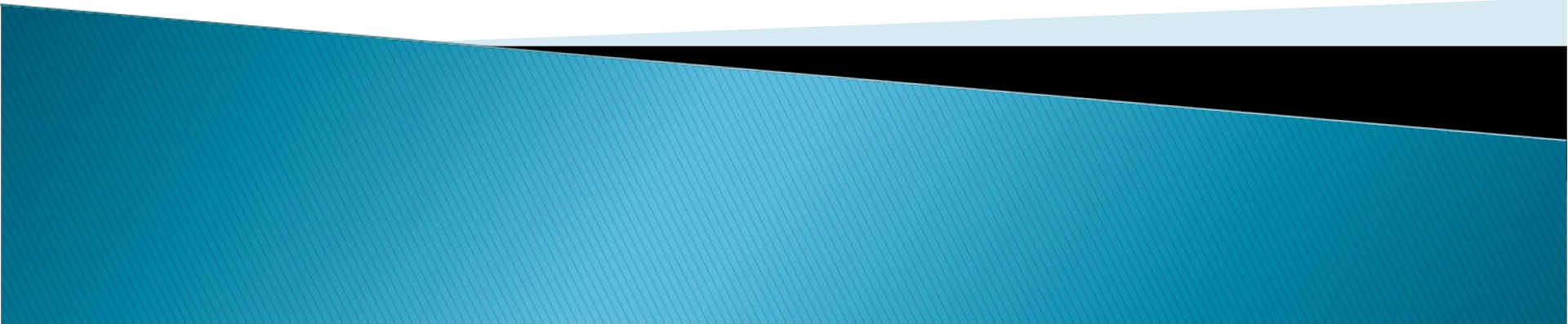


Basic Survey Design Techniques



Objectives

- ▶ To gain an overall understanding of the survey cycle.
- ▶ To understand the basics of survey design and question development.
- ▶ To consider the advantages and disadvantages of certain design choices and their effects on the survey results.
- ▶ To provide practical examples of how survey techniques have been applied and used in audits/program evaluations.

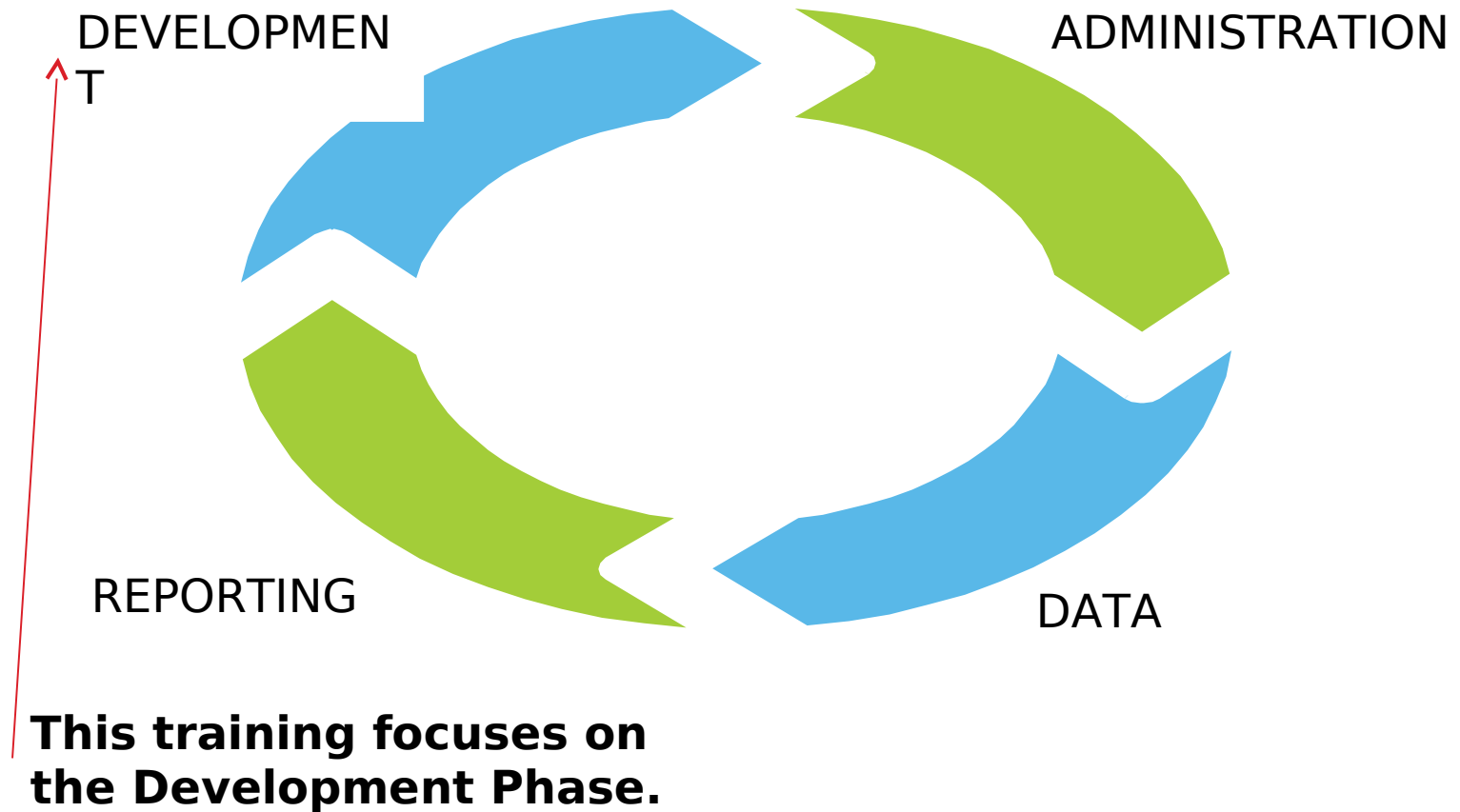
Group Discussion Question #1

What is a survey?

Group Discussion Question #2

When should I use a survey?

The Survey Cycle



Development

During the development phase, you are focusing on questions and issues related to:

- ▶ Overall survey goal and objective
- ▶ Sample selection
- ▶ Methodology
- ▶ Layout
- ▶ Questions and response sets

Administration

During the administration phase, you are focusing on questions and issues related to:

- ▶ Pre-testing
- ▶ Approach
- ▶ Survey administration
- ▶ Close and follow up

Data

During the data phase, you are focusing on questions and issues related to:

- ▶ Data entry
- ▶ Validation
- ▶ Analysis

Reporting

During the reporting phase, you are focusing on questions and issues related to:

- ▶ Reporting the results of your survey

Establishing Goals

- ▶ The first step in any survey is deciding what you want to learn.
- ▶ The goals of the project determine who you will survey and what you will ask them.
- ▶ The more specific your goals, the easier it will be to get usable answers.
- ▶ **If your goals are unclear, the results will probably be unclear.**

Sample Selection

- ▶ Surveys typically involve gathering information from a *sample* of individuals to be able to say something about the population.
- ▶ Lots of sampling methods available. The decision of which method to use is often driven by the study objectives, the level of precision required, and the resources available.
- ▶ You must have a good understanding of how your chosen sampling method affects what you can and cannot say.

SAMPLING METHODS

- ▶ **Probability Samples:** Each member of the population has a known non-zero probability of being selected. Probability methods include random sampling, systematic sampling, and stratified sampling. The advantage of probability sampling is that sampling error can be calculated. Sampling error is the degree to which a sample might differ from the population. When inferring to the population, results are reported plus or minus the sampling error.
 - **Random sampling** is the purest form of probability sampling. Each member of the population has an equal and known chance of being selected. When there are very large populations, it is often difficult or impossible to identify every member of the population, so the pool of available subjects becomes biased.
 - **Systematic sampling** is often used instead of random sampling. It is also called an Nth name selection technique. After the required sample size has been calculated, every Nth record is selected from a list of population members. As long as the list does not contain any hidden order, this sampling method is as good as the random sampling method. Its only advantage over the random sampling technique is simplicity. Systematic sampling is frequently used to select a specified number of records from a computer file.
 - **Stratified sampling** is commonly used probability method that is superior to random sampling because it reduces sampling error. A stratum is a subset of the population that share at least one common characteristic (e.g., gender or geographic location). The researcher first identifies the relevant strata and their actual representation in the population. Random sampling is then used to select a *sufficient* number of subjects from each stratum. "*Sufficient*" refers to a sample size large enough for us to be reasonably confident that the stratum represents the population. Stratified sampling is often used when one or more of the strata in the population have a low incidence relative to the other strata.
- ▶ **Nonprobability Samples:** Members are selected from the population in some nonrandom manner. These include convenience sampling, judgment sampling, quota sampling, and snowball sampling. In nonprobability sampling, it is not possible to calculate a sampling error. Thus, the degree to which the sample differs from the population remains unknown.
 - **Convenience sampling** is used in exploratory research where the researcher is interested in getting an inexpensive approximation of the truth. As the name implies, the sample is selected because they are convenient. This nonprobability method is often used during preliminary research efforts to get a gross estimate of the results, without incurring the cost or time required to select a random sample.
 - **Judgment sampling** is a common nonprobability method. The researcher selects the sample based on judgment. This is usually an extension of convenience sampling. For example, a researcher may decide to draw the entire sample from one "representative" city, even though the population includes all cities. When using this method, the researcher must be confident that the chosen sample is truly representative of the entire population.
 - **Quota sampling** is the nonprobability equivalent of stratified sampling. Like stratified sampling, the researcher first identifies the strata and their proportions as they are represented in the population. Then convenience or judgment sampling is used to select the required number of subjects from each stratum. This differs from stratified sampling, where the strata are filled by random sampling.
 - **Snowball sampling** is a special nonprobability method used when the desired sample characteristic is rare. It may be extremely difficult or cost prohibitive to locate respondents in these situations. Snowball sampling relies on referrals from initial subjects to generate additional subjects. While this technique can dramatically lower search costs, it comes at the expense of introducing bias because the technique itself reduces the likelihood that the sample will represent a good cross section from the population.

Survey Method

Your survey method refers to your method of data collection:

- In-person interviews
 - Telephone surveys
 - Mail surveys
 - Email surveys
 - Online (web-based) surveys
- Each method has advantages and disadvantages.

In-Person Interviews

PROS



- Chance for subject to “experience” the survey.
- Some people may be more willing to tolerate a longer survey if done in person.

CONS



- Significant resource investment required.
- Not as timely as other methods.

Telephone Surveys

PROS



- Faster administration.
- Potential for near universal coverage (96% of homes have a telephone).
- Ties to CATI (computer-assisted telephone interview) software.

CONS



- People are suspicious of telephone surveys (e.g., they suspect a sales call).
- Many new screening tools (e.g., caller ID, voice mail).

Mail Surveys

PROS



- Among the least expensive to administer.
- Not considered as intrusive since respondents answer at their leisure.

CONS



- Mail surveys take a long time (initial mailing, response time, follow up).
- Response rates can often be low or varied.

Online (Web-Based) Surveys

PROS



- Fast and economical.
- Data entry and validation is integrated.
- Filter and other structured question logic can be used.

CONS



- Respondents must have computer access and/or email.
- Generalization problems.
- Software limitations.
- Data confidentiality.

Survey Method

In general, your choice of survey method will depend on several factors:

- ▶ Population being targeted
- ▶ Goals of the survey
- ▶ Speed of administration
- ▶ Need to ask sensitive questions
- ▶ Cost
- ▶ Internet usage of respondents
- ▶ Literacy levels of respondents

Survey Layout – User Principles

- ▶ Navigability – movement through the survey, from one page to another, one section to another, one question to another.
 - User instructions.
 - Section breaks.
 - Question and content groupings.
- ▶ Usability – The ease with which respondents are able to use/interface with the survey.
 - Heavily affected by the chosen survey method.
 - Keep the survey short and simple if at all possible.

Survey Layout

- ▶ Provide instructions. Tell the respondent how to complete the survey, mark the questions, etc.
- ▶ Establish a logical flow to the overall survey, as well as to each section.
- ▶ Questions that the respondent is most likely to perceive as most important should come first.
- ▶ When possible, place difficult or sensitive questions toward the end of your survey.
- ▶ Use transitions or headings to denote different groups of questions.
- ▶ Put demographic questions at the end of the survey.

Question and Response Sets

- ▶ Driven by your survey goals and objectives.
- ▶ Questions and response sets are interrelated.
- ▶ Driven by how you anticipate analyzing and reporting on the data.
- ▶ **Garbage in, garbage out.**
 - You cannot simply throw together a bunch of questions and expect your survey to be successful.

10 Qualities of a Good Question

1. Is relevant.
2. Evokes the truth.
3. Asks for a response on only one dimension.
4. Can reasonably accommodate all possible responses.
5. Has mutually exclusive options.
6. Produces variability of responses.
7. Does not presuppose a current state of affairs.
8. Does not imply a desired response.
9. Does not use emotionally loaded or vaguely defined words.
10. Does not use unfamiliar words or abbreviations.

Example: Lengthy Question

Students must take standardized tests in reading, writing, and mathematics. Some teachers feel that students take too many tests. Others feel more should be required to test students in science and social studies. Do you think that your students take too many standardized tests?

What options exist to make this question less lengthy?

Example: Complex Question

Does it seem possible or impossible that the Cubs will win the World Series, or do you feel they will always find a way to lose?”

What options exist to make this question less complex?

Example: Complex Question

Including yourself, any pets you own, but not including any relatives or friends that do not live in your household for at least 320 days a year, how many household occupants did you buy groceries, by groceries we mean any item of food bought at the grocery store, for in the last year?

What options exist to make this question less complex?

Example: Leading Question

State employees love their jobs. How do you feel about your job with the State?

What options exist to make this question more neutral?

Example: Loaded Question

Do you believe that we should redistribute wealth by allowing tax credits for the rich and wealthy to expire?

What options exist to remove the loaded language in this question?

Response Sets

- ▶ Response sets are an integral part of the question design and must be carefully considered along with the wording of the question.
- ▶ Think about how you are going to code and analyze the data when putting together your response sets.

Response Sets

- ▶ Ordinal Questions – Responses have an underlying order. However, the numeric value assigned to the response may or may not have intrinsic meaning.
 - Ranking/Rating Questions – Responses are measured based on defined intervals (e.g., scale of 1 to 5).
 - ▢ Guttman Scales
 - ▢ Likert Scales
 - ▢ Symantic Differential Scales
- ▶ Nominal Questions – Responses have no underlying order. Numeric values are assigned simply as placeholders, but have no inherent meaning.

Response Sets

- ▶ Single Option Questions – The respondent can only select one response option.
- ▶ Multiple Option Questions – The respondent can select more than one response option (e.g., mark all that apply).
- ▶ Filter Questions – Used to screen and direct respondent's progression through the survey based on a set of logical parameters.
- ▶ Open-Ended Questions – The response format is *unstructured*. The respondent can enter any type of information.

Ordinal Questions

Responses have an underlying order. The numeric value assigned to the response may or may not have intrinsic meaning.

QUESTION: What is your education level?

1. Some high school
2. High school graduate
3. Some college
4. College graduate
5. Some graduate school
6. Master's Degree
7. Doctorate

Ranking Questions

Responses are measured or ranked based on some predetermined interval measure (e.g., scale of 1 to 5).

QUESTION: On a scale of 1 to 5, please rank your liking of the following vegetables from best (5) to worst (1).

	Best 5	4	3	2	Worst 1
Broccoli					
Carrots					
Potatoes					
Arugula					
Spinach					

Likert Scales

Commonly used in public opinion polling. Very useful for measuring respondents' attitudes, feelings, positions, etc. The respondent is asked to select a response in reaction to a statement:

QUESTION: Please select the response that most closely represents your agreement or disagreement with the following statement:

“The customer service representative was helpful.”

Strongly Agree	Agree	Neither Agree Nor Disagree	Disagree	Strongly Disagree
5	4	3	2	1

Guttman Scales

The respondent selects an item that best applies. The list contains items that are cumulative, so that if the respondent agrees to one, he or she probably agrees to the previous statements.

QUESTION: Please select the statement that you most closely agree with:

- ☐ I am willing to travel 1 mile to an eligibility site.
- ☐ I am willing to travel 3 miles to an eligibility site.
- ☐ I am willing to travel 5 miles to an eligibility site.
- ☐ I am willing to travel 10 miles to an eligibility site.

Semantic Differential Scales

The respondent is assessed on a set of bipolar adjective pairs.

QUESTION: Please select the option that best describes your recent customer service experience.

“The customer service representative was...”

	Very Much 5	Somewhat 4	Neither 3	Somewhat 2	Very Much 1	
Helpful						Unhelpful
Friendly						Unfriendly
Polite						Rude

Nominal Questions

Responses have no underlying order. Numeric values are assigned simply as placeholders, but have no inherent meaning.

QUESTION: What region of the country do you live in?

1. Northeast
2. South
3. Midwest
4. Mountains
5. Northeast
6. Southwest

Single Option Questions

The respondent can only select one response option.

QUESTION: What method did you use the last time you purchased airline tickets?

1. Internet site (e.g., Expedia, Orbitz, Travelocity)
2. Travel agency
3. Directly with airline
4. Other

Multiple Option Questions

The respondent can select more than one response option (e.g., mark all that apply).

QUESTION: Which of the following areas best capture the current capital construction needs in your district? (mark all that apply)

- ☐ roof repairs and replacement
- ☐ HVAC repairs and/or replacement
- ☐ correct major structural hazards
- ☐ correct major electrical hazards
- ☐ correct major mechanical hazards
- ☐ fire alarms and sprinklers
- ☐ emergency exits and lighting
- ☐ code improvements
- ☐ accessibility and ADA compliance

- ☐ improve insulation
- ☐ ice, snow, and wind protection
- ☐ repair due to water damage
- ☐ window and door replacement
- ☐ improve temperature/ventilation control
- ☐ aesthetic improvements
- ☐ acoustical control/noise reduction
- ☐ improve classroom technology
- ☐ vehicle loading and unloading zones

Filter Questions

Used to screen and direct respondent's progression through the survey based on a set of logical parameters.

QUESTION #1: Do you like ice cream?

☐ Yes (go to Question #2)

☐ No (go to Question #3)

QUESTION #2: What is your favorite ice cream flavor?

☐ Vanilla

☐ Chocolate

☐ Mint

☐ Rocky Road

QUESTION #3: Do you like cake?

☐ Yes

☐ No

Open-Ended Questions

The response format is *unstructured*. The respondent can enter any type of information.

QUESTION: Please use the space provided below for any additional comments you have about your most recent experience with customer service.

A large, empty rectangular box with a thin black border, intended for the respondent to provide additional comments.

Response Set Principles

Consistency: Responses are orderly, predictable, describable by a few rules, and therefore easy for the respondent to learn and retain.

Completeness: Response choices reasonably reflect the full range of possible answers to a question.

Failure to follow these two principles generally has two outcomes:

1. Respondents will get confused
2. Error is introduced into your survey data

CONSISTENCY PRINCIPLE

Responses are orderly, predictable, describable by a few rules, and therefore easy for the respondent to learn and retain.

QUESTION #1: Please rate your satisfaction with the product purchased:

Very Satisfied	Satisfied	Neither Satisfied Nor Unsatisfied	Unsatisfied	Very Unsatisfied
5	4	3	2	1

QUESTION #2: Please rate your satisfaction with the payment options:

Very Satisfied	Satisfied	Neither Satisfied Nor Unsatisfied	Unsatisfied	Very Unsatisfied
5	4	3	2	1

QUESTION #3: Please rate your satisfaction with customer service:

Very Satisfied	Satisfied	Neither Satisfied Nor Unsatisfied	Unsatisfied	Very Unsatisfied
5	4	3	2	1

+ QUESTION #4: Please rate your satisfaction with the delivery service:

Very Unsatisfied	Unsatisfied	Neither Unsatisfied Nor Satisfied	Satisfied	Very Satisfied
5	4	3	2	1

COMPLETENESS PRINCIPLE

Response choices reasonably reflect the full range of possible answers to a question.

QUESTION #1A: Do you support the death penalty?

- ☐ Yes
- ☐ No
- ☐ Don't know

QUESTION #1B: Do you support the death penalty or life in prison without the possibility of parole?

- ☐ Death penalty
- ☐ Life in prison without the possibility of parole
- ☐ Don't know

QUESTION #2A: Do you use regular or premium gasoline in your car?

- ☐ Regular
- ☐ Premium

QUESTION #2B: Which type(s) of fuel do you use in your car(s)?

- ☐ Regular gasoline
- ☐ Premium gasoline
- ☐ Diesel
- ☐ Other (e.g., hybrid car, electric car)
- ☐ Do not have a car

QUESTION #3A: How many hours of television do you watch in a day?

- ☐ ½ hour or less
- ☐ 1 hour
- ☐ 1 ½ hours
- ☐ 2 hours
- ☐ 3 hours
- ☐ 4 hours

QUESTION #3B: How many hours of television do you watch in a day?

- ☐ 1 hour or less
- ☐ 2 hours
- ☐ 3 hours
- ☐ 4 hours
- ☐ 5-6 hours
- ☐ 7 hours or more

Survey Pretest

The last step in survey questionnaire design is to **pretest** it.

- Ideally you should test the survey with the same kinds of people you will include in the main study.
- If this is not possible, have a few people other than the question writer try the questionnaire.

Pretests are invaluable for revealing problems with question wording, response sets, instructions, and respondent understanding.

Examples