Mission Control, We Have a Problem: Questioning the Reliability and Validity of On-site Data

David Filiberto, NMR Group, Inc.
Lisa Wilson-Wright, NMR Group, Inc.
Lynn Hoefgen, NMR Group, Inc.

ABSTRACT

Impact evaluations of compact fluorescent lamp (CFL) programs often rely on on-site saturation studies to estimate CFL adoption rates. The assumption is that on-site data collection results are more reliable and valid than telephone survey self reports at estimating CFL purchases, use and storage because a trained technician visits participant homes and records the information. A recent on-site study performed for a CFL impact evaluation draws this assumption into question. Relying on a panel approach in which the same households were visited in 2009 and again in 2010, this study revealed large variation in the number bulbs counted in the same homes between the two years, as well as in the number and types of rooms listed. Most of these differences could not be explained by room or fixture additions to the home from one year to the next, but instead most likely reflected measurement error stemming from differences in how the technicians filled out the inventory matrix in one year or another. Appropriate changes to the panel approach and redesign of the on-site data collection form may reduce measurement error and result in more valid and reliable on-site results.

Introduction

> "The quality of a survey is best judged not by its size, scope, or prominence, but by how much attention is given to [preventing, measuring, and] dealing with the many important problems that can arise." (Ferber et al. 1980)

The energy-efficiency community often turns to on-site studies to provide reliable and valid estimates of energy savings and demand parameters such as measure saturation, installation rates, and hours of use, among other factors. The operational assumption is that sending a trained technician into a home or business to count, visually inspect, or meter an efficient product results in lower levels of measurement error than approaches that rely on respondent self-reports or recall of similar data. While this assumption is sound, on-site methodologies are not infallible, but their imperfections often go unnoticed because we rarely have the ability to assess the degree of measurement error associated with any single on-site study.

This paper describes a study that revealed the fallibility of on-site studies because the authors adopted a panel study approach[1]—an opportunity to measure true change in a sample of homes, not simply estimate change based on a probabilistic comparison of data from two different samples. The panel approach involved performing on-site visits in the same homes in three different geographic areas

---

[1] A panel study is a type of longitudinal study in which variables are measured on the same sample of respondents over time. Panel studies are particularly useful in predicating cumulative effects, normally hard to analyze in single cross sectional studies. A panel study is useful when answering questions regarding the dynamics of change. On the negative, it is often difficult to recruit the same respondents to interview several times. Further discussion may be accessed online at: http://www.socialresearchmethods.net/tutorial/Cho2/panel.html

in 2009 and again in 2010 with the objective of determining the number of compact fluorescent lamps (CFL) in use in both years. The evaluators expected to utilize the observed change in the number of CFLs found in use and in storage from each visit to develop an estimate of CFL purchases that did not rely on respondent self-reports. Any CFLs found in the home between the two visits could reasonably be assumed to represent newly purchased CFLs. As this paper documents, greater-than-expected differences in changes in the number of bulbs of any type in homes revealed unexpectedly high measurement error that challenged the evaluators' approach for estimating CFL purchases. The authors discuss the likely sources of this measurement error and suggest possible changes to on-site studies—especially those relying on a panel approach—that may decrease such errors in the future.

Methodology

The on-site panel study discussed here occurred over a two-year period, 2009 and 2010.[2] The choice of an on-site panel reflected the fact that some reviewers of prior CFL impact evaluations had criticized these studies for asking respondents when they purchased the CFLs found in their homes. These reviewers argued that asking respondents to recall when they purchased individual CFLs introduced a degree of self-report error with which they were uncomfortable. In response, the evaluation team and their clients decided in 2010—after the initial 2009 visits took place—to conduct a panel study by revisiting some of the homes taking part in the 2009 study in 2010 to determine, with an anticipated higher level of accuracy, the change in use and storage between the two years and applying this information to estimate CFL purchases. They believed that a direct comparison over time of CFLs found in the homes of a panel of identical respondents from 2009 to 2010 would reduce the error associated with customer recall of purchases thereby providing a more reliable and valid estimate of CFL purchases between the two visits.

The evaluators recruited the initial 2009 sample of on-site households through a 2009 random digit dial (RDD) survey. Respondents to the RDD survey were offered incentives to participate in these initial on-site visits to their homes. The amount of the incentive was determined based on the evaluation team's prior experience with identifying the optimal incentive to induce participation as well as the cost of living of the area in which on-site surveys were conducted. The evaluators then randomly selected the initial on-site sample from among willing householders. Upon deciding to perform the panel study in 2010, the evaluators called a random selection of the 2009 participants and offered them an additional incentive to take part in the 2010; incentive amounts were based on the number of sockets observed in the home in 2009 and the cost of living in the area. Table 1 provides a summary of the number of on-sites completed in each area in each year and the incentive offered to participants in 2009 and 2010. Note that this paper limits its analysis to the collective 261 households visited in both 2009 and 2010; the 2009 visit is called the "first" visit and the 2010 visit is the "revisit".

Table 1. Revisit Panel Sample Size and Incentive Amounts

Incentive	**Area 1**		**Area 2**		**Area 3**	
	2009	**2010**	**2009**	**2010**	**2009**	**2010**
Incentive	$100	$100-150	$150	$150-250	$100	$100-200
Number of On-sites	203	132	100	65	99	64

[2] This paper focuses on the results from one client; however a separate client also adopted a panel study and experienced similar issues as those discussed here.

While visiting each respondent's home on-site, a trained technician used a previously developed inventory matrix to record the total number of lighting sockets in the home, the number of CFLs in use, and the number of CFLs in storage (Figure 1).[3] They also collected additional information not directly pertinent to the issues discussed here. There were slight differences between the 2009 and 2010 data collection forms (*e.g.*, the codes used to classify sockets or bulb types), but the data collected on CFL use and storage were kept fairly consistent between the two years. Both the 2009 and 2010 instruments were pre-tested in households in the service territories and revised based on the results of those pre-tests. The data collection contractor developed a training program for on-site technicians. All technicians had to attend a training session in person or via webinar and pass tests on session content before being approved to perform the 2009 and 2010 on-site visits. However, it is critical to note that, in most homes, a different technician collected the data in 2009 and in 2010.

Figure 1. On-site installed lighting data collection form*

	Fixture Information					**All Bulbs**					**CFL Bulbs Only**						
Room	**Fixture #**	**Fixture Type**	**# of Sockets**	**Dim**	**3-way**	**# of Bulbs**	**Bulb Type**	**Bulb Shape**	**Socket Type**	**Watts**	**Manu-facturer**	**Model #**	**Dim**	**3-way**	**When Obtained**	**Store Type**	**Notes**
See Codes		**See Codes**		**Y/N**	**Y/N**		**I /F/ CFL/ LED/ H/S/ X/O**	**T/G/ A/B/B ug/S/ C/ Tub/ O**	**S/P/G/ Can/O**				**Y/N**	**Y/N**	**1=2010 2=2nd Half 2009 3=1st Half 2009 4=<2009**	**See Codes**	

* Not all of the data collected are not discussed in this paper.

Evaluators also performed on-sites in homes recruited in 2010. Recruitment of "replacement sample" is a standard practice in panel research. The results of data collected from these newly visited homes are not discussed here, but the existence of these visits is introduced as they were ultimately combined with the revisited sample in the final analysis after verifying that the two samples did not systematically differ from each other. Likewise, the client for these efforts agreed to contribute both the revisit and new visit on-site data to a multistate modeling effort. Again, the results of this effort are not discussed here, but the existence of the study is raised in the conclusions.

Analysis

In any primary data collection effort, evaluators must take into consideration the principle threats to reliability and validity of collected data. To avoid poor, unreliable and invalid data, which can often result in costly and time consuming efforts to rectify, evaluators should be cognizant throughout the data collection design process of four main error types. According to Salant and Dillman (1994), a questionnaire offers generalizable results if it successfully avoids coverage, non-response, measurement and processing error. Groves (1989) defines these errors as follows:

- Coverage and sampling error: error that results from the failure to give some members of the population any chance of selection into the sample,

[3] Multiple clients across the nation use a very similar instrument for their own saturation studies and a similar instrument continues to provide data to the multi state modeling effort discussed in the text.

- Non-response error: error that results from the failure to collect data on all members of the sample,
- Measurement error: error that results from the failure of the recorded responses to reflect the true characteristics of the respondents,
- Editing and processing error: error that results from the failure to convert responses accurately into an analysis file.

An on-site interview faces several challenges to avoid and minimize non-response error while simultaneously limiting measurement error. A poorly trained or poorly prepared interviewer or auditor will in many cases collect unreliable and invalid data, and if left unchecked, can seriously undermine results. The American Association for Public Opinion Research (AAOPR 1997) lists twelve considerations that constitute best practices and standards when designing an effective data collection effort. With respect to technicians charged with interviewing respondents face-to-face, the AAOPR suggests that interviewers be carefully trained with proper interviewing techniques and that they are familiar with the subject matter of the survey. Standardizing interviewer performance is one of the best methods to minimize measurement error.

In this CFL impact evaluation, many of the four error types were minimized with appropriate sampling techniques, a carefully designed questionnaire and appropriate probes to elicit valid responses (i.e., measuring what they were intended to measure), survey pre-testing, technician training, and adherence to careful data processing protocols. However, despite these efforts to minimize error, in comparing the 2009 on-site results with the revisited on-site results from the same homes in 2010, the evaluators realized that the data still suffered from measurement error.

Comparison of 2009 and 2010 On-site Data

In the effort to estimate, with a high level of accuracy, the changes in CFL use and storage between the 2009 and 2010 program years, the on-site lighting study results were initially examined for obvious measurement error. The primary method of checking for the accuracy of collected information in this study was to compare the number of bulbs in each household revisited across years. If the change in bulb counts remained relatively stable (operationalized as a change of less than or equal to 10 bulbs) from one year to the next, it could be assumed the data collection was reliable and accurate.[4] Substantial differences in bulb counts (operationalized as a change of more than 10 bulbs) in the absence of an additional or removed room(s) or fixture(s) led the evaluation team to suspect measurement error.

As explained earlier, the on-site visits in 2009 and 2010 relied on similar training materials and approaches, data collection forms, and on-site methodologies and protocols. Since a trained technician recorded in each participant's home the installed lighting, the rooms where the fixtures were located, and the number and types of bulbs in the fixtures, a straightforward method was developed to uncover any differences. The evaluation team took the following steps to estimate the change in bulb counts:

1. Sorted each respondent by a uniquely identifiable case number for both years.
2. Counted the number of bulbs for each household in each year.
3. Compared 2009 bulb counts to those in 2010 and identified bulb count differences greater than ten.

[4] The choice of exploring households with changes of more than 10 bulbs was arbitrary and chosen because it was a "round number" and seemed an appropriate choice for one-year changes in lighting to homes that had not been renovated. The evaluators could not use standard techniques to identify outliers (*e.g.,* box plots or standard deviations) because the number of homes demonstrating large changes in the number of sockets skewed the results of such techniques.

4. More closely examined households where the difference was greater than ten by comparing counts in each room for both years.
5. Identified any room or fixture additions or deletions from year to year that might explain the difference.

In reviewing the 2009 and 2010 on-site revisit data with the above method, the evaluators noticed large variation in the number of bulbs counted across the two years (Table 2). Ideally, these differences would be used to estimate the number of purchases of specific bulb types from year to year, but the differences were so large in some houses that the evaluators suspected measurement error was responsible. Specifically, bulb counts varied by more than 10 in 11 percent of the Area 2 households, 21 percent of Area 1, and 42 percent in the Area 3 homes.

Table 2. Bulb count differences from 2009 to 2010

Bulb Count Differences	**Area 1**	**Area 2**	**Area 3**
Total revisits	132	65	64
Number revisits differing > 10	28	7	27
% Differing > 10	21%	11%	42%
Number differing 1 to 10	89	52	33
Maximum difference in > 10 bulbs*	38	39	68

* Maximum differences exclude any items such as differential treatment of holiday lights that were identified as large sources of difference.

In order to explain the discrepancies, the team compared room and fixture types listed in Excel spreadsheets for 2009 and 2010 provided by the data collection contractor, recording any differences between the datasets for the two years. The evaluators identified multiple room-type discrepancies, such as room types listed in 2009 that were not inventoried in 2010 and vice versa. They also searched for discrepancies in the number and type of fixtures. Although fixtures have more potential than rooms to be installed or removed from one year to the next, the evaluators concluded that the changes in the number and types of fixtures exhibited a too large degree of variation. In particular, far more bulbs had apparently appeared or disappeared from homes than seemed reasonable over the course of the year.

Given the differences in bulb numbers across years in the Excel spreadsheets, the evaluation team next reviewed the hard copy data collection forms filled out by the field technicians in each year in order to uncover discrepancies in the Excel datasets that could have been due to data-entry errors. A handful of cases had identifiable data-entry errors, but these errors did not significantly reduce the number of cases with bulb count differences greater than 10.

The evaluators found that bulb count discrepancies in Area 1 and Area 2 were within acceptable tolerances for the needs of the evaluation. At this point, the clients and evaluation team had to make a difficult, pragmatic choice. The time already spent to address the unexpectedly high changes in bulb counts had stressed the evaluation budget, and the actual data analysis had not even begun. Since Area 3 revealed almost one-half of respondents having bulb count differences greater than ten, the evaluators focused their remaining resources on gathering additional information about the change in lighting in the more problematic Area 3. To do so, the evaluators revisited or called the problematic households in Area 3 and asked them additional questions about individual rooms, fixtures and bulbs to establish whether they had put additions on their homes or had added or removed several fixtures and bulbs during the course of the year. Unfortunately, this additional effort did not yield conclusive explanations for changes in socket counts for most of the questionable households in Area 3; instead, the data pointed to remaining measurement error.

Based on these efforts, the evaluation team came to a number of conclusions about differences in observed bulb counts from one year to the next. The evaluators expected that a household might add or remove a small number of fixtures and bulbs from year to year, thus affecting the number of bulbs in the home; this represents the primary reason for conducting the revisits and must be considered as a source of legitimate difference between the two visits. Yet, the team also concluded that it was not realistic to assume that 11 percent to 42 percent of households had added or moved more than ten bulbs from their homes in only one year without verifiable additions or reductions in room or fixture number. In summary the evaluation team identified several sources of inconsistency between last year's and this year's data that contributed to the large shifts in bulb counts observed in some homes:

- Reliance on different technicians in the 2009 and 2010 efforts without careful communication to the 2010 technician about irregularities faced by the 2009 technicians when visiting the same homes. Because of this situation, one technician may have made one set of decisions (or errors), and the second technician made a different set of decisions (or errors).

- Unexplained large differences in the reported number of bulbs within a room or room type, suggesting that technicians in one year or another had overlooked fixtures in the room, taken different paths through the home or had come to different conclusions about the nature of bulbs, fixtures, or rooms.[5]

- Inclusion of rooms and fixtures in one year but not another, suggesting that at least one of the technicians may have been denied access to a room or overlooked them. Examples of situations in which this may have occurred include denial of access to attics or overlooking the existence of a closet or garage in one year or another.

- Keying errors from the hard copy to the analysis file.

Table 3 illustrates the final adjusted sample disposition taking into account the reanalysis efforts mentioned above. Despite the effort to review each disputed case in the dataset in all three areas, examination of the hard copy lighting forms from all three areas for error, and revisiting/calling back several households in Area 3, the evaluation team was only able to resolve one case in Area 1 and six in Area 3 with bulb count differences greater than ten.[6] The measurement error effectively decreased the sample size in Areas 1 and 2 to the minimum levels required for statistical precision and the needs of the evaluation.[7] The evaluation team was able to create a more valid and reliable set of adjusted on-site revisit CFL inventory data by careful review of homes with questionable bulb count differences than otherwise would have been analyzed.

---

[5] For example, classifying a room as an "office" in one year and a "den" in another made comparisons across years difficult.
[6] Although the evaluators were not able to resolve any of the Area 2 cases in question, enough of a sample remained to yield reliable results for the needs of the evaluation.
[7] Statistical precision of a sample can be defined as "the closeness with which it can be expected to approximate the relevant population value" (Cohen 1988). As sample size decreased, the standard error increases or gets more imprecise.

Table 3. Final disposition of sample for analysis after data cleaning efforts

Sample Before and After Cleaning	Area 1		Area 2		Area 3	
	Before	After	Before	After	Before	After
Number revisits differing > 10	28	27	7	7	27	21
% Differing > 10	21%	20%	11%	11%	42%	33%

Conclusion and Recommendations

Strict adherence to minimizing measurement error before as well as after data are collected can often resolve difficulty. This evaluation followed standard on-site protocols used previously by these data collection contractors and contractors across the nation, including detailed training of the technicians and pretesting of the instrument. Even with careful adherence to these protocols, measurement error still occurred. The evaluation team discovered this error only after completing a panel study. On-site work has rarely relied on a panel study which is why similar measurement error has not come to light previously. The nature of the inconsistencies identified in this revisit, on-site-panel approach led the evaluators to reexamine methods and identify what changes, if any, could have been made to limit measurement error and further improve accuracy in future on-site visits and panel studies. Most of the differences in bulb counts between the 2009 visit and the 2010 revisit could not be explained by changes in the number of rooms or fixtures in the home. Given that the data collection forms in both years had similar layouts, instructions and protocols, with only minor additions and changes in language, the evaluation team ruled out the data collection form as a potential source of measurement error. Instead, measurement error most likely resulted from differences in how the technicians filled out the inventory matrix in one year or another. These differences were costly. The high level of measurement error necessitated additional data collection for some homes and required the evaluators to eliminate some households from the sample. It also challenged our ability to draw conclusions about changes in CFL use and storage, and, therefore, number of CFLs purchased by revisit households between the two visits.

Despite the detection of considerable measurement error in the revisited on-site saturation study, the evaluators believe that the results drawn from the study remain valid and reliable for the following three reasons. First, despite the reduction in *revisited* on-site sample size resulting from the data reanalysis efforts, the evaluators combined the on-site sample with on-site sample drawn from the aforementioned homes newly visited in 2010 (and visited only once) when describing the CFL market in the three comparison areas. This combination yielded sample sizes that were 33 percent to 50 percent larger than those for the revisit households presented in this paper. These larger sample sizes were sufficient to meet precision and sampling error at the 90/10 level.

Second, although the data were collected for a single client, the client agreed to take part in a larger multistate modeling effort. Comparisons to the overall multistate modeling sample demonstrates that data from these three areas fit patterns in CFL use and saturation observed in other areas across the nation, despite the fact that the multistate data were collected by different contractors using slightly different protocols.[8]

Finally, the observed trends in use and saturation were similar among the areas under study here and the other long-running program area in the multistate sample that also had revisits. This similarity

---

[8] See additional papers at this August 2011 International Energy Program Evaluation Conference (IEPEC) by Russell, *et al.* and Albee, *et al.* resulting from this effort.

existed both before and after reanalysis. Importantly, the other areas had similar data collection issues, and for this reason, the data were reanalyzed in a parallel manner in all three areas in which revisits occurred.

The evaluators considered the sources of the measurement error and lack of comparability between the first visit and the revisit and arrived at the following conclusion and recommendations. A quality data collection instrument ensures not only that data are collected carefully, but also builds checks and verification practices into its design to avoid measurement error. Even carefully trained on-site field technicians who follow a protocol vigilantly will still find that they are faced with decisions about how to code a particular fixture mount, room type, or bulb shape. Two skilled technicians could legitimately interpret such information differently. One technician visiting in 2009 takes a unique path through the home and makes decisions about how to code fixtures, rooms, and bulbs, and could have been barred by the householder from entering certain parts of the house. A second technician who performed the revisit in 2010 may have taken an entirely different route through the home and come to different conclusions about how to code the fixture, room, or bulb. This technician may have gained access to parts of the home that had been off limits the year before. Technicians may have also overlooked some fixtures from one year to the next, leading to inaccurate bulb counts.

In developing the research design, the on-site methodologies for the first visit did not take into consideration the possibility that another technician would again visit the home at a later date to collect similar information. Therefore, the protocols did not require keeping the types of information that would facilitate replication of the initial on-site visit data (*e.g.,* a sketch of the house, the path taken to inventory, listing of rooms that could not be inventoried and why).

The efforts have led to the conclusion that a significant amount measurement error in on-site saturation studies stems from the concerns mentioned above, as well as, the decision to create a panel *after* fielding the initial 2009 on-sites. The evaluators make several recommendations for future on-site panel studies that they believe could minimize measurement error:

- The adoption of an on-site panel study approach should occur *prior* to the fielding of any surveys; tacking a panel onto an existing study will almost certainly lead to difficulties, because protocols and instruments will not be designed with a repeat on-site visit in mind.
- One specific recommendation for all on-sites—whether they are one-time visits or panel approaches—is for the technician to make a sketch of the house and denote on the sketch the path they took through the home looking for information. Rooms, including closets, should be labeled using a standard code so that the data collection can be more easily verified—or replicated—at a later point. Moreover, room types such as bedrooms, bathrooms, or closets should have a number associated with their label, so that a later technician will know exactly which bedroom is Bedroom One. Denoting the path taken through the home reduces another potential source of variation, enhancing data collection efforts. Theoretically, the path through the home should not matter as long as the home is well diagrammed and the rooms numbered. However, in some home arrangements, entering from one doorway can increase the chances of seeing—or failing to see—certain fixtures. The evaluators believe that following the same path provides yet another way to bring consistency in data collection.
- If the householder allows, the technician could also take digital photographs of rooms and of any questionable fixtures or bulb types that could help in categorizing lighting products.
- On-site technicians should also be instructed to note any areas of the home they were not allowed to visit and why. If technicians revisiting the home later gained access to these areas, they could note that this room was in fact part of the home at the time of the prior

visit, but that it was not inventoried during the prior visit. Likewise, the revisit technician should note his or her inability to access rooms that had been inventoried previously.

- Institute a repeat visit as part of the data collection for a quality control spot check of a sub-sample of revisited homes.
- Instruct on-site technicians to take a second look around all rooms they have inventoried and verify that they have counted all lighting products in the room, including less noticeable ones that may be under cabinets, in closets, or installed in built-in bookcases. Additional training might stress that such fixtures are easy to overlook.

Some of these recommendations may seem self evident and in fact may be used in other commercial and residential programs such as energy audits. However they are not common practice in on-site lighting saturation studies. Taking the above recommendations into consideration in designing an on-site lighting panel study may potentially reduce measurement error and result in more valid and reliable on-site lighting results in future studies.

Despite the challenges of accurately collecting and representing observed data from on-site lighting studies, the evaluators believe that, with appropriate changes, the panel approach can enhance the year-to-year comparability (and the reliability and validity) of CFL program impact estimates. Most importantly, the evaluators must plan a panel study from the start in order to design a data collection instrument with repeat visits in mind. Requiring the first visit technician to sketch and photograph the home, denote the path taken through the home, label all rooms, and note which parts of the home they could not access for the inventory provides a solid foundation for reducing ambiguity and measurement error in both one-time visits as well as repeat visits. The technician visiting a home would use the sketch and information to guide the visit in the home, noting any substantial differences from one visit to the next. These steps would provide the vital communication needed between the technician performing the first visit and the technician performing the second visit, if the same technician is not available to visit the same home in both years, as if often the case. Additionally, as a quality control measure, a randomly selected subset of the revisited homes would be visited a third time to ensure the accuracy of initial data collection efforts.

Reducing systematic measurement error (or biases in the measurement) in on-site studies can be costly. The careful data collection in both the first and second visits may require using technicians with more experience than are often used in lighting saturation studies. Likewise, additional quality control visits will add to the on-site data collection cost. Yet, when the evaluation budget allows, the authors of this paper believe the costs are justified. In fact, the efforts of the current evaluation team to rectify the measurement error expended considerable evaluation resources—both time and money. In this case additional visits to homes, call backs, and analysis resulted in cost overages for data collection contractors and the evaluators; more careful planning and quality control would have added to the initial evaluation budget but would have likely reduced the need for the extensive data reanalysis required to address the measurement error.

The reliability and validity of data collected in the field is often judged by how associated errors are minimized. In the effort to improve on-site lighting data collection, the focus on enhanced technician training, standardized data collection instruments, additional visits tied together with extensive planning must be balanced with cost. There is a point where additional resources will only improve the quality of the data marginally. The authors believe implementing the above measures will provide a good hedge against undue measurement error. The point at which these efforts are no longer cost effective however, requires further study.

References

American Association for Public Opinion Research (AAPOR). 1997. *Best Practices for Survey and Public Opinion Research and Survey Practices AAPOR Condemns.* http://www.aapor.org/Best_Practices.htm. Ann Arbor, Michigan: American Association for Public Opinion Research.

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ferber, R. P. Sheatsley, A. Turner, and J. Waksberg. 1980. *What Is a Survey?* http://client.norc.org/whatisasurvey/downloads/pamphlet_1980.pdf . American Statistical Association.

Groves, R.M. 1989. Survey Errors and Survey Costs. New York: Wiley Interscience.

Salant, P. and D. Dillman. 1994. How to Conduct Your Own Survey. New York: John Wiley and Sons.