Formulating a Hypothesis

by Sophia

| ☰ | WHAT'S COVERED |

In this lesson, we will take a closer look at the third step of the scientific method for sociology: formulating a hypothesis, which is the step in which sociologists turn their research question into something testable. We will also explore how formulating a hypothesis strengthens our problem solving skill. Specifically, this lesson will cover:

1. Variables

Once you've decided on your research questions and completed your background reading, you will select variables to study and a hypothesis to test. This is where you begin to put your problem solving skills into action.

A **variable** is a characteristic that varies throughout the population as a whole and which can be used to study differences between people and groups. Population variables can include age, class, income level, level of education, race, veteran status, gender, employment status, whether one drives, whether one smokes, country of origin, language, citizen status, region of the country, city or country dweller, or marital status. These variables are different for each individual, but you can batch together large groups of people who all share a certain variable or set of variables. You can also see how variables impact each other by identifying them and sorting the data.

Focusing on particular variables allows you to isolate those characteristics in order to analyze the influence of these characteristics on the population's experience.

> **IN CONTEXT**
>
> If you are studying the effects of wealth transfer through generations, you might look at the relationship between your subjects' income and education levels and their parents' education levels. You might also want to know if all levels of parental education and income have the same effect.
>
> What might you hypothesize about the relationship between the two? You might hypothesize that if a subject's father was educated, the subject will be as educated or higher. You might also hypothesize that if a subject's father was educated, the subject will be likelier to earn a higher income. But there are other variables at play here too: age, gender, race, location, the presence of other similarly educated family members, and many others.

In formulating your hypothesis, you make a statement about how the variable "father's education" is related to the variable "subject's education level." Keep in mind that not all variables are created equal. Some are very critical in explaining a subject's education level, and some aren't, meaning that they don't strongly relate to the outcome that you're trying to explain.

There are two different kinds of variables:

An **independent variable** is the factor that causes the change, or the outcome. You can think of it as the cause. In the example above, the independent variable is the subject's father's education level. It is what drives the change. The **dependent variable** is the effect or the variable that is influenced by the other. In the example, the dependent variable is the subject's education and income level. You are hypothesizing that the father's education level affects their child's education and income level.

📄 TERMS TO KNOW

Variable
A characteristic such as age, education, income, or marriage status that can vary throughout the population.

Independent Variable
The cause of the change, or what drives the change in the dependent variable.

Dependent Variable
The effect of the change; a variable changed by other variables.

2. Formulating a Hypothesis

People commonly try to understand the happenings in their world by finding or creating an explanation for an occurrence, which is what we referred to earlier as common sense. Social scientists may develop a hypothesis for the same reason.

A **hypothesis** is a testable, informed guess about predicted outcomes between two or more variables; it's a possible explanation for specific happenings in the social world and allows for testing to determine whether the explanation holds true in many instances, as well as among various groups or in different places. The hypothesis will often predict how one form of human behavior influences another. The independent variable is the cause of the change, or the variable that influences the other variable. The dependent variable is the effect, or variable that is changed. It depends on the independent variable.

⭐ BIG IDEA

The hypothesis is the researcher's guess—based on background research—about the answer to the research question. A hypothesis often concerns how one thing affects another, which is another way of saying how the independent variable impacts the dependent variable.

In putting together their hypotheses, researchers establish one form of human behavior as the independent variable and observe the influence it has on a dependent variable.

↗ EXAMPLE How does gender (the independent variable) affect income (the dependent variable)? How does religion (the independent variable) affect family size (the dependent variable)? Or to switch it around, how is annual income (the dependent variable) affected by level of education (the

independent variable)?

It is important to note that at this stage we are suggesting relationships between variables, or correlation. We are not yet suggesting that one variable is the cause of another, just that one variable changes when another variable changes in a predictable way.

Examples of Dependent and Independent Variables		
Hypothesis	Independent Variable	Dependent Variable
The greater the availability of affordable housing, the lower the homelessness rate.	Affordable housing	Homelessness rate
The greater the availability of math tutoring, the higher the math grades.	Math tutoring	Math grades
The greater the police patrol presence, the safer the neighborhood.	Police patrol presence	Safer neighborhood
The greater the factory lighting, the higher the productivity.	Factory lighting	Productivity
Individuals with college degrees or higher are less likely to live below the poverty line.	College education	Likelihood of living below the poverty line

As the table shows, an independent variable is the one that influences the other variable. Rather than being "right," sociologists are interested in the relationships between variables. If we were to examine the last example, what other variables might come into play? Would we see similar patterns of income for all college-educated people or are there disparities for racial and ethnic minorities? Gender minorities? First, we must move into the next research steps: designing and conducting a study and drawing conclusions. You'll learn more about these types of research methods in the next section of the course.

▶ 🧩 **Problem Solving: Skill Tip**

We know that correlation does not equal causation; meaning that just because two variables are connected in how they change, that does not mean that one variable is the cause of the other variable. From an example in the table, we cannot assume that an increase of affordable housing lowers the homeless rate; there may be other variables to consider. However, examining the relationships between several variables provides us with information to understand issues on a deeper level. This information enables us to form an educated hypothesis and work toward solving problems.

📄 **TERM TO KNOW**

Hypothesis
A testable, informed guess about predicted outcomes between two or more variables.

3. Sampling

What happens after you gravitate towards a topic, come up with a hypothesis, and hypothesize a relationship between an independent variable and a dependent variable? Most likely it won't be practical to plan on studying an entire population of a city or country. You need to use a sample of the population as a whole.

A **sample** is a smaller group of subjects that ideally represents the population as a whole. You use a sample because it is impossible to go and ask everyone in the whole population, so you have to take a slice of the whole population. The goal, then, is to have a representative sample where all facets of interest of the study are included. The only requirement is that the sample be random.

> ↗ EXAMPLE In your study about your father's education and your income, you wouldn't have a representative sample if you analyzed data from 100 people with highly educated fathers and two people with fathers who didn't finish high school. How could you make conclusions on just the two? A better procedure would be to find out what percentage of the population has finished high school, college, and graduate study, and recruit subjects in those same percentages. If your study is focused on your city, and 40% of adults in your city have finished high school only, then you will want to build a sample where around 40% of your subjects have fathers who finished high school only.

> 🚩 HINT
>
> When selecting a sample of a population for a study, the goal is to select a sample that is representative of the entire population.

One effective way to get a sample is through a technique called **snowball sampling**. In snowball sampling, you find your initial respondents or subjects through acquaintances that you already have in your network. You then use those acquaintances to find their acquaintances, and so on, and the process snowballs.

> ▶ 🧩 **Problem Solving: Skill Reflect**
>
> Consider samples you may study as a sociologist, such as certain demographics. What are some potential questions you can explore with these samples? How might the information help you identify and solve problems?

> 📄 TERMS TO KNOW
>
> **Sample**
>   A smaller group of subjects that ideally represents the larger population as a whole.
>
> **Snowball Sampling**
>   A sampling technique where initial subjects are found through acquaintances, and later subjects are found through acquaintances of acquaintances.

> 📋 SUMMARY
>
> In this lesson, you learned about how sociologists go about **formulating a hypothesis**, including

establishing independent and dependent **variables**. You saw why **sampling** is a useful approach to making a huge population something small enough to work with but still representative. You also strengthened your problem solving skill by beginning to consider educated solutions to problems in society.

Best of luck in your learning!

📄 TERMS TO KNOW

Dependent Variable
  The effect of the change. a variable changed by other variables.

Hypothesis
  A testable educated guess about predicted outcomes between two or more variables.

Independent Variable
  The cause of the change, or what drives the change in the dependent variable.

Sample
  A smaller group of subjects that ideally represents the larger population as a whole.

Snowball Sampling
  A sampling technique where initial subjects are found through acquaintances, and later subjects are found through acquaintances of acquaintances.

Variable
  A characteristic such as age, education, income, or marriage status that can vary throughout the population.