

## **Interview & Questionnaire Research**

Obtain information across a large (or representative) set of individuals (surveys) or from a selected set of individuals using interviews and/or questionnaires.

The results are based on self report. Thus, as with observational research, they can provide a basic set of data for further work.

This is the approach used in opinion and voter polls. However, as a first step to investigating changes in memory over the life span, it has also been used with elderly people to investigate memory. One finding was that the elderly seem to rely less on mnemonic devices and more on external aids than do young adults.

Surveys provide more accurate answers to questions than non-scientific alternatives such as anecdotal evidence and word of mouth.

The *key problems* here are *who to include* in the study and *how to minimize reactivity (and maximize accuracy)* in the participants' answers.

*Who to ask.* The issues here will be oriented around surveying as few people as possible while still getting enough people to be confident that the data are similar to the target population.

*What and How to ask.* Interview and Questionnaire research is based on the idea that people can and will give honest answers to the questions. However, some topics are very sensitive and people may provide answers that reflect cultural norms. We will examine how studies are conducted and how the questions are constructed to minimize reactivity.

# I. Sampling and Surveys

*Who do you survey?* In most cases, we can not ask everyone. Even the U.S. Census uses sampling for its long form questionnaire.

Our goal – to understand how everyone in the target population would answer the questions.

Our problem – there are usually too many people in the population and it is not feasible or possible to ask all of them.

Our solution – *sampling*. A subset of the population is given the survey. If we do this well, their answers should be similar to those that we would get from the target population.

## A) Population and Sample Size

1. The term population here does not mean all people. It means the target group of people from whom we draw our sample and to which we would like our data to generalize.

For example, if the target population is voters in New York for a survey about opinions on state aid for schools, then this excludes certain individuals: those under 18, those who are not residents of New York, those who are not citizens of the U.S., those convicted of a felony and on parole, others who do not vote.

Once the target population is identified, then we can determine *how many people* we need to sample to produce reliable data. We can also determine how best to “choose” the individuals.

## 2. Sample Size and Accurate Results.

The general rule: the more people we sample, the more likely our results will be an accurate representation of the population.

This is based on the laws of probability. As an example, consider the following. I have two quarters. One is a fair coin. If I flip it, it will come up heads half the time and tails the other half. The second coin is biased. When flipped, it will land heads 90% of the time.

a. I choose one coin and flip it. It comes up heads. How confident are you that you can tell which coin I chose?

b. I choose one coin and flip it 10 times. It comes up heads 6 times and tails 4 times. How confident are you that you can tell which coin I chose?

With larger samples (e.g. b. above), we are more likely to get a distribution of heads and tails that reflects the true nature of the coin (50-50 or 90-10).

However, we don't want to sample more people than we have to. That is, survey research takes time and effort. We want to sample only enough people to be confident of the data to a particular precision.

a) Confidence intervals. Since the data are from a sample, and not the entire population, we can estimate the “margin of error” or confidence interval for our data.

In voter poles, especially leading up to an election, they often say that the numbers have a 3 percent margin of error. That is, if the pole showed that 47% of voters favored Bush and 47% favored Gore, then for each candidate the actual percentage for the population is likely to be between 44% and 50% with a high (95%) degree of confidence.

The confidence interval gives the range of values (around the sample value) in which the population value is most likely to occur.

b) **Sample Size.** Sample size is the most important factor in determining the confidence interval. Based on the size of the population, and the size of the confidence interval that we want (the precision of our estimates), we can determine the size of the sample that we need.

Size of population	Precision of estimate		
	+/-3%	+ /-5%	+/-10%
2,000	696	322	92
5,000	879	357	94
10,000	964	370	95
50,000	1,045	381	96
100,000	1,056	383	96
over 100,000	1,067	384	96

This table shows the *sample size* needed for a 95% confidence level with each of three precisions. It is based on relatively conservative assumptions about other factors besides sample size.

As the table shows, for a  $\pm 10\%$  confidence interval, the sample size is not influenced very much by the population size. For the other confidence intervals, the sample size grows a bit with population size, but levels off quickly.

*Sample size does not change substantially with population size.*

Greater precision in the estimates of the population values requires larger samples. As the table shows, to increase precision from  $\pm 10\%$  to  $\pm 3\%$  (cut the error by two-thirds) requires a sample about ten times as large.

These estimates are based on the idea of randomly sampling individuals from a population and some idea about the range of possible answers to the questions (and the homogeneity or heterogeneity of the population).



## B) Sampling Techniques

### 1. Probability or Random Sampling

*The only factor that ultimately determines which members of a population are sampled is chance.*

We can do this three ways: simple, stratified or cluster.

a) *Simple random sampling.* If every member of a population has an equal chance of being included and chance determines who is surveyed, this is simple random sampling.

For example, you start with a list of all the residential phone numbers in Erie County (business numbers have been screened out). From this set, 100 are chosen randomly. This is a random sample of households (with land-line phones) in Erie County.

b) *Stratified Random Sampling* - Identify subgroups within the population. Within each of these strata (subgroups), sample randomly. The sample size in each subgroup needs to be large enough to make it reliable.

For example, a college has a student population of 10,000 and 5 percent are Hispanic. If we do a stratified sample and want to include Hispanics, then we should have a sufficient size sample of this subgroup (Hispanics) to be representative. For example, while a 100 person sample might represent the bulk of the student body very accurately, the 5 Hispanics in a proportional stratified sample might *not* be representative of Hispanics. 5 would be too few for us to be confident that their data is representative. We would “oversample” this subgroup (e.g. 50-90 rather than 5) to make our total sample representative of the subgroups.

c) *Cluster Sampling* – What if you do not have access to lists of individuals? What if you want to do a random sample of the students at UB, but you can not get a listing of the students?

In cluster sampling, you identify clusters of individuals and then randomly choose clusters. All individuals in each cluster are surveyed.

Returning to our UB example, we could treat classes taught in one semester as clusters and randomly choose a set of classes. We would approach the instructors and ask if we could use class time to administer the survey. We would need to be careful to screen respondents so that no one filled out the survey more than once.

This can be done in a series of stages. To survey teachers on the use of technology, we could start out by randomly selecting a set of states, then a set of school districts in the states. Now, we randomly select schools in the district and survey all teachers at each selected school.

## 2. Nonprobability (Nonrandom) Sampling – Convenience Sampling

*Sampling is based on factors other than chance.*

a) *Haphazard sampling* is a type of convenience sampling. In essence, you sample individuals where you can conveniently find them. If someone conducts a survey by approaching patrons at a shopping mall and asking them to fill out a questionnaire, this is a form of haphazard sampling.

This approach is limited because it is restricted to those individuals who have a reason to be at the sampling location at that time.

What would be some of the characteristics likely for the individuals to be surveyed if the location was the Walden Galleria Mall at 10 am on a Tuesday? Would this sample be representative of the larger population of Erie County?

b) *Quota sampling* involves recruiting participants from subgroups in proportion to their presence in the population.

For example, we could collect responses from males and females while making sure that each corresponds to about half of our convenience sample. However, the recruiting is still non random. Thus, while the sample is more balanced with respect to subgroups than the haphazard sample above, it still has most of the problems of haphazard sampling.

## C) Problems in Sampling

If the ideal sample is to be representative of the target population, then it should be randomly drawn (simple, stratified, cluster) from the entire population. Then, everyone sampled should respond.

Problems include:

1. The actual population that you sample from may not be the target population. This is the issue of the *sampling frame*. If you are doing a telephone survey in the hours of 9am to 5pm on weekdays, then you will not reach:

- a) people with out phones
- b) people not home during the day
- c) people with unlisted numbers

These differences could have a small impact on the results or a very large impact.

2. *Not everyone responds.* This is a potentially huge problem. The people who do not respond could differ systematically from those who do (e.g. in income, age, education, marital status, etc.). The lower the response rate, the bigger the potential problem.

This can distort the results so that they are not representative of the target population.

Some considerations:

a) Response rates are higher for telephone surveys than mail surveys.

b) A postcard indicating the survey is coming and a follow-up mailing increase response rates.

c) Call back for unanswered phones or arrange a more convenient time for those who can not respond at that time,

d) Incentives for responding such as entry in a raffle.

e) Response rates are higher if people think that their participation is important.

## D) Some Reasons for Convenience Sampling

Haphazard or Quota (convenience) sampling is, in some cases, as appropriate as random sampling.

1. Sometimes the researcher is not interested in precise estimates for the population. Instead, they are looking for relationships among the variables that they are measuring within their sample.

2. The convenience sample is drawn from the target population.

3. Some basic psychological processes studied with surveys are likely to be the same across different groups, so a convenience sample is sufficient.

4. *HOWEVER*, it can also generate a response from only those individuals with a strong view on the questions being asked. In this case, the sample is *NOT* representative.



## II. Construction of Questions

The key idea here is to keep an eye on the goal of the research. Very few people like filling out a long questionnaire. To keep it short, the research objectives need to be well defined so that only the questions needed to address the objectives are on the survey.

### A) Question types

As an aid in thinking about this, what types of topics might we ask questions about?

1. Facts and Demographics - Factual questions ask people for information about themselves. Basic principle is to ask those needed to verify representativeness of sample and to address the research objectives. Do not ask questions just because you can.

For example, asking about age or gender or party affiliation (voters) are factual questions.

2. Behavior – These questions focus on past, present and future behavior. Ask those needed for research only.

Examples include: *How often do you exercise per week? Do you smoke?*

3. Attitudes and Feelings – These questions focus on how people evaluate and think about topics. Again, ask only those necessary for the research objectives.

Examples include: *Should the tax cut be larger? Should more money be spent on education? On a five point scale, how satisfied were you with the service work done on your car?*

## B) Question Wording

There are some general rules for constructing questions.

1. *Simplicity*. Keep the questions simple and easy to comprehend. Avoid jargon and technical terms. Make sure that the participant has sufficient background to answer the question.

For example, if you are going to ask whether someone agrees or disagrees with a ballot proposition, you will need to provide them with the proposition to read before they can answer the question.

2. Avoid Double-Barreled Questions. These are questions that ask about two things at once. In this case, the participant's answer may be to one part or the other or the participant may simply not respond because they do not have a single position on the two questions.

Example: *Should more money be devoted to preschool and health care for children?* This is really **two** questions that should be asked separately.

3. Avoid Loaded Questions. Questions should be designed to allow alternative responses. Questions designed to elicit a particular response are loaded questions and have no place in scientific research.

Example: *Do you favor eliminating wasteful spending in the military?* is a loaded question. *Do you favor reducing military spending?* gets at some of the same information but does not pre-judge the topic of the question.

The key here is to avoid words and wording with negative emotional and moral connotations or positive emotional or moral connotations. These bias the answers.

4. Avoid Negative Wording. People find questions with negatives confusing.

Example: Mark your degree of agreement, on a five point scale, with each of the following.

*The voters should not approve the proposed school budget.*

Agreeing with this statement means disagreeing with the proposal. To avoid the potential for confusion, questions should be framed in the affirmative (not negative) form:

*The voters should approve the proposed school budget.*

*-or-*

*The voters should reject the proposed school budget.*

5. “Yea-saying” and “Nay-saying”. If all of the questions require a yes response to agree with the topic, and a person answers yes, it could mean that they agree. It could also mean they are simply answering yes to everything. If the participant consistently answers no, does this mean that they are consistently disagreeing or “just saying no”.

To check this, **multiple questions** on a topic which require different responses to reflect agreement with the topic **are asked**.

Example. In addition to our question about the school budget, above, we might ask about whether they agree with:

*The proposed school budget contains too much spending.*

Agreeing with both this and the statement that the budget should be passed appears to be “yea-saying”. Agreeing with passing the budget and disagreeing with this would be a more consistent view on the school budget.

### **III. Response Alternatives for Questions**

What alternatives are offered to the participant to respond to the questions? There are two ways to allow participants to respond: open ended and closed ended (forced choice). Each has their own advantages and disadvantages.

#### **A) Open-Ended Responses**

An open-ended question and response is one where the participant is free to answer in any way they choose.

For example, we could ask: *What do you do when you want to remember something?*

In responding, the participant is free to put their answer in their own words.



1. Advantages – The biggest advantage of the open-ended approach is that you are more likely to get a response that accurately reflects what the participant is currently thinking about the question. In this respect, the open-ended question gets responses that have “ecological validity”.

2. Disadvantages – The open ended question also has a few problems. First, the researcher must devise a system for coding the content of the participants’ answers and check the reliability of the coding. This is time consuming.

Second, sometimes participants can’t think of their own answer or their answer can’t be coded using the coding scheme (the answer may not make sense).

## B) Closed-Ended Questions and Responses

When a set of response alternatives is given to the participant and they choose one alternative for each question, these are close-ended questions. Yes/No questions and multiple choice are examples of closed-ended questions and responses.

1. Number of alternatives. The choice here depends upon the goals of the survey. In voter preferences and public opinion surveys, a simple, two-alternative choice such as yes/no or agree/disagree is often sufficient.

However, humans can often make fine grained distinctions. In behavioral research, a scale is used to allow the participant to indicate the degree of their response.

2. Rating scales. In a rating scale, the participant is offered a set of alternatives that range from one extreme to the other. Examples include:

Strongly Agree      \_\_\_\_\_      \_\_\_\_\_      \_\_\_\_\_      Strongly Disagree

Very Confident      \_\_\_\_\_      \_\_\_\_\_      \_\_\_\_\_      Not Very Confident

---

In the first (a Likert scale), the participant chooses a spot along a 5 point scale. In the second, the participant can choose any point along the line. The response is coded by measuring the distance (e.g. with a ruler) and converting to a percentage of the total distance.

The same type of scale should be used consistently across questions. Avoid changing the type of scale or the number of alternatives from question to question – it is confusing.

3. The Semantic Differential. This scale was devised to assess the meaning of concepts and has been applied to attitudes and feelings: Participants rate a concept on a set of 7-point scales between bipolar (opposite) adjectives (see text for example).

The various adjective pairs cluster into three basic dimensions: *evaluation* (good-bad, wise-foolish); *activity* (active-passive, slow-fast); and *potency* (weak-strong, hard-soft).

The ratings on these scales and dimensions can be used to compare participants' concepts and assess such things as the persuasiveness of a message.

4. Nonverbal scales. With young children or others who can not read, an iconic (nonverbal) scale can be used.



5. Labeled scales. Sometimes, we provide labels for the points on the scale to more clearly define how the participants are to use the scale.

For Example:

_____	_____	_____	_____	_____
Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree

Sometimes, labels are provided to get participants to use a part of the potential range to make finer distinctions. Here is a typical scale used for ranking candidates' potential in graduate school:

<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
Lower	Upper	Upper	Upper	Upper
50%	50%	25%	10%	5%

The scale is done this way since many candidates are very bright and motivated. The attempt is being made to make finer distinctions among the candidates.

Some terms that might be used, such as: never, rarely, sometimes, frequently may have different meanings for different people. In this case, we could quantify the scale with labels.

*Example:*

How often do you exercise for at least 20 minutes?

- \_\_\_\_\_ Less than twice a week
- \_\_\_\_\_ About twice a week
- \_\_\_\_\_ About four times a week
- \_\_\_\_\_ About six times a week
- \_\_\_\_\_ At least once a day

An alternative response scale for people who exercise less frequently might be:

- \_\_\_\_\_ Less than once per month
- \_\_\_\_\_ About once a month
- \_\_\_\_\_ About once every two weeks
- \_\_\_\_\_ About once a week
- \_\_\_\_\_ More than once per week

By being more specific in our labels, we reduce the variability in how participants *interpret* the scale. Furthermore, the range of values on the scale should be chosen to cover the likely range of participants' responses. If the scale does not coincide with their choices, it may create biases in their responses.



## **IV. Finalize the Questionnaire**

A) Neat, consistent appearance. Typed, no spelling errors. Clean, easy to read and follow layout. Use the same set of response alternatives consistently (avoid switching scales from question to question).

Consider the order in which you ask the questions. Group questions that ask about similar topics together. Put important and interesting questions at the beginning to get people interested.

B) Pilot test the questionnaire. Get a small group of participants to fill it out. Ask them to think aloud as they go through it. This means that they are telling you how they interpret the question and how they are choosing to answer. This will allow you to catch questions that are not working as intended and modify them and verify that people understand how to use the response alternatives.

## **V. Administer the Survey**

There are two basic ways to administer a survey:

1. Written
2. Interview

### **A) Written Questionnaire**

The survey is presented in written format to the participants. They read the questions and write their responses.

1. These can be administered in three ways:
  - a) In person – One or more individuals at a time.
  - b) Mail – Send to individuals.
  - c) Internet – Send a link to potential participants or post with topic listed for a search engine to find.

## 2. Advantages:

- a) Relatively inexpensive to administer.
- b) Allow anonymous responses, assuring participant of confidentiality (reducing reactivity).

## 3. Disadvantages:

- a) Require participant to be able to read.
- b) Boring.

#### 4. Special Considerations:

- a) Group administration allows researcher to answer questions. Often done when people are present for another reason so they are a “captive” audience and more likely to finish.
- b) Mailed surveys have very low response rate. Typically require additional steps (such as follow-up, small incentive) to boost response rate.
- c) Internet surveys: Data indicate that results are comparable to those collected with other mailed surveys. Issue is how to reach a representative sample.

## B) Interviews

A person asks the participant the questions and records their answers.

1. This can be done in three ways:
  - a) Face-to-face
  - b) Telephone
  - c) Focus groups – Group interview.
  
2. Advantages:
  - a) Higher response rate. People are more likely to respond to another human. Interviewer can try to convince people to participate.
  - b) Interviewer can provide clarification, ask follow-up questions as necessary. This results in more of questions being answered.

### 3. Disadvantages:

- a) Higher cost than written because of personnel involved.
- b) Potential for *interviewer bias*. Since interviewer is a human, they may alter their voice or behavior as they ask the questions. This can interact with the participants' responses, conveying approval or disapproval. In open-ended questions, this can be especially problematic. This requires careful screening and training of interviewers to reduce such biases.
- c) Reactivity in how individual answers questions.

#### 4. Special Considerations

- a) Face-to-face interviews have higher costs (time and travel). Higher response rate and importance of data must justify cost.
- b) Telephone interviews allow use of a computerized script to reduce interviewer bias. Data can be directly entered on computer for rapid analysis.
- c) Focus groups allow interaction. With open-ended questions, this can trigger different responses and produce a wealth of data. Requires skilled interviewer to elicit participation from all members of group and keep group on task. Group members can be chosen for particular qualities (expertise, interest in topic).

## **VI. Surveys to Study Changes over Time**

In addition to using a survey to sample data at one point in time, they can be used to track changes.

The basic approach is to ask the same questions at different points in time:

A) Sample new groups. Consumer confidence in the U.S. economy is tracked this way. Changes are one economic indicator.

Attitudes of groups in society (aspirations of college freshmen, teenager views on smoking and alcohol, etc.) can be tracked.



B) Panel study tracks a set of individuals over time. In this way, we can study changes in individuals. For example, do attitudes and behaviors of people in a long term relationship change more or less (or become more like their partner) than attitudes for people whose relationship does not last?

The major difficulty here is maintaining contact with the same people over time.

## **Answers for Chapter 6 Sample Questions**

1) – a; 2) – d; 3) – d

## Chapter 7 Sample Exam Questions

- 1) If you want your results to generalize to the general population, then your subjects: a) should only be volunteers b) should be selected by random sampling c) should always be paid for their participation d) b & c above
  
- 2) Cluster sampling involves: a) random selection of naturally occurring groups of individuals b) non-probability sampling (sampling that is not based on random selection) c) a form of quota sampling d) b & c above
  
- 3) The response rate for a survey is important because: a) with lower response rates there is a greater likelihood that biases will distort the results b) high response rates indicate that the individuals who responded are likely to be representative of the target population c) high response rates guarantee that there will be no biases in the questionnaire responses d) all of the above

4) Response rates on surveys are generally higher when: a) an interview is used instead of a mailed questionnaire b) an incentive is used for people who participate c) the participants believe that the survey and its results are important d) all of the above

In the text (Chapter 7, page 134) is a table with a set of questions. Figure out what is wrong with each question.