

The Survey Process: With an Emphasis on Survey Data Analysis

LCDR Walter W. Kulzy, III, United States Navy, walter.w.kulzy.mil@mail.mil
Dr. Ronald D. Fricker, Jr., Naval Postgraduate School, rdfricker@gmail.com

Introduction

Correctly executing a survey and subsequently conducting a rigorous analysis is a complex undertaking requiring deep domain knowledge of how to design a survey, how to appropriately select the sample to be surveyed, how to field the survey, and ultimately how to appropriately and methodically analyze the resulting data. Elsewhere (Fricker et al., 2012; Anderson and Fricker, 2015), we have focused on particular analytical techniques. In this article, we step back and review the entire survey

process, from defining the initial survey objective all the way through to conducting the analysis and reporting the results.

The survey process can be divided into six stages:

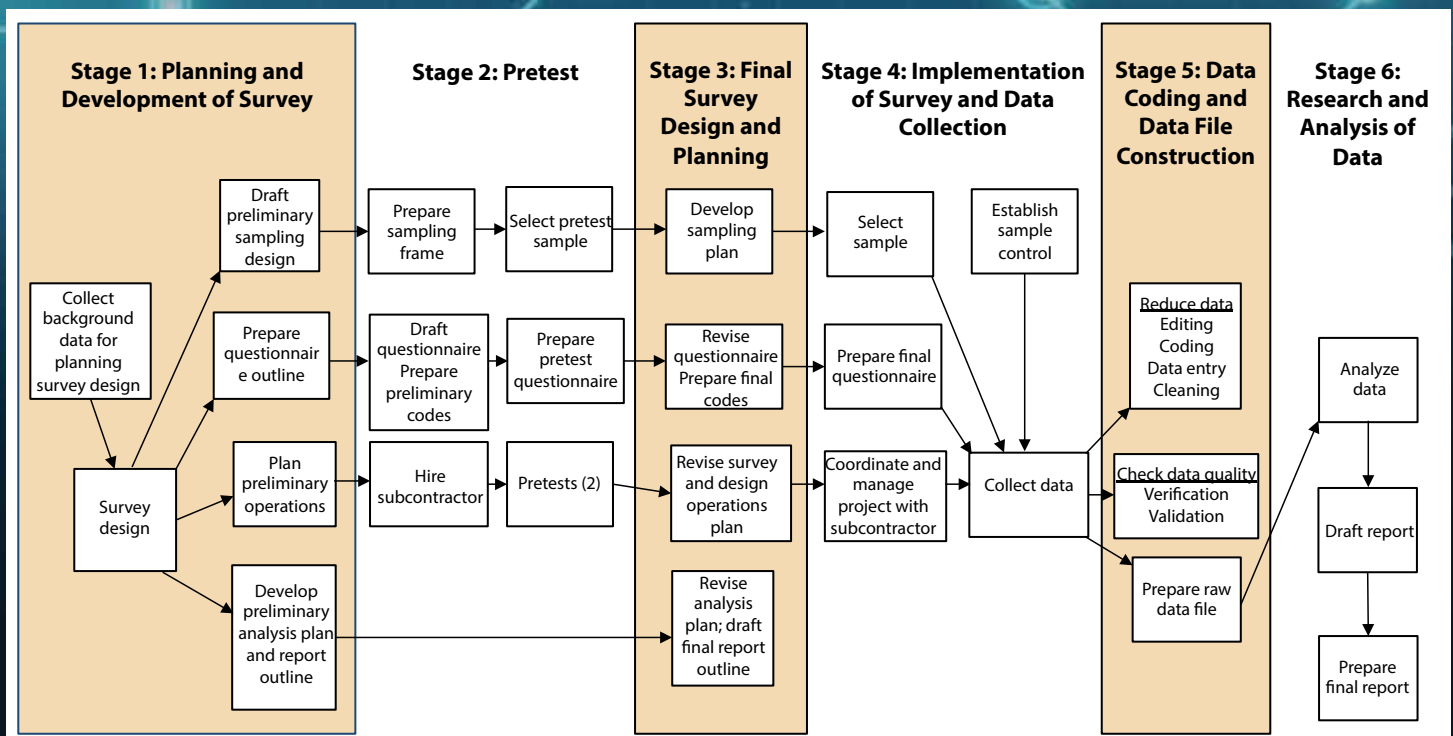
1. Planning and development
2. Pretesting
3. Final design and planning
4. Implementation/fielding
5. Data coding
6. Analysis and reporting

Each of these can be further broken down into substages and more detailed steps. The purpose of this paper is to both provide an overview of major stages in conducting large-scale surveys, and delve further into the substages of data coding, analysis, and reporting.

An Aside: Survey Expertise and Excellence

Before we describe the details of the survey process, it is important to recognize that surveying requires expertise in three broad areas:

Figure 1. Steps in conducting a survey. Source: Scheuren (2004, 8) . Used with permission.



survey and sample design, fielding methods, and analysis. As shown in Figure 2, well-done, rigorously conducted surveys require excellence in all three areas and, conversely, poorly executed surveys result from weakness in one or more of the areas. For example, no amount of excellence in the way a survey is fielded and analyzed can compensate for poor questionnaire or sample design. Similarly, a poor or incorrect analysis completely negates all the effort put into a well-designed and fielded survey. Good surveys only exist at the intersection of good design (both questionnaire and sample), good fielding, and good analysis.

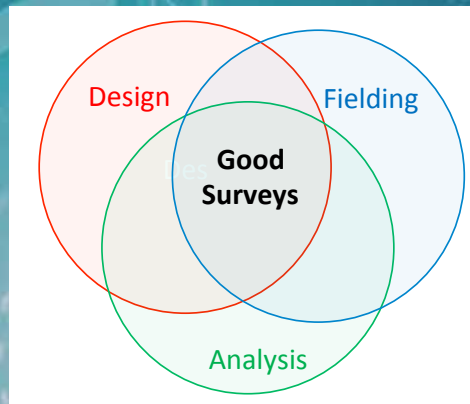
Now, we note that not everyone is or can be an expert in all three areas. However, because surveying is almost always a team effort, it is key that the team contains experts in all three of these areas. For example, as Figure 1 shows, large survey efforts are almost always contracted out to a survey company and some of the expertise—particularly the fielding expertise—will lie with the contractor. That said, it is critical that the government have sufficient expertise in all areas to appropriately select and oversee contractor performance.

The Survey Process

As we walk through each stage of the survey process, it is important to recognize that surveys rarely proceed as perfectly linear as shown in Figure 1. Instead, decisions made or required in future stages often require revisions to those made in previous stages. That is, the survey questionnaire may initially be designed assuming a particular type of fielding. Should those fielding plans have to be modified, which can happen for any number of reasons—budgetary, operational, etc.—then the survey team may have to revisit questionnaire

design. More generally, every survey effort requires compromises, often because of real-world constraints, particularly financial constraints, and a good survey team must continually assess the best approach for conducting a particular effort and make trade-offs, sometimes retroactively. Flexibility throughout the survey process, and a willingness to revisit and revise the products of

Figure 2. Good surveys result from excellence in design (both questionnaire and sample), fielding, and analysis.



prior steps, are key to conducting an outstanding survey.

Stage 1. Perhaps the most important step in Stage 1, and one that is not shown, is the explicit statement of the survey objective. Easier said than done, this is useful for a number of reasons:

- to ensure the entire survey team understands the purpose of conducting the survey,
- to ensure that the survey sponsor concurs with the team's intent, and
- to prevent survey mission creep.

The last is particularly important as large-scale surveys take a long time to conduct and, for any number of reasons, the process can lose focus

without an explicit objective.

Next, “items” are drafted with the objective, fielding method, and final analysis in mind. An item is a survey question and its associated response scale. It is important to note that proper response scale design is just as important as good question design, and a good response scale typically includes options for respondents to indicate they either “don’t know” or “refuse to answer” each question.

Questions that do not support the survey objective should be ruthlessly culled from the draft. It is absolutely critical that the questionnaire design effort involve both those who will actually field the survey and those who will analyze it. A good survey cannot be designed in the absence of fielding and analysis considerations (see Figure 2) and those who will ultimately execute and analyze the survey always provide useful improvements to questionnaire design.

Part of the involvement of analysts at this stage will be in drafting an analysis plan that will specify how the data collected will answer the survey objective. Here again, questions that do not contribute to the analysis should be ruthlessly culled from the questionnaire. Also, in this stage, the preliminary sampling methodology is created, usually by a statistician with expertise in survey sampling. This sampling specifies the number of respondents and how survey respondents will be selected from the population. It also provides initial estimates of the survey margin of error.

Finally, for large-scale surveys, this stage will also involve executing a contracting action to hire the survey subcontractor, where the contractor may then be tasked to prepare the draft questionnaire, sampling

methodology, and/ or the analysis plan for government approval.

Stage 2. There is both a science and an art to question and questionnaire development. For this reason, in Stage 2 it is important to understand who will be answering the questions and then to carefully pretest the preliminary questions developed in Stage 1 on individuals who are as similar as possible to those who will ultimately take the survey. It is almost impossible to overstate the importance of conducting rigorous pretesting of the survey questionnaire and then subsequently revising and re-pretesting it. We have never seen a preliminary questionnaire survive the pretesting process intact. Good pretesting significantly improves survey questions and questionnaire quality; rigorous pretesting is the mark of a professional survey effort.

Stage 3. Upon completion of Stage 3, all lessons learned and any other necessary changes are incorporated into the final questionnaire, the final sampling plan, the final operations plan, and the final analysis plan. It is at this point that the process is fully and formally documented and the planning is locked down prior to execution. Think of this stage as the generation of a final campaign plan that incorporates all the necessary trade-offs and lessons learned from the preparatory wargames, exercises, and coordination.

Stage 4. This is the execution stage. Here the actual sample of potential respondents is drawn according to the sampling plan and they are asked to complete the questionnaire. This is, of course, easier said than done. Within this step are a host of important activities that have to do with maximizing response rates and ensuring data quality. If fielding

is conducted by a professional survey company, the company will have standard processes and procedures that likely require little or no government involvement. However, it is always good practice to require the company to outline these procedures during the contracting process and to ensure that source selection is based at least in part on the company demonstrating mastery of these methods. It is also important that the contract requires the survey company to submit a post-fielding report that describes how the survey was actually fielded, their quality control procedures, survey response rates, any problems encountered in the field, and any deviations from the sampling plan.

Stage 5. For analysts, this stage is perhaps the most labor intensive and the level of effort required to transform the raw survey data into an analytical data file that is useful for analysis often takes first-time survey analysts by surprise. We'll discuss this stage in more detail in the next section; here we simply note that there are a number of important steps in this stage. The first is simply quality assurance, where the analyst should do some basic checks of the raw data to look for anomalies that may then require clarification from the fielding organization. One important quality check is to look for excessive "missingness," by both question and respondent that can uncover potential issues. After quality checking the data, the analyst will almost surely have to "recode" the data to make it useful for analysis (more on this shortly). However, a key point is that without rigor applied in this stage, the actual analysis could be filled with errors and unsubstantiated inferences. As Groves et al. (2013, 306) point out, the "act of coding itself can produce statistical errors. This can

have noticeable effects on survey instruments."

Stage 6. Whereas stage 5 is used to "clean" and prepare the data, it is in this stage that assumptions are made, algorithms are constructed, and statistics are used to craft and generate insights into a population. It is also in this stage that the analysis is made digestible to the decision maker via a clear and concise presentation of findings and insights. In many ways, distilling the analysis into an appropriate presentation can be more art than science, but it is just as valuable as properly conducting the survey because a decision maker failing to act because of poor presentation is just as a disastrous as a decision maker taking the wrong action because of a poor survey.

Analyzing Survey Data

Coding and Data File Construction

Critical to good survey analysis is the preparation of the data, which is, as mentioned above, a generally labor-intensive process.

Prior to coding or recoding the data, input and other errors will first need to be identified and corrected to the greatest extent possible. The goal is to find errors in the data, perhaps arising from data entry errors, or from data transcription errors, or from a multitude of other ways in which human error can innocently or less innocently result in bad data. For example, "stone curbing" is the act of a survey fielder filling in the survey in lieu of properly fielding it to actual respondents. There are accepted methods such as analyzing the demographic data to identify potentially false data (Munro et al., 2008).

Identifying data that may need cleaning is most fundamentally accomplished using different kinds of checks, including:

- **Missing data checks:** Are some questions too often missing responses, or have respondents routinely refused to answer one or more questions? Which respondents are missing responses to many questions?
- **Range checks:** Do the data for a given question or demographic fall within the expected range? For example, for an age variable, do some values fall outside the range of ages that would be reasonable for survey respondents?
- **Outlier checks:** Are one or more observations very different in some dimension from the other data? For example, do some of respondents come from regions outside of the survey area?
- **Logic and consistency checks:** Is the data internally consistent for each respondent? For example, data might be checked to see if all those who say they are pregnant are also female.

Data coding and recoding.

Depending on the analytical approach to be taken with the data, it may be necessary to recode Likert scale responses (e.g., “strongly agree” to “strongly disagree” responses) to a numerical scale. Similarly, one may want to recode binary yes/no response questions to 0/1 indicator variables. It may also be useful or necessary to create groupings to reduce a long list of categorical responses into a more manageable or meaningful shorter list. Furthermore, there may be a necessity to recode nonnumeric data to ensure it is consistent, particularly if the original data entry was done by more than one individual. For example, words

with alternate spelling (“theatre” or “theater”) will have to be corrected if in the analysis one will eventually want to summarize the number of times a word (“theater”) is mentioned.

Often a bigger challenge with recoding is deciding how to handle “don’t know” and “refuse to answer” responses. Depending on the question/response scale design, a response of “don’t know” may be equivalent to a neutral value or it may be equivalent to a nonresponse or even a “refuse to answer” response. There is no simple answer for how to make these types of recodes and they must be made on a question-by-question basis in the context of the particular survey and associated respondents. Key, returning to the issue of questionnaire design, is to ensure that respondents are generally allowed “don’t know” and “refuse to answer” responses to each question so that the analyst can make explicit choices.

Finally, we recognize that there are many other other issues related to whether it is appropriate to recode variables and then how to properly use them in an analysis that we won’t attempt to go into here. For example, when is it appropriate (or, perhaps, is it ever appropriate?) to treat an ordinal Likert scale-based variable as a continuous variable? Here we will just say that some type of recoding is almost always required to conduct an analysis and those choices have to be carefully made according to an accepted set of principles.

Preparing the analytical data file.

Liberal use of find-and-replace in Excel and “unique() and replace() functions in R” are two of several methods to assist in quality checks and recoding. Regardless of how one chooses to conduct the recoding,

any changes made to the data must be traceable, auditable, and, if necessary, reversible. It is critical that any changes to the original data are known and tracked and that a copy of the original data is maintained. In our work, we find it most useful to work in a software package that allows for some sort of scripting language so we can generate and regenerate the analysis file simply by reading in the original data and outputting a new analysis file. The script then becomes the audit trail for any and all cleaning, coding, and recoding actions.

Depending on the analysis, completing the analytical data file may also require “imputing” missing values. Missing values can be due to a respondent failing to respond to a question or being unable to answer due to the lack of a desired response to a question. Imputation is within sample inference, where Groves et al. (2013, 303) describe imputation as “placing an estimated answer into the item’s data field.” There are number of ways to conduct imputation with a range of complexities, the discussion of which is a subject for another article.

Data Analysis and Presentation of Results

Although Figure 1 reduces analysis to one box in Stage 6, this step can be both very involved and complicated. For any large-scale survey, the analysis will likely involve: (1) calculation of survey weights, (2) data reduction, (3) univariate and multivariate analysis and modeling, and (4) presentation of the results.

Calculating survey weights. For almost all surveys with a complex sampling design, weights must be used to do the correct inference from sample to population. Essentially, survey weights adjust for the fact that in complex sampling each respondent

could have a different probability of being sampled. See Anderson and Fricker (2015), and references therein, for additional discussion, including how raking can be useful for adjusting survey weights.

Data reduction. Often the sheer number of survey questions can be analytically overwhelming. As a result, it may be desirable to summarize the results from multiple questions with a single measure, either because one wants to or needs to reduce the dimensionality of the raw survey data into a dataset that is smaller and more analytically manageable, or because the population characteristics of interest are best measured using multiple questions. For the former, principle components analysis (PCA) may be a useful analytical technique and, for the latter, factor analysis can be useful. Although PCA and factor analysis are often both described as factor analytic methods, at least in the social science literature, they are methodologically quite different. For survey analysis, we frequently prefer factor analysis because it is designed to identify substantively meaningful groups of questions that can be very helpful in conducting the survey analysis and communicating the results. See Fricker et al. (2012) for an explanation of factor analysis.

Univariate and multivariate analysis and modeling. Whether univariate techniques, multivariate techniques, or a combination of the two is appropriate for analyzing a given survey depends on the survey objective. Regardless of the methodology, it is important that it is correctly done where, with complex sampling, both will generally require the use of survey weights and special software from which to calculate the correct standard errors. Said another way, except in special cases,

complex sampling requires more advanced statistical methods to conduct the correct inference from sample to population.

An analyst may not have a much time to produce a sophisticated analysis, particularly in the field. For example, the decision maker may be looking to quickly confirm an intelligence analysis or picture of the environment. For this “snapshot” of the operating environment, univariate analysis may be the most appropriate approach wherein the analyst can quickly compute a frequency distribution or similar product to articulate “the bottom line.” Similarly, the analyst may rely on single questions within the survey to generate a measure of effectiveness (MOE) for the command’s assessment. Thus, when time is of the essence, univariate analysis may offer a timely and good enough product.

Second, a challenge with univariate analyses, particularly with long and/or complicated questionnaires, is that it requires the consumer of the univariate analysis to mentally synthesize the results into an overall summary. This can be both mentally challenging and fraught with human error. We have seen surveys with 150 or so questions turn into slide decks containing hundreds of slides with as many bar and pie charts. In these cases, various types of multivariate analysis can be both informative and help the analyst and decision maker separate the important and informative from the mundane and noninformative results.

Presenting the results. As mentioned earlier, survey analysis is as much an art as it is a science, and it is in this final stage where the art shines. For, as exciting as multivariate analysis may be for the analyst, an

inability to communicate thoughtful insights to the decision maker in a logical and concise manner has the potential to negate all the hard work in conducting the survey. Key to this is communicating to the decision maker how the survey data address the original objective.

When presenting survey results, there are some important “dos” and “don’ts.” First, do present the results in terms of a narrative that describes how the survey addresses the original survey objective. Conversely, don’t subject the decision maker to a presentation that consists of marching though univariate results, question-by-question, in the order they were asked in the survey. The goal is to synthesize the results for the decision maker so he or she can understand what the data say about the objective.

However, second, when presenting results to the decision maker, provide him or her with the actual question wording. Particularly when presenting univariate plots summarizing individual question results, put the actual wording of the question on the slide. That way the decision maker can interpret the data in the context of the question itself. Also, it’s generally a good idea to provide the decision maker with a handout that tabularizes all of the survey results, question by question, so that he or she can delve into specifics as desired, but without the analyst having to brief question by question.

Third, somewhere at the beginning of the presentation, communicate to the decision maker measures of survey quality. This should include the survey response rate, an overall margin of error measure so the decision maker has some idea of the inferential uncertainty in the data, and any problems or issues that arose in

the fielding process (e.g., perhaps some locations inaccessible or were some subpopulations uncooperative or less responsive).

Fourth, as in any analytical briefing, the slides and associated graphics should follow presentation best practices and the details of complex analyses should be relegated to back-up slides. One practice we have found particularly helpful is including a verbal takeaway on each slide that summarizes what that summary graphic and/or statistics on that slide mean. In particular, the analyst/briefer should not assume that all in the audience will be equally adept at digesting statistics and quantitative material. To the extent that the survey contained open-ended questions, respondent quotes that appropriately capture and summarize the analytical findings can be particularly powerful and can be helpful to include on briefing slides.

Finally, particularly with survey data, the presenter should come fully prepared with a suite of back-up slides that address questions that might come up during the briefing. These often involve more detailed “slicing and dicing” of the data so the briefer can address questions as they arise. To be effective, the presenter must be intimately familiar with all the survey results so he or she can address questions that come up throughout the briefing and discussion of the analysis results.

Summary

This article should have made clear that conducting a rigorous survey requires many steps, where the

quality of the data requires executing all of the survey stages and associated steps well. Indeed, as we have described, just the appropriate analysis of survey data requires the successful completion of many steps. Furthermore, note that every survey is unique; there is no single recipe for designing, fielding and analyzing surveys and each is and should be a one-off tailored to the particular objective, population, and situation at hand.

For additional reading, we recommend Scheuren (2004) as a general primer, Dillman et al. (2009) for an introduction to survey design and fielding, and Lohr (1999) for a good introduction to survey sampling and analysis.

References

- Anderson, L., and Fricker Jr., R. D. 2015. Raking: An Important and Often Overlooked Survey Analysis Tool, *PHALANX*, to appear; preprint available at <http://faculty.nps.edu/rdfricke/docs/RakingArticleV2.2.pdf>.
- Dillman, D. A., Smyth, J. D., and Christian, L. M. 2009. *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, 3rd ed., John Wiley & Sons.
- Fricker Jr., R. D., Kulzy, W. W., and Appleget, J. A. 2012. From Data to Information: Using Factor Analysis with Survey Data, *PHALANX*, vol 45, No 4, 30–34.
- Groves, R. M., Fowler Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangeau, R. 2013. *Survey Methodology*, John Wiley & Sons.
- Lohr, S. L. 1999. *Sampling: Design and Analysis*, 2nd ed., Duxbury Press.

Munro, N., and Cannon, C. M. 2008. *Data Bomb*. National Journal Group Inc.

Scheuren, F. 2004. *What Is a Survey?* <http://www.amstat.org/sections/srms/pamphlet.pdf>. Accessed Feb. 12, 2015.

About the Authors

Walter W. Kulzy, III, is a Lieutenant Commander in the United States Navy. He holds a bachelor's of arts degree in economics from Virginia Tech and a master's of science degree in operations research from the Naval Postgraduate School. He deployed to the Persian Gulf for the Iraq War in February 2003, to the Diyala Province in Iraq in support of Operation Iraqi Freedom in 2007, and is currently assigned to the United States Central Command's J8- Analysis and Requirements Division.

Ronald D. Fricker, Jr., is currently a professor in the Operations Research Department of the Naval Postgraduate School and is soon to be the head of the Statistics Department at Virginia Tech. He received his PhD in statistics from Yale University. An author of more than 70 papers, monographs, and book chapters, Professor Fricker is a Fellow of the American Statistical Association and an Elected Member of the International Statistical Institute. He is on the editorial boards of *Statistics, Politics, and Policy* and the *International Journal of Quality Engineering and Technology*.