

Module 4 – Study Designs II (Surveys)



Original Author: Jonathan Berkowitz, PhD

PERC Reviewer: Amy Plint, MD

Table of Contents

Table of Contents.....	1
Overview	2
Introduction	2
Objectives.....	2
Key Concepts.....	2
Activities.....	2
Quick Links	3
Highlighted Tools	3
Task Checklist.....	3
Reading	3
Module 4: Study Designs II (Surveys).....	5
Background	5
Steps in planning a survey.....	6
How to Begin.....	6
How to Plan Your Survey Administration.....	6
How to Get Good Population Coverage.....	7
How to Choose a Random Sample.....	8
How to Determine Sample Size.....	10
How to “Plan In” Quality	11
How to Format	12
How to Schedule (The Dillman Technique).....	13
Designing the Questionnaire Items.....	14
Writing Good Questions.....	14
Open-ended	15
Close-ended with Ordered Choices	16

Close-ended with Unordered Choices	16
Does the Order of Response Choices Matter?.....	18
Shortcuts to Avoid.....	19
Using Survey Results	20
Conclusion.....	21
The Last Question: Why should you ask questions?.....	22

Overview

Introduction

Surveys (or questionnaires) provide a speedy and economical means of determining facts about peoples' knowledge, attitudes, beliefs, expectations and behaviours. They are widely used in research about patients and health care providers. Sir Ronald A. Fisher, one of the great statisticians of the past century said, "Nature will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other topic has been discussed." This module will address the development of a survey tool and survey administration.

Objectives

You will be able to:

- Determine whether a survey design is suitable to answer the research question
- Choose the most appropriate and feasible method of administration
- Select an appropriate sample size
- Establish a sampling frame and choose a random sample
- Write suitable questionnaire items
- Choose and appropriate format for a questionnaire tool
- Identify sampling and non-sampling errors
- Know what types of statistical analyses are possible

Key Concepts

- Construct a survey tool (questionnaire) that properly addresses the research questions
- Determine how to administer the survey to get a valid and large enough sample of responses

Activities

- Determine whether survey design is appropriate to answer the research questions
- Create a survey tool by writing items and/or locating relevant existing questionnaires
- Set up a method of survey administration
- Establish a sampling frame and select a random sample

Quick Links

- Online survey creation (free): <http://www.surveymonkey.com/Pricing.asp>

Highlighted Tools

On this Extra Resources page you will find several PERC surveys – both protocols and the actual surveys. All brought to you by the kind folks at CHEO.

<http://www.columbia.edu/~mvp19/RMC/M4/ExtraMaterials.htm>

- There are many Internet sites with advice on survey design, question construction, sampling methods, etc. Typing the keywords “designing survey questions” into Google gives about 3 million hits! Here are a few useful sites.
 - <http://www.statpac.com/surveys/>
 - <http://www.statpac.com/surveys/sampling.htm>
 - <http://www.isixsigma.com/library/content/t000702.asp>

Task Checklist

1. Prepare a draft survey tool using SurveyMonkey
2. Establish the sampling frame
3. Develop the protocol for how to reach the prospective respondents, how to receive completed surveys, and how to maximize response rate
4. Articulate the potential limitations and sources of bias

Reading

Main reference

- Hulley, SB, Cummings, SR, et al. (2001). *Designing Clinical Research*, Second Edition; Lippincott Williams and Wilkins. -- Chapter 15. Designing Questionnaires

Supplementary references

- McDowell, Ian & Newell, Claire. (1996). *Measuring Health: A Guide to Rating Scales and Questionnaires*, 2nd edition; Oxford University Press
- Streiner, D.L. & Norman, G.R. (1989). *Health Measurement Scales: A Practical Guide to Their Development and Use*; Oxford University Press
- Sudman, S. & Bradburn, N.M. (1982). *Asking Questions; A Practical Guide to Questionnaire Design*; Jossey-Bass

- Rea, L.M. & Parker, R.A. (1992). *Designing and Conducting Survey Research: A Comprehensive Guide*; Jossey-Bass
- Salant, Priscilla & Dillman, Don A. (1994). *How to Conduct Your Own Survey*; John Wiley & Sons, Inc.

Module 4: Study Designs II (Surveys)

Background

The word **survey** is used most often today to describe a method of gathering information from a sample of individuals in order to learn something about the large population from which the sample has been drawn. The word comes from Anglo-French – “sur”=over + “veer” = see/view.

Surveys have a wide variety of purposes and they can be conducted in many ways – over the telephone, by mail, or in person. But all surveys have certain characteristics in common.

In a bona fide survey, the sample is chosen using chance processes so that each member of the population has a measurable chance of selection. In this way the results can be reliably projected from the sample to the larger population. By contrast, self-selected opinion polls may be misleading since participants are not scientifically selected; hence, persons with strong opinions (often negative) are more likely to respond.

Information is collected using standardized procedures so that every individual is asked the same questions in more or less the same way. The survey’s intent is not to describe the particular individual who, by chance is part of the sample, but to obtain a composite profile of the population. The sample size depends on the statistical goals and resources available for the survey. Although the formal sample size calculations depend on the type and goal of the survey, there are a few simple rules that can be used. It is astonishing to think that a properly selected sample of 1000 to 1500 individuals can produce accurate estimates even for an entire country with a very small margin of error. That underscores the value of surveys in modern society. For details on choosing a sample size, see the Sample Size Estimation Module.

Surveys provide a speedy and economical means of determining facts about peoples’ knowledge, attitudes, beliefs, expectations and behaviors.

Survey methods can be classified in many ways. One classification is by size and type of sample – special population groups, geographical area, etc. Another classification is by method of data collection – mail, telephone interview and in-person interview. Chart audits – extracting data from sample of medical and other records – come under the heading of survey methods.

A third classification is by survey content – voter preferences, consumer spending, transportation habits, health issues.

Steps in planning a survey

Many interrelated activities are involved in planning a survey.

How to Begin

As with any research, the first step is to articulate the objectives of the investigation. As always, be specific, clear-cut and unambiguous.

Surveys usually require self-reported information. Researchers ask respondents for information on personal characteristics or attributes, knowledge, behaviour, attitudes, and beliefs. Ask yourself whether self-reported information is sufficient? Is it reasonably reliable and believable? Here is a simple, but important question to start with! Can the required information even be collected by a survey? Is the information better gathered from a database of previously collected information? One exception to this is the chart audit, which can also be thought of a proxy survey; the chart plays the role of the subject or person and provides equivalent information.

Also, ask yourself whether your research question is primarily descriptive in nature; that is, is it an assessment of “the way things are” at a particular point in time. If you are considering an intervention that involves a comparison of groups at one point in time, or a comparison of one group at two points in time, then you may use a survey as the data collection instrument but your survey design would be considerably different. See the module on Experimental Designs.

How to Plan Your Survey Administration

As mentioned in the Introduction, survey methods can be classified by method of data collection – mail, telephone interview and in-person interview. Chart audits – extracting data from sample of medical and other records – also come under the heading of survey methods.

Mail surveys can be relatively low in cost. But problems exist in their use when insufficient attention is given to getting high levels of cooperation. Mail surveys can be most effective when directed at particular groups, such as members of a professional association.

Telephone interviews are an efficient method of collecting some types of data and are particularly well-suited when timeliness (not timelessness!) is a factor and the length of the survey is limited. In-person interviews are much more expensive than mail or telephone surveys, but they may be necessary, especially when complex information is to be collected or when health, age, language or education barriers are involved.

Some surveys combine methods: for example, the telephone may be used to screen or locate eligible respondents and then set up in-person interviews.

A new choice for administering surveys is to do so through a Web site or e-mail. Electronic questionnaires have the advantages that the data are already captured electronically and can be entered directly into a database, and that missing and out-of-range values can be rejected. However, they can't be used, yet, to reach the broadest population. Have a look at Survey Monkey at <http://www.surveymonkey.com/Pricing.asp>

The decision on mode of survey administration often comes down to a trade-off between cost, time and level of non-response. Once the mode of survey administration has been determined a questionnaire instrument can then be developed and pre-tested.

How to Get Good Population Coverage

A critical element of any survey is to locate (or cover) all members of the population being studied so that they have a chance to be sampled. To achieve this, a list – called a sampling frame – is usually constructed.

In a mail survey a sampling frame could be all of the postal addresses; in a telephone survey it may be the list of names and telephone numbers.

A sampling frame can also consist of geographic areas if no suitable population list exists.

The quality of the sampling frame – whether it is up-to-date and complete – is the dominant feature for ensuring adequate coverage of the desired population to be surveyed. If the population is not properly “covered” then the generalization of survey results is called into question.

Don't ignore or avoid this step, as so many novice researchers do!

How to Choose a Random Sample

Any "good" survey uses some form of random sampling. Random sampling is based on theory of probability and statistics. Reliable and efficient estimates of population parameters can be made. Whether simple or complex, the goal of a properly designed sample is that all of the units in the population have a known, positive chance of being selected.

There are four main methods of choosing a random sample.

Simple Random Sampling. Each unit in the sampling frame has an equal chance of being selected for the sample. Begin by assigning ID numbers sequentially (starting from 1) to all the units in the sampling frame. Then select a subset of the numbers using random number tables or random number generators on the computer. (e.g. Excel has one). A simple way is to add a column of random numbers to the spreadsheet and then sort the spreadsheet in ascending order by that column. The first N rows will be the N ID numbers you want for your sample. Alternatively, some software will put a list of numbers in random order – that makes it much easier for you; simply select the first N rows.

Systematic Sampling. This is a useful approximation to simple random sampling. Instead of putting the sampling frame in random order, just take every nth unit. For example, if your sampling frame has 1000 units and you wanted 200 of them for your sample, you would select every 5th unit. Be sure to choose a random starting place. If you reach the end of the sampling frame before your sample is complete, just loop back to the beginning and keep going.

Both simple random sampling and systematic sampling assume that the sampling frame is not set up in an order that is related to what is being investigated. A sampling frame in alphabetical order is best for these types of sampling.

Although Simple Random Sampling is the purest and easiest form for analysis, it may also be the most inconvenient method of sample selection. For example, suppose you want to do a province-wide survey. The sampling frame would be composed of residents from all geographic areas of B.C. SRS might select residents of the Lower Mainland in entirety. How could you ensure representation of rural and semi-rural areas? Use Stratified Random Sampling.

Stratified Random Sampling. Divide the sampling frame into subgroups or strata. For example, classify the population by whether they live in large urban, semi-urban, semi-rural and rural areas. Then select a simple random sample within each stratum. Usually the sample size in each stratum should be proportional to the sizes of the stratum. Although this will give you protection against a geographically unrepresentative sample, it may still be very inefficient and expensive, and send you traveling all over the province.

This leads to the final method, Multi-stage Sampling. In the province-wide survey situation, divide the province into large geographic areas. Then select a small sample of these areas, say two or three, at random. Next, divide the large geographic areas into smaller sub-areas, such as electoral constituencies or census metropolitan areas (CMAs). Once again choose a subset of these sub-areas by random selection. Repeat as often as necessary! This will lead to a final set of small areas, chosen by multiple random processes, which you can sample intensively. This method is particularly useful if your surveying requires personal interview.

The sample plan must be described in sufficient detail to allow a reasonably accurate calculation of sampling errors. This makes it valid to draw inferences about the entire population that the sample represents.

What exactly is meant by “inference”? Here’s a little story. The Statson Family sees a herd of black cows grazing on farmland alongside the highway leading out of Belltown. Norman comments, “Since these cows are black I infer that ALL cows are black.” Tina replies, “I think all you can conclude is that THESE cows are black.” The two kids have the last word. They say, “Actually, all you can say with complete certainty is that one side of these cows is black; you can’t see the other side!” Inference is the art of knowing just how far you can generalize your results.

Ideally, the sample size chosen for a survey should be based on how precise the final estimates must be. In practice, usually a trade-off is made between the ideal sample and the expected cost of the survey.

Remember that a large sample that is not drawn by chance methods usually leads to erroneous results. Convenience samples may be convenient but they usually have considerable bias. Think of

phone-in radio talk shows; no matter how many calls a station receives the results cannot be generalized to the entire population. Would you want the callers speaking for you?

How to Determine Sample Size

See the section called “Sample Size Estimation for Descriptive Studies” in the Sample Size Module.

A simple rule of thumb is that the margin of error in a survey is approximately equal to the inverse of the square root of the sample size. That means that a sample of size 100 gives a margin of error of about +/- 10%; a sample size of 400 gives a margin of error of about +/- 5%; and a sample size of 1000 gives a margin of error of about +/- 3%. Sample sizes larger than 1000 give little meaningful improvement in the margin of error.

Note that this sample size is the NET sample size; that is, the number of returned, completed and usable surveys. But not everyone you ask to respond to your survey will do so – a fact that I’m sure comes as a complete surprise to you. So we refer to a response rate; that is, the percentage of people who are asked to participate who actually do.

Different modes of survey administration have different response rates, with mail surveys having the lowest rate. Think about the number of times you’ve received a survey in the mail and the proportion of times you actually completed it and sent it back in (in a reasonable time!).

Response rates have a negative converse, called non-response rates, and the resulting non-response bias. Simply put, non-response bias arises when the people who do respond to your survey have very different answers to those who do not respond.

Rules of thumb vary, but for mail surveys I suggest setting the number of potential respondents at double the number you hope to get returned. Using a procedure of reminder letters, etc. (see Dillman Method below), well-worded questions, an attractive format, and a reasonable length you should be able to get to a response rate of 50 to 60%. So with 2N surveys sent out you should be able to get N completed surveys returned.

When a higher response rate is necessary, options include monetary incentives and telephone reminders.

The problem with this suggestion is that many journals would not publish a study with a 50% response rate. We have found that many journals typically expect at least 70% response rate (and often closer to 80%)!! To get this kind of response rate requires good planning and some sort of incentive structure.

The rationale for aiming for a 60% response rate is as follows. If 60% reply, then only 40% did not; of these, a sizable majority did not respond for reasons of time, lack of interest, lost survey, etc. – reasons that are not connected to the questions you are asking. The bias arising from those who did not answer because they have such divergent views is likely to be small. Remember that in an election, if one party (in a two-party system) gets 60% of the vote, they are deemed to have recorded a “landslide victory.”

How to “Plan In” Quality

Devise ways to keep respondent mistakes and biases to a minimum. For example, memory is important when the respondent is expected to report on past events, so don’t force them to report events that may have happened too long ago to be remembered accurately. Are any of the questions too sensitive? Do they unduly invade privacy? Are they too difficult even for a willing respondent to answer?

For a quality product checks must be made at every step to ensure that the sample is selected according to specifications, that interviewers do their work properly, that information is coded accurately; and that data entry is done correctly, etc.

Coding, data entry and transcription operations are subject to human error and must be carefully and rigorously controlled through verification processes, either on a sample basis or 100% basis. You don’t want a six-year-old grandfather in your data set!

Long questionnaires can lead to respondent fatigue and errors from inattention, refusals, and incomplete answers. They may also contribute to higher non-response rates in subsequent surveys involving the same respondents.

How to Format

The shorter the survey, the more likely it is to get people to respond to it. Time is a precious commodity for most of us. If it can comfortably fit on one page, great!

If it is a multi-page survey, consider print it in booklet format; use standard 8.5" x 11" sheets of paper folded in half to form a 5.5" by 8.5" booklet. This size will fit into a 5 7/8 x 9" envelope, which meets Canada Post's dimensions for First Class Standard mail.

Use the front and back pages (i.e. the cover pages) for material that will stimulate interest in the questionnaire. Do not put questions on it. All instruments must have instructions specifying how they should be filled out.

For each question, particularly if its format differs from other questions, give clear instructions on how to respond (perhaps give an example to demonstrate).

Group the questions in major subject areas together and introduce them by heading or short descriptive statements.

Questions in any topic area that are most likely to be sensitive or objectionable to respondents should come after the less sensitive or objectionable ones.

Questions that the respondent is most likely to see as useful should come first, and the least useful last.

Consistent with the previous statement, put simple demographic questions about age, sex, birth date, etc. at the end of the questionnaire. However, there is a school of thought that says you should put the demographic questions at the beginning to "warm up" the respondent to the process of answering questions. That way you at least have demographic information even if the respondents think the next questions are too sensitive. Keep your study population and type of questions in mind before making this decision.

The first question should be easy to answer, relevant, applicable, and interesting to everyone!

Keep the front cover simple – a study title, directions, study sponsor, etc. Often an organization’s ethics review board will state explicitly the requirements for a cover letter for a survey. Make sure that the title should make the topic sound interesting! And make sure you clearly indicate where and how completed surveys are to be returned. If the sponsoring organization has a letterhead or logo, use it! (Ethics review boards usually make this mandatory.)

For sensitive information, consider using a reply card instead of an identification number on the survey. The reply card will contain identifying information and is to be returned separately to tell the researchers that the survey has been completed. That way the survey itself has no identifying information and assures confidentiality.

How to Schedule (The Dillman Technique)

The time allotted for a survey varies with type of survey and situation. If it is brief and the data are to be collected by telephone from a list already available, the survey can be done in two or three weeks. With longer surveys, more individuals and mail distribution it could take months.

Here are some tips for mail surveys, from Don Dillman.

- Select a mailing date early in the week.
- Avoid mailing close to holidays or anytime in December.
- Plan a follow-up sequence. Dillman’s method (called TDM) involved three carefully timed mailings after the original mail-out of the survey.
- One week afterward, send a postcard reminder to everyone; this is a thank you to respondents and a reminder to non-respondents.
- Three weeks afterward, send a letter and replacement questionnaire to non-respondents only. The cover letter should say that the completed survey has not been received, and asks for participation.
- Seven weeks afterward (if you can wait that long, and have the budget for it), a final mailing is sent, possibly by certified mail, to emphasize the survey’s importance. Another replacement questionnaire is enclosed.
- You can add mailings or telephone calls to increase response rate.
- For mailing to a large sample, consider applying to Canada Post for a Business Reply Application, this will save postage costs as only those surveys and reply cards that are returned are paid for.

Very Important Note: If the survey is being distributed at a clinic, office, or other medical setting, try befriending the office staff. This is by far the greatest single method for increasing participation, particularly if you are not physically present. You will need the help of the office staff at multiple times during the research.

Designing the Questionnaire Items

Surveys are often a combination of custom-written questions and existing previously validated health measurement scales

Books such as “Measuring Health: A Guide to Rating Scales and Questionnaire, 2nd edition” (McDowell and Newell), give a wide range of public domain measurement tools for physical disability and handicap, social health, psychological well-being, depression, mental status testing, pain, general health status, and quality of life.

And nothing beats a good literature search for existing instruments; remember the old cliché about not re-inventing the wheel. Imitation is the sincerest of flattery, even in research!

Some tools are not public domain and require payment to the developers; budget accordingly. Your survey tool will also involve some items you write yourself. Question construction has elements that often appear to be just common sense, but when they are implemented they may involve some subtlety. For example, it is common sense to require that the concepts be clearly defined and questions unambiguously phrased.

The question, “Were you robbed during the last six months?” is apparently straightforward but the question is, in fact, ambiguous. Many respondents are unaware of the legal distinction between robbery (involved personal confrontation of the victim by the offender) and burglary (involving breaking and entering but no confrontation).

Writing questions is an art and deserves a section of its own. See following for a discussion of Writing Questions.

Writing Good Questions

Questions can usually be classified as asking for one of two types of information.

1. What people do, or what they are – their behaviour or attributes.

e.g. Do you ever eat doughnuts? 1 Yes 2 No

2. What people say they want, or what they think is true – their attitudes or beliefs.

e.g. In your opinion, should doughnut consumption be regulated by government?

1 Yes 2 No

Survey questions can also be classified by response format – open-ended or close-ended.

Open-ended questions seek the answer in respondent's own words; there are fewer limits imposed by the researcher. There are no response choices given for the respondent to choose from. These are analogous to essay-type questions on an exam.

Close-ended questions have the respondents choose one or more pre-selected answers. They are quicker and easier to respond to and to analyze. These are analogous to multiple-choice or true/false questions on an exam.

Let's examine these in more detail, with examples.

Open-ended

“What would you like to see offered on the menu of your local donut shop?”

“In your opinion, why did Deviation Donuts open on the outskirts of Belltown instead of in the city centre?”

Open-ended questions take the least amount of effort for the researcher to write, but they can be very demanding for respondents. Respondents have to recall past experiences or comment on issues they may not have considered. Their range of possible responses is very great. Open-ended questions do not provide accurate measurements or consistent comparable information across the whole sample. Preparing the responses for entry into a computer takes a great deal of time; you need to make a master list of responses and then code or assign numbers to each response.

When are open-ended questions useful?

- At the end of a survey, you can ask “Is there anything else you would like to comment on about the subject of this survey?”

- If researchers have little prior knowledge about a topic and so cannot specify response choices.
- If they follow a close-ended question (see below) and explain why a particular answer was selected.
- For estimation – for example, “How many hours per day do you watch TV?”
- For precise information – for example, “What make of car do you drive?”

Close-ended with Ordered Choices

“How do you feel about this statement? ‘This community needs more doughnut shops.’ (Please circle the number of your response.)”

1. Strongly Disagree
2. Mildly disagree
3. Neither agree nor disagree
4. Mildly agree
5. Strongly agree

“What is your present age? (Please circle the number of your response.)”

1. Under 20 years
2. 20-39 years
3. 40-59 years
4. 60-79 years
5. 80 years or over

The distinguishing feature of this kind of question is that each choice represents a gradation of a single concept. They are quite specific, less demanding for the respondent and much easier to code and analyze.

Close-ended with Unordered Choices

“Which best describes the kind of building in which you live? (Please circle the number.)”

- A one-family detached house
- A one-family house attached to at least one other
- An apartment building
- A Mobile home
- A tent

Answer choices are provided but they don't fall on a continuum; respondents choose from discrete unordered categories. Respondents have to consider all the answer possibilities, so it takes a little more time than ordered choices, but not much more. Coding and analysis takes about the same time.

You can also use of this type of question for ranking items, as follows:

"Rank the following types of donuts starting with 1 as your most favorite."

RANK (1 through 7)

- Plain _____
- Glazed _____
- Iced without sprinkles _____
- Iced with sprinkles _____
- Cream-filled _____
- Fruit-filled _____
- Everything on and in it _____

Or: "From the choices above, which is your favorite, second favorite and third favorite from among these choices? (Put the appropriate letter in the box.)"

- () Favorite
- () Second favorite
- () Third favorite

Questions in which respondents must rank unordered response choices are very difficult. Use them carefully and judiciously, especially in telephone surveys!

Partially Close-ended

"Which of the following areas of expenditure do you want to have the highest priority for improvement in this community?"

1. Streets and Roads
2. Sewage Treatment
3. Parks
4. Other (Please specify) _____

Here, answer choices are provided but respondents have the option of creating their own responses. This format rarely yields many additional responses but occasionally it generates new information.

None of the four questions structures is inherently best. Each has merits. Knowledge of the four types will assist you as you sort out what you are really trying to find out. Also, thinking about how you will analyze and report the results will help you choose the question type.

See Appendix for more examples of question format.

Does the Order of Response Choices Matter?

Order can matter, but how it matters differs by survey method. In mail surveys people are more likely to choose from among the first categories listed. In telephone and face-to-face interview they are more likely to choose from among the last. This is called category order effect. The longer the list the less likely it is that the respondent will review all of them. Try to keep your lists from getting too long.

Measuring attitudes and beliefs requires special attention. To deal with uncertainty of responses from time to time, a scaling technique is used. That is, multiple questions are written to address complex issues and then the answers are combined into a single numerical estimate using statistical techniques. Most psychosocial scales of pain, depression, general health, etc. use this approach.

The final step in writing good questions is to consider the exact wording. Salant and Dillman write:

“No one has ever succeeded in compiling a perfect list of rules for writing good questions, although many have tried. The reason such a list is hard to make is that principles that seem sensible often get in each other’s way and only confuse us. For example:

- Be specific.
- Use simple words.
- Don’t be vague.
- Keep it short.
- Don’t talk down to respondents.
- Don’t be too specific.

“Use simple words” and “Don’t talk down to respondents” are both reasonable suggestions, but they get in the way of “Keep it short”. Likewise, the combination of “Don’t be vague” and “Don’t be too specific” puts even the most experienced surveyor in a quandary.

Since no one can provide a set of absolute rules, consider using questionnaires that have worked in other surveys on a similar topic.

Writing good questions usually takes more than one, two, or even three sittings. The shorter the final questionnaire, the longer it takes to get there.

The book by Sudman and Bradburn (*Asking Questions*) has good advice and much more detail on writing good questions.”

To summarize: Avoid vague wording, vague responses, abbreviations or jargon that respondents may not understand, bias from a slanted introduction or unequal comparison or unbalanced response choices (e.g. too little, about right, slightly too high, moderately too high, far too high), sensitive questions (e.g. about income information), double-barreled questions, and answers that are not mutually exclusive. Other than those “minor” pitfalls, writing questions is easy

My advice is: Put yourself in the role of respondent. Is the format attractive? Is it unambiguous? Are there clear directions regarding “branching” questions (i.e. If response to Question 1 is “Yes” go to Question 10, if “No” continue with Question 2.)? Here is a list of steps in creating questionnaires:

- Make a list of variables
- Borrow from other instruments
- Write a draft Revise Pretest (see shortcut #2 following)
- Shorten and revise again
- Repeat the pretest, shortening and revision as often as necessary!
- Code the response choices
- All of the above could/should be done in consultation with others, particularly colleagues in the subject area, and a statistician (for advice on response format, coding, etc.)

Shortcuts to Avoid

There are four types of shortcuts: failure to use a proper sampling procedure; no pretest of the field procedures; failure to follow up non-respondents; inadequate quality control. Here are some examples of failure.

- Using a convenience sample instead of one with a probability design. (e.g. Literary Digest presidential poll of 1936)

- A pretest is the only way of finding out if everything “works”; especially if the survey uses a new procedure or a new set of questions. It is impossible to foresee all possible misunderstandings or biasing effects. The pretest is a small-scale pilot study, usually in the population you are planning to survey, to test the feasibility of the intended techniques or to perfect the questionnaire concepts and wording.
- Failure to follow up non-respondents: it is not uncommon for the initial response rate to most surveys to be under 50 percent. You should plan to return to sample those who did not respond the first time, and then contact a sub-sample of persistent non-respondents to learn why. A low response rate does more damage to the validity of a survey’s results than a small sample since there is no valid way of scientifically inferring the characteristics of the population represented by the non-respondents.

Using Survey Results

Statistics derived from a survey will rarely correspond exactly with the unknown truth. Every survey has a margin of error; the size of margin of error that can be tolerated depends on the recommendations to come out of the data. Sources of error can be classified as Sampling Errors or Non-sampling Errors.

Sampling Errors are due to taking a subset of the population; and, if probability methods are used to select the sample (subset), the size of the sampling error can be estimated. That’s what the margin of error is. If the sample is chosen by convenience or subjective means, no assessment of the margin of error is possible.

Non-sampling errors have no simple and direct method of being estimated. Non-sampling errors can be classified into two groups – firstly, random types or errors whose effects approximately cancel out if fairly large samples are used; secondly, biases, which tend to create errors in the same direction and thus accumulate over the entire sample. With large samples, possible biases are the main cause for concern about survey quality. The main contributing causes of bias are:

Sampling operations – errors or omissions in sampling frame or disproportionate sample rates

- Non-interviews – non-response
- Adequacy of respondent – e.g. caregiver must act as proxy
- Respondent understanding of the concepts

- Lack of knowledge
- Concealment of the truth
- Loaded questions
- Data processing errors
- Conceptual problems
- Interviewer errors

It is good practice to report on the percent of the sample that could not be interviewed and as many of the other factors listed as practicable.

Analysis of survey data begins with descriptive statistics: summaries of each item and scale, with percentages and means, and corresponding confidence intervals where appropriate.

Descriptive bivariate analysis can be done by comparing responses to key outcome variables with demographic variables. For example, do responses differ significantly for male or female respondents? A statistician can discuss these options with you.

Conclusion

Dillman summarizes survey research in two ways. Good survey research requires:

- A large enough sample of respondents drawn randomly, so that sampling error is kept to an acceptable level
- A sample selected in such a way that all members of the population of interest have an equal chance of being selected, thus limiting coverage error
- Questions written and arranged in ways to avoid measurement error
- A response rate high enough to lessen concern about non-response error.

Here is one more way (the last one I, promise!) of summarizing the components of survey research.

Try answering the following questions.

- What is the purpose of my survey? What information do I need?/How will I use it?
- What method seems to be best suited for this particular survey?
- How should I select the sample? What kind of coverage problems are likely to occur?
- How large should the sample be?

- What specific questions will I ask?
- How will I implement my survey to get the highest possible response rate?
- In what form will I present the final results?

The Last Question: Why should you ask questions?

The Last Answer: Why not?