adhere to protocols and avoid falsification in any form.

Survey organizations can prevent or reduce interviewer falsification by providing a supportive environment for interviewers, designing studies appropriately, maintaining conspicuous deterrents, proactively seeking to detect any instances of falsification, and responding firmly and speedily to interviewer misconduct. Together, these actions create an environment of organizational integrity that will keep interviewer falsification rare and minimize its potential harm to data.

2. What Is Interviewer Falsification?

“Interviewer falsification” means the *intentional* departure from the designed interviewer guidelines or instructions, unreported by the interviewer, which could result in the contamination of data. “Intentional” means that the interviewer is aware that the action deviates from the guidelines and instructions.

Falsification includes:

- a. fabricating all or part of an interview - the recording of data that are not provided by a designated survey respondent and reporting them as answers of that respondent;
- b. deliberately misreporting disposition codes and falsifying process data (e.g., the recording of a refusal case as ineligible for the sample; reporting a fictitious contact attempt);
- c. deliberately miscoding the answer to a question in order to avoid followup questions;
- d. deliberately interviewing a nonsampled person in order to reduce effort required to complete an interview; or
- e. otherwise, intentionally misrepresenting the data collection process to the survey management.

There is, thus, a continuum of severity of falsification. Some incidents of falsification can seriously damage the statistical conclusions of the survey. Others lead to more minor damage.

This statement of current best methods focuses on minimizing falsification of survey data. However, it does not address acts of falsification by survey staff other than interviewers (e.g., project leaders, sampling statisticians, interviewer supervisors, coders, data processors). It omits other types of interviewer falsification that affect the employer-employee relationship (e.g., misreporting hours worked). Similarly, falsification of data does not include the common and unintentional errors of measurement that arise in the question-answer situation or mistakes by the interviewer in recording answers due, for example, to failure to understand or remember the protocol. It requires the interviewer to know that the data being recorded are false at the time they are recorded or to become aware of this after the fact and not acknowledge the errors. Thus, determining that falsification has taken place involves some judgment about the intention of the interviewer.

Prevalence of Falsification. Interviewer falsification has long been recognized in survey research, both in the published literature and in the professional practices that have been developed over the years to prevent and detect it. Certain characteristics of interviewers, characteristics of respondents or sampled units, and features of the interview task affect the likelihood that falsification will occur. The various processes, practices, and incentives all play roles in making falsification more or less likely.

The survey profession has developed highly effective means for the prevention, deterrence, detection, and correction of falsification in its various forms. These tools are

standard practice for most large survey projects, especially those that are federally funded and conducted by established survey organizations. This statement of current best methods reflects the accumulated experience and prevailing practices of organizations that have developed effective methods to enhance interview quality.

The various control practices are actively followed in most survey organizations, so the prevalence of falsification is quite low. The literature suggests that where appropriate methods are used, interview falsification is rare, involving only a small percentage of interviewers and a substantially smaller percentage of interviews. Still, most survey organizations acknowledge that they have experienced falsification on one or more studies they have undertaken over the years.

Acts of falsification by one or a handful of interviewers—while always a serious breach of the norm of data integrity—rarely threaten the overall study objectives or alter findings in any significant or meaningful way. Because of the way interviewer assignments are made, falsification of an interview or sample contact does not contaminate data collected by the vast majority of interviewers who faithfully follow the protocol. There have been occasions when falsification in a survey is substantial and widespread, but these situations are highly unlikely in studies that follow the methods described in this document. Whenever falsification is detected, survey researchers have an obligation to attempt to repair the integrity of the data.

Interviewer-assisted data collection typically occurs in two environments - centralized telephone facilities and through face-to-face contact of sample members in the field. Differences in these two environments affect the ability to prevent and detect interviewer falsification. For instance, falsification can occur in both settings, but implementing procedures that prevent and detect falsification are more readily accomplished in centralized telephone facilities where behavior can be observed and monitoring of interviewers is routine, feasible, and less costly. Most practitioners believe that falsification is more rare in surveys that are conducted exclusively in centralized telephone facilities than in those conducted by dispersed field staff. As will be described below, the recommended detection and deterrence measures differ between these data collection modes.

Obviously, interviewer falsification is non-existent in self-administered surveys that are sent directly to respondents, either on paper or electronically, and directly returned. However, falsification is possible when interviewers are tasked with distribution and collection of self-administered instruments in the field.

3. What are Effective Ways of Preventing Falsification?

Data integrity is a product of organizational integrity. As with all quality assurance processes, the assurance of data integrity in survey research has many components. In interview studies these include:

- a. articulation of values, goals and rules by research managers and supervisory personnel;
- b. selection and training of interviewers;
- c. the manner in which interviewers are compensated, supervised, evaluated, and rewarded;
- d. overall project design and budget;
- e. the design and execution of programs of deterrence and detection, including the use of advanced technologies;
- f. appropriate use and analysis of process data and data from actual interviews to detect anomalies; and
- g. appropriate actions in response to suspected and proven instances of falsification.

Effective control of falsification is not the result of any single method, but of the combined aspects of the study-specific environment in which interviewers conduct their work.

Researchers often require interviewers to obtain very high response rates from reluctant populations, to use complicated and long questionnaires, to take auxiliary measurements, and to accomplish these tasks efficiently. Indeed, many of these decisions are made to maximize the quality of the survey results. Unfortunately, they have the undesirable consequence of increasing interviewer stress and thus the risk of falsification. Examples of organizational factors that may affect the prevalence of interviewer falsification include:

- a. hiring and training practices that ignore falsification threats,
- b. inadequate supervision,
- c. lack of concern about interviewer motivation,
- d. poor quality control,
- e. inadequate compensation,
- f. piece-rate compensation as the primary pay structures,
- g. excessive workload, and
- h. off-site isolation of interviewers from the parent organization.

Examples of design factors, some of which may be necessary for high quality surveys, that may influence falsification include:

- a. interviewer-delivered incentives for respondents,
- b. interviewer production bonuses,
- c. long, complex interviews or data collection protocols,
- d. difficult to locate or enumerate sample members, and
- e. short field periods.

Family pressures, financial problems, and health issues can also affect job performance. Such stresses can contribute to the risk of interviewer falsification. Attentive supervisors may note these circumstances and can take steps to minimize falsification risks.

Preventing Interviewer Falsification. Researchers can promote conditions that prevent interviewer falsification by creating an organizational environment that encourages honesty, discourages falsification, enhances morale, and values data quality.

To reduce the risk of falsification, organizations should consider several procedures. They should inform clients about the problems of burdensome survey instruments. They should include information about the prohibition of falsification in recruiting, hiring, training, and supervising interviewers. When hiring interviewers, reference checks are recommended, and criminal background checks may be advisable. Organizations should require all newly hired interviewers to sign a pledge of ethical behavior and should describe the consequences of falsification. The pledge should describe the importance of data integrity and the consequences to interviewers of falsifying data.

Organizations should acknowledge that production quotas, some pay structures, and the use of production incentives may increase the probability of falsification. These protocols are necessary and desirable in some situations, but the falsification detection procedures should reflect the increased risk inherent in these situations.

A primary way organizations can prevent falsification is through observation and verification. Organizations must inform interviewers that their work will be monitored and/or verified, as this awareness can serve as an effective deterrent to falsification.

4. What Are Effective Ways of Detecting Falsification?

Fabricated interviews are generally easier to detect than falsification of individual data elements. Since detecting the latter requires more extensive verification, there is an inherent tension between controlling costs of verification efforts and increasing the probability of detection.

Procedures for detecting interviewer falsification include observation of the data collection process, recontacting respondents, and ongoing review of administrative, process, and interview data. These methods are typically implemented by a supervisor, an independent interviewer, and/or an other more highly trained and experienced staff member. All staff involved in these activities should have an unquestioned commitment to the identification of falsification.

Observational Methods. “Observation” means that another staff member sees or hears interactions between interviewers and respondents. Common observational methods include: silent monitoring (e.g., audio; visual; screen capture) in centralized phone facilities and audio taping or digital audio recording in field surveys. Where monitoring is used, interviewers must know that they will be monitored but should not know when they will be monitored. In centralized facilities, unobtrusive monitoring is usually a key part of routine quality control, mostly focusing on observing compliance with interviewing guidelines. In these facilities, monitoring alone is generally sufficient for detection and deterrence of falsification.

Recontact Methods. Recontact methods to detect falsification are generally used in field surveys. Common modes of recontacting respondents include mail, telephone, and face-to-face. Once an interview has been completed, the recontact efforts should commence as soon as possible. There are consistent differences among the three methods on cost and recontact response rates. Face-to-face is the most expensive mode, yet it generally achieves the highest response rates; mail is least expensive and generally achieves the lowest response rates. Whereas face-to-face recontact is preferable, a mixed-mode approach that includes a face-to-face component is often more cost-effective. For cases where recontact by phone is not possible, face-to-face verification is recommended over mail methods.

Data Analysis Methods. Data analysis for the purpose of detecting falsification includes identification of outliers on interview length, disposition coding, daily or weekly production, and key questionnaire items. Data analytic methods permit the organization to target interviewers who appear more likely to have falsified data and to then increase the monitoring and verification of these staff. The identified cases are usually investigated further by recontact of respondents.

Selection Procedures. Observation and verification procedures should apply to all interviewers and continue throughout the entire data collection period. Typically, 5-15% of the interviews are monitored and/or recontacted. Identifying cases for verification should include a combination of both random selection (using probability sampling) and targeted selection. The random portion of the observation/verification sample should be designed to provide estimates of the prevalence of falsification. The targeted portion is focused on detection (e.g., identifying suspicious cases and investigating whether falsification occurred). To identify unusual and suspicious outcomes, survey researchers can review process and administrative data, as well as data from completed interviews. In many cases, new interviewers are given more attention. Similarly more attention often is given during the start-up and again during the concluding phases of the survey field period.

Longitudinal surveys, where the same unit is interviewed repeatedly over time, may require special recontact designs in consideration of respondent burden.

Recontact Questionnaires. At a minimum, detection systems should make a determination of whether an interview actually took place. Once that is established, a small set of factual questions can be re-asked or confirmed such as:

- a. household composition and/or other eligibility requirements,
- b. mode of data collection,
- c. length of interview,
- d. payment of incentive, if any,
- e. use of computer during data collection,
- f. key topics discussed, and
- g. key items, especially those that govern large skips in the interview.

5. What Are Effective Actions to Take When There Is Evidence of Falsification?

Personnel Actions. If falsification is suspected, survey researchers should conduct an investigation by reviewing other work of the interviewer in question. To protect the integrity of the data in instances where some evidence of falsification is obtained, the researcher should remove the interviewer in question from all data collection activities until the issue is resolved. If a preponderance of evidence indicates any falsification, the researcher should initiate personnel actions, according to the organization's policies. Researchers should know and follow the relevant personnel and research integrity

policies and practices that apply to their organization. For serious or continuing falsification, the profession's standard is dismissal without the possibility of rehire.

Repair. Survey organizations should not deliver data that are known to be falsified. When an interviewer is found to have falsified data, the researchers should remove all contaminated data and attempt to replace them with valid data where practical. Replacement is not always possible; for example, falsified data from a pre-election voter study cannot be replaced post-election.

6. What Are Appropriate Procedures for Documenting and Disclosing Results of Falsification Detection Efforts?

As with all study procedures, researchers have an obligation to document and disclose the results of efforts aimed at preventing and detecting falsification. The researcher makes this information available as an important component of understanding the quality and accuracy of a survey. Typically this would appear as part of the technical documentation.

At a minimum the documentation would include:

- a. the sample design and selection method for the monitoring and verification protocol. This would include the sampling rate and number of sample units for probability samples;
- b. non-probability selection methods should be described to the extent feasible. The researcher can legitimately withhold from disclosure certain features of targeting procedures. For example, releasing the criteria on which targeting procedures are based could significantly weaken the effectiveness of the detection program;
- c. the methods of detection employed. This would include the mode, procedure used (e.g., monitoring or recontact), and timing of the detection effort;
- d. summarized falsification detection results, including:
 1. response rate, if appropriate. For example, if mail is used as part of a falsification detection program, the mail-back rate should be documented.
 2. number of cases found to be falsified. This would include separate results for falsification of entire interviews and of selected questionnaire items.
 3. documentation of the levels and methods of repair.
 4. statistics on personnel actions (e.g., number and percent of interviewers dismissed because of falsification evidence).

The report should provide estimates of the falsification rate from the probability sampling components of falsification detection programs. For situations where non-probability methods (e.g., targeting) are also employed, discussions of falsification rates must distinguish between these sources of information.

By disclosing information about interviewer falsification and falsification control measures, survey organizations underscore their continuing commitment to data quality and research integrity.