

# 9. QUANTITATIVE METHODS

By Amy Liu

“Research is creating new knowledge.”

–Neil Armstrong, astronaut

**A**n effective methods section is one of the most important tools that a writer can use to help the reader assess the advantages and limitations of a research paper, and to present the evidence that should eventually lead us to trust the results of the research.

The first three sections of any research paper often include:

- 1) Introduction or Statement of Problems and Significance of the Study (see Sarabia’s chapter in this manual for more details),
- 2) Literature Review or Development of Research Questions and/or Hypotheses (see Shigihara and Sarabia’s chapter in this manual for more details ), and
- 3) Methods or Methods of Data Collection and Analysis.

In the introduction, students need to formulate empirical research questions that are feasible, socially important, and scientifically relevant. Then, in the literature review, students need to complete a *thorough* review of relevant literature and the problems to demonstrate

why their study is important and to develop the research questions and/or hypotheses. Then, the reader needs to use an existing dataset and/or collect their own data (qualitative and/or quantitative) to test the hypotheses or study the research questions. In the methods section, a reader can critically assess advantages and disadvantages of the data and analysis to reach the final conclusions of the research. This chapter will focus on the methods section of the research paper.

The methods section explains to the reader what data (information) was used to answer the research question and which technique was used to analyze the data. Hence, methods often refer to:

- 1) methods of **data collection** (how the data was obtained), and
- 2) methods of **data analysis** (how the information was analyzed).

This chapter will focus on this first part of the methods of data collection in quantitative projects (see Morris’ chapter in this Manual for more details on the second part of methods section of data analysis).

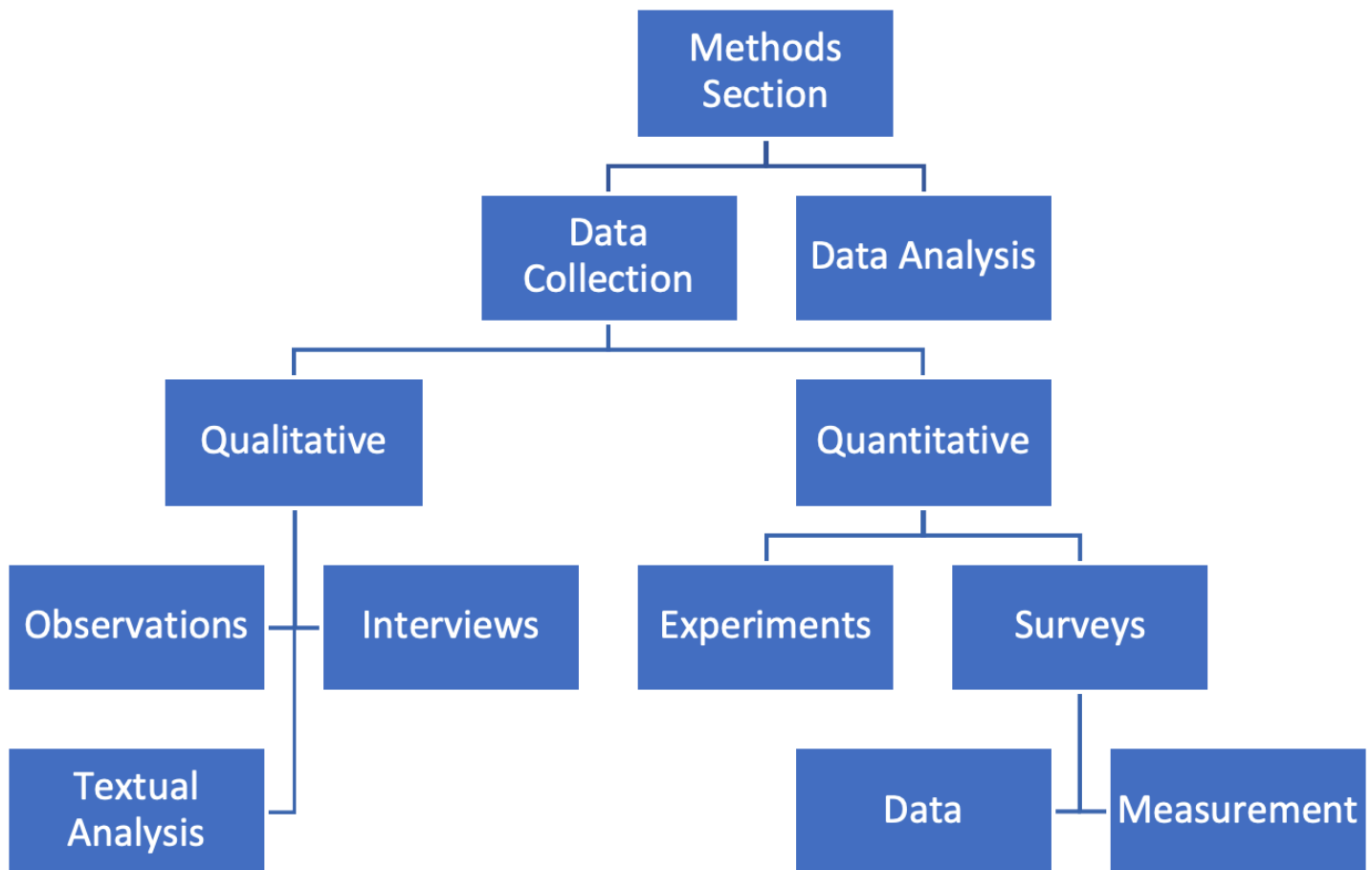
## METHODS OF DATA COLLECTION

In this section, a writer needs to tell readers what method(s) are used to collect the data. Methods of data collection are often divided into quantitative and qualitative methods. Quantitative data is often collected through surveys or experiments—this data is often numerical, and measures quantities (for example, determining the factors that shape social isolation); while qualitative data is often obtained through observations, field work, in-depth interviews, focus interviews, or documents—this data is often descriptive, and measures mechanisms or processes (for example, determining the social processes that lead to social isolation).

Each method of data collection has its advantages and limitations, so this information is critical for the reader to assess the advantages and limitations of the data used and the variables/concepts for the study.

Usually the data collection section also addresses two issues: data and measurement. Since surveys are one of the most often used tools in sociology, the rest of this chapter will focus on surveys as a method of data collection to show the different elements that should be included in the methods section for a quantitative research paper.

Figure 1: Summary of the Methods Section



## 1. Data

In this section, a writer should provide enough information so the reader can determine the reliability and validity of the data as a whole. Reliability implies consistency while a test is valid if it measures what it's supposed to measure. To effectively evaluate the validity and reliability of the data and each variable, a writer should at least provide the following information:

- 1) the organization that collects the data and when the data was collected;
- 2) what is the purpose of the study;
- 3) research design: Is the data cross-sectional, trend, or panel design;
- 4) sampling methods, including sample size, sampling frame, and sample units;
- 5) population of the study and the response rate.

Here is one example that describes the data of a study:

Data from the *2006 Sacramento State Annual Survey of the Region* was used to examine the relationship between availability of affordable health care and age, income, and race. This study is a computer-assisted telephone survey of 1,106 adult residents aged 18 and older. Cluster sampling was utilized to collect data from randomly selected households in the counties of El Dorado, Placer, Sacramento, and Yolo, which consists of the Sacramento region. The telephone interviews were conducted in English and Spanish by more than 30 students at the Institute for Social Research, California

State University, Sacramento. Data was collected from February 17, 2006 to March 8, 2006. This is our 6<sup>th</sup> year collecting the data.

This regional survey was designed to capture quality of life and public opinion concerning a wide range of local, state, and national issues. The population of the study consisted of the residents of the Sacramento region. The sampling frame was the list of household phone numbers, while the primary sampling unit was the household, and the secondary sampling unit is the same as the unit of analysis and the element was the individual resident aged 18 and older. The random sample of 1,106 adult residents was quite representative of the population when compared with the 2000 U.S. Census, and the margin of error at the 95 percent confidence level was  $\pm 3$  percentage points.

Here is another example of data used for a different study:

The General Social Survey (GSS) was conducted every year from 1972 to 1994 (except in 1979, 1981, and 1992) and every other year since 1994. It was carried out by the NORC (National Opinion Research Center) at the University of Chicago. The population of the study is adults age 18 or above living in the United States. Area probability design is used to randomly select respondents in households across the United States to take part in the survey.

This survey includes demographic information and respondents' opinions on a wide range of topics from attitude toward divorce to government spending. This cross-sectional study allows us to

monitor changes and trends in American society. The survey is a face to face interview and takes about 90 minutes to administer. As of 2014, 30 national samples with 59,599 respondents and more than 5,900 variables have been collected.

## **2. Measurement**

A writer needs to clearly and explicitly present all the research questions and/or hypotheses for the study. This kind of background information is needed for the reader to use a dataset that would include all the variables in the study. Each variable used for the study should be of high quality to study each research question and/or to test each hypothesis. This section can be omitted if this information is explicitly included at the end of the literature review. This will provide the necessary information to evaluate how each variable in the dataset is conceptualized, operationalized, and measured. Measurement is the process by which the writer assesses opinions, thoughts, feelings, or behaviors. Conceptualization is a process whereby a concept (thought, opinion/view, behavior) is defined. Operationalization is a process whereby a writer specifies how a variable will be measured.

Here is an example of how a research question is conceptualized, operationalized, and measured:

This study will test the following three hypotheses:

- 1) Younger Americans are more likely to consider affordable healthcare a big problem than older Americans.
- 2) Americans with less income are more likely to consider affordable healthcare a big problem than those with higher income.

- 3) Minorities are more likely to consider affordable healthcare a big problem than non-Hispanic Whites.

This study will use the following control variables to test the following hypotheses:

- 4) Women are more likely to consider affordable healthcare a big problem than men.

- 5) Americans with lower levels of education are more likely to consider affordable healthcare a big problem than Americans with higher education.

- 6) Americans with more liberal political views are more likely to consider affordable healthcare a bigger than Americans with more conservative political views.

The dependent variable in this study is attitude towards affordable healthcare. The categories that respondents were given to answer the question "Availability of healthcare you can afford?" were: "Big Problem" (=1), "Somewhat of a problem" (=2), and "Not a problem" (=3), and were recoded to "Big Problem" (=3), "Somewhat of a problem" (=2), and "Not a problem" (=1).

Age, the first independent variable, was measured by asking respondents the question, "In what year were you born?" The second independent variable is income level and was measured by asking respondents, "Which of the following categories best describes your total household income for 2005, before taxes?" The response categories were "Less than \$20,000" (=1), "\$20,000 but less than \$30,000" (=2), "\$30,000 but less than \$50,000" (=3), "\$50,000 but less than \$75,000" (=4), "\$75,000 but less than \$100,000" (=5), and "\$100,000 or more" (=6). These categories were

recoded to "Less than \$30,000" (=1), "\$30,000 to less than \$75,000" (=2), and "\$75,000 or plus" (=3).

The final independent variable is measured by asking, "Do you consider yourself to be...?" The response categories were "Hispanic/Latino" (=1), "White/Caucasian" (=2), "Black/African American" (=3), "Asian" (=4), "Pacific Islander" (=5), "American Indian or Alaska native" (=6), "Multi-racial" (=7), and "Other" (=8). These categories were recoded to "Non-Hispanic White" (=1), and "Other racial groups" (=0).

The three control variables used as predictors are sex, educational level, and political view. Sex was measured by asking, "Are you "Male" (=1) or "Female" (=2)? After recoding, the categories became "Male" (=0) and "Female" (=1). Educational level was measured by asking "What is the highest year of school you have completed?" The response categories were "Less than high school (grade 11 or less)" (=1), "High school diploma or GED" (=2), "some college" (=3), "Bachelor's degree" (=4), and "Graduate or Professional degree" (=5). Finally, political view was measured by the question "Would you consider yourself to be politically "Liberal" (=1), "Middle of the Road" (=2) or "Conservative" (=3).

Notice that, in the previous example, the hypothesis of "Younger Americans are more likely to consider affordable healthcare a big problem than older Americans" is conceptualized as how age shapes views/opinions about health care.

Figure 2: Relationship between two variables



More concretely, the operationalization (or definition) of "younger" is measuring age (asking people how old they are) and the operationalization (or definition) of "healthcare" is measuring views/opinions about health care (asking people "Availability of healthcare you can afford?"). Age was measured as an interval variable (year of birth) and views about health care was measured as an ordinal variable (a scale from 1 to 3).

### SUMMARY

After the writer finishes this section, ask yourself: Do you, as a reader, get enough information about the data so you can thoroughly analyze the data and have the confidence to say that the data and the variables are appropriate (reliable and valid) to answer the research questions for this study? If the answer is yes, you have done a good job. If the answer is no, you need to revise your methods.