Chapter 11
Nonexperimental Quantitative Research

(Reminder: Don't forget to utilize the concept maps and study questions as you study this and the other chapters.)

Nonexperimental research is needed because there are many independent variables that we cannot manipulate for one reason or the other (e.g., for ethical reasons, for practical reasons, and for literal reasons such as it is impossible to manipulate some variables). Here's an example of an experiment where you could not manipulate the independent variable (smoking) for ethical and practical reasons: Randomly assign 500 newborns to experimental and control groups (250 in each group), where the experimental group newborns *must smoke* cigarettes and the controls do not smoke.

<u>Nonexperimental research</u> is research that lacks manipulation of the independent variable by the researcher; the researcher studies what naturally occurs or has already occurred; and the researcher studies how variables are related.

- Despite its limitations for studying cause and effect (compared to strong experimental research), nonexperimental research is very important in education.

Steps in Nonexperimental Research

The pretty much the same as they were in experimental research; however, there are some new considerations to think about if you want to be able to make any cause and effect claims at all (i.e., that an IV--->DV).

1. *Determine the research problem and hypotheses to be tested*. Note: it is important to have or develop a theory to test in nonexperimental research if you are interested in making any claims of cause and effect. This can include identifying mediating and moderating variables (see Table 2.2 on page 36 for definitions of these two terms).

2. *Select the variables to be used in the study*. Note: in nonexperimental research you will need to include some control variables (i.e., variables in addition to your IV and DV that measure key extraneous variables). This will help you to help rule out some alternative explanations.

3. *Collect the data*. Note: longitudinal data (i.e., collection of data at more than one time point) is helpful in nonexperimental research to establish the time ordering of your IV and DV if you are interested in cause and effect.

4. *Analyze the data*. Note: statistical control techniques will be needed because of the problem of alternative explanations in nonexperimental research.

5. *Interpret the results*. Note: conclusions of cause and effect will be <u>much weaker</u> in nonexperimental research as compared to strong experimental and quasi-experimental research because the researcher cannot manipulate the independent variable in nonexperimental research.

When examining or conducting nonexperimental research, it is important to watch out for the post hoc fallacy (i.e., arguing, after the fact, that A must have caused B simply because you have observed in the past that A preceded B).

- By the way, post hoc or inductive reasoning is fine (i.e., looking at your data and developing ideas to examine in future research), but you must always watch out for the fallacy just mentioned and you must remember to empirically test any hypotheses that you develop after the fact so that you can check to see whether your hypothesis holds true with new data. In other words, after *generating* a hypothesis, you must *test* it. (This last point goes back to Figure 1.1 on page 18 showing the research wheel.)

Independent Variables in Nonexperimental Research

This includes variables that cannot be manipulated, should not be manipulated, or were not manipulated.

- Here are some examples of categorical independent variables (IVs) that cannot be manipulated—gender, parenting style, learning style, ethnicity, retention in grade, personality type, drug use.
- Here are some examples of quantitative IVs that cannot be manipulated—intelligence, age, GPA, any personality trait that is operationalized as a quantitative variable (e.g., level of self-esteem).
- It is generally recommended that researchers should not turn quantitative independent variables into categorical variables.

Simple Cases of Causal-Comparative and Correlational Research

Although the terms causal-comparative research and correlational research are dated, it is still useful to think about the simple cases of these (i.e., studies with only two variables). There are four major points in this section:

1. In the simple case of causal-comparative research you have one *categorical* IV (e.g., gender) and one quantitative DV (e.g., performance on a math test).
    - The researcher checks to see if the observed difference between the groups is statistically significant (i.e., not just due to chance) using a "t-test" or an "ANOVA" (these are statistical tests discussed in a later chapter; they tell you if the difference between the means is statistically significant; they are discussed in chapter 16).

2. In the simple case of correlational research you have one *quantitative* IV (e.g., level of motivation) and one quantitative DV (performance on math test).
    - The researcher checks to see if the observed correlation is statistically significant (i.e., not due to chance) using the "t-test for correlation coefficients" (it tells you if the relationship is statistically significant; it is discussed in chapter 16).

- Remember that the commonly used correlation coefficient (i.e., the Pearson correlation) only detects linear relationships.

3. It is essential that you remember this point: Both of the simple cases of nonexperimental research are seriously flawed if you are interested in concluding that an observed relationship is a causal relationship.
   - That's because "observing a relationship between two variables is not sufficient grounds for concluding that the relationship is a causal relationship." (Remember this important point!)

4. You can improve on the simple cases by controlling for extraneous variables and designing longitudinal studies (discussed below).
   - And once you move on to these improved nonexperimental designs, you should drop the "correlational" and "causal-comparative" terminology and, instead, talk about the design in terms of the research objective and the time dimension (which is discussed below, and summarized in Table 11.3)

The Three Necessary Conditions for Cause-and-Effect Relationships

It is essential that your remember that researchers must establish three conditions if they are to make a defensible conclusion that changes in variable A cause changes in variable B. Here are the conditions (which have been stated in previous chapters) in a summary table:

TABLE 11.1
The Three Necessary Conditions for Causation

Researchers must establish three conditions if they are to conclude that changes in variable A cause changes in variable B.

Condition 1:	Variable A and variable B must be related (the relationship condition).
Condition 2:	Proper time order must be established (the temporal antecedence condition).
Condition 3:	The relationship between variable A and variable B must *not* be due to some confounding extraneous or "third" variable (the lack of alternative explanation condition).

Applying the Three Necessary Conditions for Causation in Nonexperimental Research

Nonexperimental research is much weaker than strong and quasi experimental research for making justified judgments about cause and effect.

- It is, however, quite easy to establish condition 1 in nonexperimental research—just see if the variables are related For example, Are the variables correlated? or Is there a difference between the means?.
- It is much more difficult to establish conditions 2 and 3 (especially 3).
- When attempting to establish condition 2, researchers use <u>logic and theory</u> (e.g., we know that biological sex occurs before achievement on a math test) and <u>design approaches</u> that are covered later in this chapter (e.g., longitudinal research is a strong design for establishing proper time order).
- <u>Condition 3 is a serious problem in nonexperimental research</u> because it is always possible that an observed relationship is "spurious" (i.e., due to some confounding extraneous variable or "third variable").
- When attempting to establish condition 3, researchers use <u>logic and theory</u> (e.g., make a list of extraneous variables that you want to measure in your research study), <u>control techniques</u> (such as statistical control and matching), and <u>design approaches</u> (such as using a longitudinal design rather than a cross-sectional design).
- The rest of the chapter will be explaining these points.
- To get things started, you need to understand the idea of <u>controlling for a variable</u>. Here is an example: first, Did you know that there is a correlation between the number of fire trucks responding to a fire and the amount of fire damage? Obviously this is not a causal relationship (i.e., it is a spurious relationship). In Figure 11.2 below, you can see that after we control for the size of fire, the original positive correlation between the number of fire trucks responding and the amount of fire damage becomes a zero correlation (i.e., no relationship).

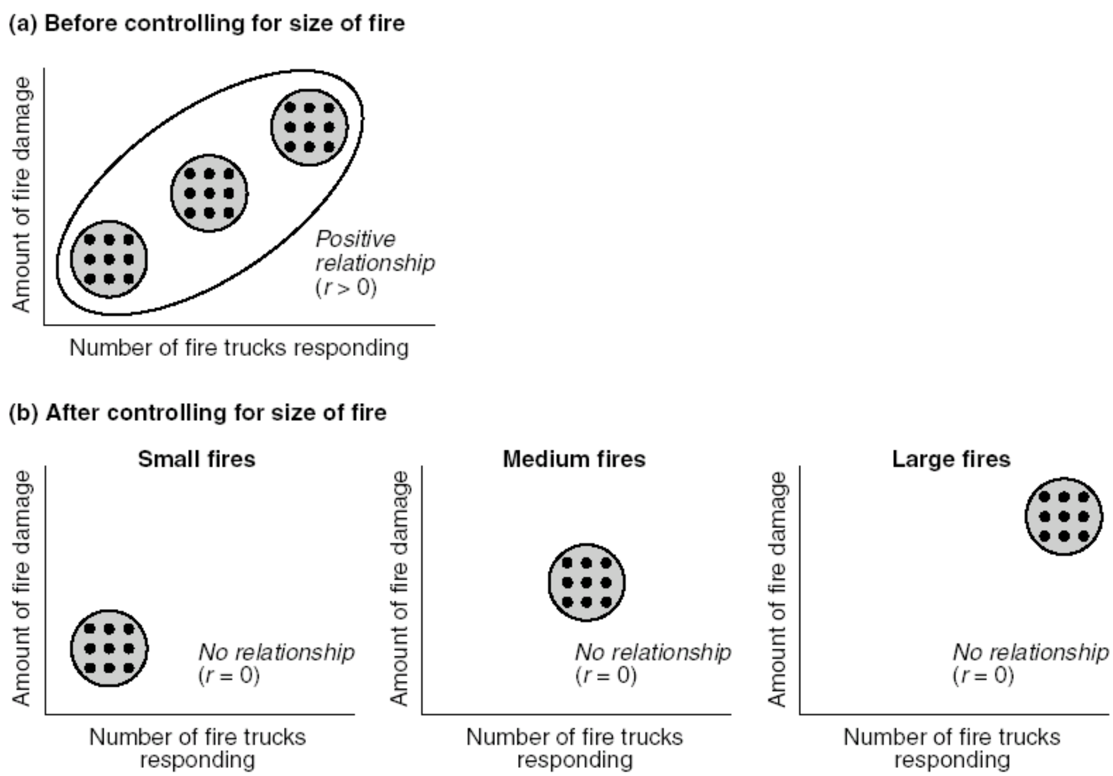

FIGURE 11.2 Relationship between amount of fire damage and number of trucks responding before and after controlling for size of fire. We controlled for size of fire by examining the original relationship at different levels of size of fire. The original relationship disappears.

- Here is one more example of <u>controlling for a variable</u>: There is a relationship between gender and income in the United States. In particular, men earn more money than women. Perhaps this relationship would disappear if we controlled for the amount of education people had. What do you think? To test this alternative explanation (i.e., it is due not to gender but to education) you could examine the average income levels of makes and females ate each of the levels of education (i.e., to see if males and females who have equal amounts of education differ in income levels). If gender and income are still related (i.e., if men earn more money than women at each level of education) then you would conclude make this conclusion: "After controlling for education, there is still a relationship between gender and income." And, by the way, that is exactly what happens if you examine the real data (actually the relationship becomes a little smaller but there is still a relationship). Can you think of any additional variables you would like to control for? That is, are there any other variables that you think will eliminate the relationship between gender and income?

Techniques of Control in Nonexperimental Research

We discuss three ways to control for extraneous variables in nonexperimental research.

1. Matching.
   - A "matching variable" is an extraneous variable you wish to control for (e.g., gender, income, intelligence) and you are going to use it in the technique called matching.
   - If you have two groups (i.e., your IV is categorical), you could attempt to find someone like each person in group one on the matching variable and place these individuals into group two. In other words, you could in effect construct a control group.
   - If your IV is a quantitative such as level of motivation and you want to see if motivation is related to test performance, you might decide to us GPA as your matching variable. To do this, you would have to find individuals with low, medium, and high GPAs at the different levels of motivation as shown in the following table.

You could do this by finding people for each of the cells of the following table:

	Low Motivation	Medium Motivation	High Motivation
Low GPA	15 people	15 people	15 people
Medium GPA	15 people	15 people	15 people
High GPA	15 people	15 people	15 people

- Technically speaking, matching makes your independent variable and the matching variable uncorrelated and unconfounded. What this means is that if you still see a relationship between your IV and your DV you can conclude that it is not because of the matching variable because you have controlled for that variable.

2. Holding the extraneous variable constant.

- If you use this strategy, you will include in your study participants that are all at the same constant level on the variable that you want to control for. For example, if you want to control for gender using this strategy, you would only include females in your research study (or you would only include males in your study). If there is still a relationship between your IV and DV (e.g., motivation and test grades) you will be able that the relationship is not due to gender because you have made it a constant (by only including one gender in your study).

3. Statistical control (it's based on the following logic: examine the relationship between the IV and the DV at each level of the control/extraneous variable; actually, the computer will do it for you, but that's what it does).

- One type of statistical control is called partial correlation. This technique shows the correlation between two quantitative variables after statistically controlling for one or more quantitative control/extraneous variables. Again, the computer program (such as SPSS) does this for you.

- A second type of statistical control is called ANCOVA (or analysis of covariance). This technique shows the relationship between a categorical IV and a quantitative DV after statistically controlling for one or more quantitative control/extraneous variables. Again, you just have to figure out what you want to control for and collected the data; the computer will actually do the ANCOVA for you.

Now I am going to talk about the two key dimensions that should be used in constructing a nonexperimental research design: the time dimension and the research objective dimension. (Note that these dimensions eliminate the need for the terms correlational and causal-comparative in nonexperimental research.)

The Time Dimension in Research

Nonexperimental research can be classified according to the time dimension. In particular, Figure 11.3 shows and summarizes the three key ways that nonexperimental research data can vary along the time dimension; in cross-sectional research the data are collected at a single point in time, in longitudinal or prospective research data are collected at two or more time points moving forward, and in retrospective research the researcher looks backward in time to obtain the desired data. .

FIGURE 11.3 Depiction of cross-sectional, longitudinal, and retrospective research designs

Design Type	Depiction	Example
Cross-sectional	O_1	Data are collected at one point in time on several variables such as gender, income, and education.
Longitudinal	$O_1 \quad O_2 \ldots O_n$	Data are collected in a forward direction over time on one or more variables such as gender (O_1), IQ (O_1), discipline problems in middle school (O_1), high school GPA (O_2) and dropout status (O_2).
Retrospective	$O_{T-n} \ldots O_{T-1} \quad O_T$	Data are collected that represent present and past status on variables such as dropout (O_T), use of drugs (O_{T-1}), and GPA (O_{T-1}).

O stands for collection of data on independent variables, control variables, and/or dependent variables. n stands for the final time period data are collected for the longitudinal design, "T" stands for the current time, "T–1" stands for a time in the past, and "T–n" stands for the beginning time point in the past examined in the retrospective design.

Classifying Nonexperimental Research by Research Objective

The idea here is that nonexperimental can be conducted for many reasons. The three most common objectives are description, prediction, and explanation.

- Descriptive nonexperimental research is used to provide a picture of the status or characteristics of a situation or phenomenon (e.g., what kind of personality do teachers tend to have based on the Myers-Briggs test?).
- Predictive nonexperimental research is used to predict the future status of one or more dependent variables (e.g., What variables predict who will drop out of high school?).
- Explanatory nonexperimental research is used to explain how and why a phenomenon operates as it does. Interest is in cause-and-effect relationships.

One type of explanatory research that I want to mention in this lecture is called theoretical modeling or causal modeling or structural equation modeling (those are all synonyms). Causal modeling (i.e., constructing theoretical models and then checking their fit with the data) is commonly used in nonexperimental research.

- Causal modeling is used to study direct effects (effect of one variable on another). Here is a way to depict a direct effect: X -----> Y
- Also used to study indirect effects (effect of one variable on another through an intervening or mediator variable). Here is a way to depict an indirect effect of X on Y: X -----> I ----->Y
- A strength of causal modeling in nonexperimental research is that they develop detailed theories to test.
- A weakness of causal modeling in nonexperimental research is that the causal models are tested with nonexperimental data, which means there is no manipulation, and you will recall that experimental research is stronger for studying cause and effect than nonexperimental research.
- Also, causal models with longitudinal data are generally better than causal models with cross-sectional data.

Classifying Nonexperimental Research by Time and Research Objective

So we talked about two key dimensions for classifying nonexperimental research: the time dimension and the research objective dimension. Notice that these two dimensions can be crossed, which forms a 3-by-3 table, which results in 9 types of nonexperimental research. Here is the resulting Classification Table:

TABLE 11.3 Types of Research Obtained by Crossing Research Objective and Time Dimension

	Time Dimension		
Research Objective	Retrospective	Cross-Sectional	Longitudinal
Descriptive	retrospective, descriptive study. (type 1)	cross-sectional, descriptive study. (type 2)	longitudinal, descriptive study. (type 3)
Predictive	retrospective, predictive study. (type 4)	cross-sectional, predictive study. (type 5)	longitudinal, predictive study. (type 6)
Explanatory	retrospective, explanatory study. (type 7)	cross-sectional, explanatory study. (type 8)	longitudinal, explanatory study. (type 9)

Source: Adapted from Johnson (2001).

If the above table seems complicated, then note that all you really have to do is to remember to answer these two questions:

1. How are your data collected in relation to time (i.e., are the data retrospective, cross-sectional, or longitudinal)?
2. What is the primary research objective (i.e., description, prediction, or explanation)?

Your answers to these two questions will lead you to one of the nine cells shown in the above table.