

Chapter 8 Answers to Study Questions

8.1. What is a confounding variable, and why do confounding variables create problems in research studies?

An extraneous variable is a variable that MAY compete with the independent variable in explaining the outcome of a study. A confounding variable (also called a third variable) is a variable that DOES cause a problem because it is empirically related to both the independent and dependent variable. A confounding variable is a type of extraneous variable (it's the type that we know is a problem, rather than the type that might potentially be a problem).

8.2. Identify and define the four different types of validity that are used to evaluate the inferences made from the results of quantitative studies.

1. Statistical conclusion validity.

- Definition: The degree to which one can infer that the independent variable (IV) and dependent variable (DV) are related and the strength of that relationship.

2. Internal validity.

- Definition: The degree to which one can infer that a causal relationship exists between two variables.

3. Construct validity.

- Definition: The extent to which a higher-order construct is well represented (i.e., well measured) in a particular research study.

4. External validity.

- Definition: The extent to which the study results can be generalized to and across populations of persons, settings, times, outcomes, and treatment variations.

8.3. What is statistical conclusion validity, and what is the difference between null hypothesis significance testing and effect size estimation?

Statistical conclusion validity is the degree to which one can infer that the independent variable (IV) and dependent variable (DV) are related and the strength of that relationship.

- Null hypothesis significance testing (a major topic in Chapter 16) is used to determine whether we can reject the null hypothesis (which says there is NO relationship present) and accept the alternative hypothesis (which says there IS a relationship). Note that when we reject the null hypothesis, the researcher says that the relationship is statistically significant.
- Effect size estimation involves the use of some type of effect size indicator (such as the percentage of variance explained, the size of the correlation, the size of the difference between two group means, etc.) to inform you of the size or strength of an observed relationship.
- In other words, null hypothesis testing tells us whether we have observed a real (i.e., non-chance) relationship, and an effect size indicator tells us how strong a significant relationship is.

8.4. What is internal validity, and why is it so important in being able to make causal inferences?

Internal validity is defined as the “approximate validity with which we infer that a relationship between two variables is causal” (Cook and Campbell, 1979, p.37). Often in research we want to be able to make causal inferences (i.e., state that two variables are causally related). To do this, we must have internal validity which is obtained through the use of design features and control techniques. The best designs are the strong experimental designs, and the best control technique is random assignment to groups. Note that it is essential for us to be able to make causal inferences because doing so helps us to know how to improve the world (e.g., find effective teaching practices, find ways to help people reach positive mental health, etc.). If you listen to your everyday language, you will see that cause and effect is embedded in your daily thinking.

8.5. What are the two types of causal relationships, and how do these two types of causal relationships differ?

1. Causal description involves describing the consequences of manipulating an independent variable.
2. Causal explanation involves more than just causal description. It also involves explaining the mechanisms (e.g., see the discussion of intervening/mediating variables in Chapter 2) through which and the conditions (e.g., see the discussion of moderator variables in Chapter 2) under which a causal relationship holds. To see more on mediating and moderating variables, (look at Table 2.2 or [just click here](#)).

8.6. What type of evidence is needed to infer causality, and how does each type of evidence contribute to making a causal inference?

The three necessary conditions for cause and effect are 1) Variable A and variable B must be related (the relationship condition), 2) Proper time order must be established (the temporal antecedence condition), and 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). If you are going to argue that causation is occurring, then you must address each of the three conditions. You must also make sure that none of the threats to internal validity discussed in the chapter represents an alternative explanation for the research results.

8.7. What is an ambiguous temporal precedence threat, and why does it threaten internal validity?

If you look again at the three necessary conditions for cause and effect listed in the last question, you will see that ambiguous temporal precedence simply means that you have not met condition two (i.e., you have not established proper time order with your variables). For example, if cancer was observed to occur before smoking, you would have failed to meet the requirement of proper time order (smoking must occur before the onset of cancer if you plan on arguing that smoking causes cancer).

- Ambiguous temporal precedence is formally defined as the inability of the researcher (based on the data) to specify which variable is the cause and which is the effect.

- If you cannot meet this necessary condition but your variables are related then you need to just say that the two variables are related (i.e., you cannot say that they are causally related).

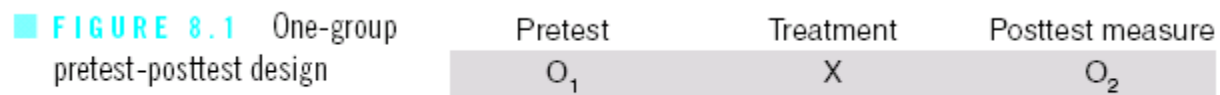
8.8. What is a history threat, and how does it operate?

Whenever you measure your dependent variable with a pretest followed by implementation of a treatment followed by the measurement of the dependent variable again at the posttest, you should worry about the history effect. You hope to conclude that the difference between the pretest and the posttest is due to the treatment, but the history threat can cause problems.

- The history threat refers to any event, *other than the planned treatment* event, that occurs between the pretest and posttest measurement and has an influence on the dependent variable.
- If both a treatment and a history event occur between the pretest and posttest, you will not know whether the observed difference between the pretest and posttest is due to the treatment or due to the history event. In short, those events are confounded.

8.9. What is a maturation threat, and how does it operate?

Let's assume again you are using the design shown in Figure 8.1 and shown here:



In the above design, the effect of the treatment is estimated by the change measured from the pretest to the posttest on the outcome (i.e., dependent) variable.

Maturation is a problem that can threaten the researcher's ability to conclude that the treatment caused or produced the change from pretest to posttest.

- Maturation is any physical or mental change that occurs over time that affects performance on the dependent variable.
- Children are especially prone to maturation because they are naturally changing so rapidly.
- In short, if you have a maturation effect operating, it is confounded with the treatment and you do not know whether the change observed from pretest to posttest is due to the treatment or simply due to maturation.

8.10. What is a testing threat, and why does it exist?

The testing effect is another threat that can occur when using the design shown above in Figure 8.1.

- Testing is any change in the scores on the second administration of a test that results from having previously taken the test.

- Again, in the one-group pretest-posttest design shown in Figure 8.1, testing would be a threat if the participants were affected by having taken the pretest. That effect would be confounded with the treatment effect.

8.11. What is an instrumentation threat, and when would this threat exist?

An instrumentation effect is another problem that can occur when using the design shown in Figure 8.1.

- Instrumentation is any change that occurs in the way the dependent variable is measured over time.
- Again, in the one-group pretest-posttest design shown in Figure 8.1, instrumentation would be a threat to internal validity if the way the dependent variable was measured changed from time one (pretest) to time two (posttest). The effect would be confounded with the treatment effect.

8.12. What is a regression artifact threat, and why does this threat exist?

Another problem that can occur when using the design shown in Figure 8.1 is the regression artifact effect (sometimes called “regression to the mean”).

- Regression artifacts is defined as the tendency of very high scores to become lower over time and for very low scores to become higher over time.
- Again, in the one-group pretest-posttest design shown in Figure 8.1, regression artifacts would be a threat if you had selected participants with extremely high scores (e.g., on the SAT). This is because some of these high scorers probably did a little better than they would normally do, and their scores will be a little lower when they take the test again. This regression artifact would be confounded with any treatment effect.

8.13. What is a differential selection threat, and when would this threat exist?

Differential selection is defined as selecting participants for various treatment groups that have different characteristics.

- This is not a threat to the design we have been discussing so far (i.e., the one-group pretest-posttest design shown in Figure 8.1). That’s because that design does give you a read on the change for the people in the study.
- This is a threat for the design shown in Figure 8.2 (shown below), as long as there is no random assignment to the groups (because random assignment will prevent differential selection from occurring because it will, on average, make the groups the same).

■ **FIGURE 8.2** Two-group design comparing an experimental group that receives a treatment condition with a control group that does not receive the treatment condition

	Treatment	Posttest
Group A	X	O ₂
Group B	—	O ₂

- When you have two or more groups (and no random assignment to the groups), any difference observed between the groups might be due to the characteristics of the people in the different groups rather than the treatment. In other words, the selection variables such as those shown in Table 8.1 might be the real reason that the groups differ. In short,

you cannot conclude that the observed differences between the groups at the posttest is due to the different treatments because it is confounded with participant characteristics

8.14. What is meant by an additive and interactive effect as a threat to internal validity?

Additive and interactive effects refers to the fact that the threats to internal validity can sometimes combine to produce a bias in the study which threatens our ability to conclude that the independent variable is the cause of the differences in the dependent variable.

- One example of this kind of threat is called the selection-history effect.
- The selection-history effect occurs when an event occurs in a multi-group design (such as the one shown in Figure 8.2) that differentially affects the different comparison groups. For example, if someone came into one group's room and shouted the president has been shot by did not go into the other group's room, we would expect a differential effect.
- Another example is the selection-maturation effect.
- The selection-maturation effect occurs when an event occurs in a multi-group design where the participants in one of the groups experience a different rate of maturation than the participants in a different group.

8.15. How does differential attrition threaten internal validity?

Attrition simply refers to the fact that participants sometimes drop out of a research study.

- Differential attrition can occur in a multi-group design(not a single group design), and it is defined as the differential loss of participants from the various comparison groups.
- This is a problem in the design shown in Figure 8.2 because the groups can become different because of the people dropping out rather than just the treatment. In other words, the differences due to differential attrition and the differences due to the treatments are confounded.

8.16. What is external validity, and why is it important?

External validity is the degree to which the results of a study can be generalized to and across populations of persons, settings, times, outcomes, and treatment variations. In short, external validity has to do with generalizing.

8.17. What is population validity, and why is it difficult to achieve?

Population validity is the degree to which the results of the study can be generalized to individuals that were *not included* in the study. It is difficult to achieve because, first, in experimental research it is usually not feasible to randomly select from the target population (e.g., how would you get a random sample of people with dyslexia?). Also, even if we get a random sample of the accessible population (i.e., the research participants who are available for participation in the research study), we still would often find that the accessible population is different from the target population (the larger population to whom the study results are to be generalized).

8.18. What is ecological validity?

Ecological validity is the degree to which one can *generalize* the results of the study across different settings and different contexts.

8.19. What is temporal validity?

Temporal validity is the degree to which one can *generalize* the results of the study across time (e.g., do results found previously still apply and will results found today apply in the future?).

8.20. What is treatment variation validity, and why can this be a threat to external validity?

Treatment variation validity is the degree to which one can *generalize* the results of the study across variations of the treatment (i.e., if the treatment were varied a little, would the results be similar?).

8.21. What is outcome validity?

Outcome validity is the degree to which one can *generalize* the results of the study across different but related dependent variables (e.g., if a study showed an effect on self-esteem would it also show an effect on the related construct of self-efficacy?).

8.22. What is construct validity, and how is it achieved?

Construct validity is the degree to which a construct is represented (i.e., measured well) in a research study. Basically, in all research studies we want to have good measurement.

8.23. What is operationalism, and what is its purpose?

Operationalism refers to the process of representing constructs by a specific set of steps or operations. In other words, we want to measure things well and we want to make it clear to our readers exactly how we carried out our measurement (so they can judge for themselves how well our measurement was).

8.24. What is multiple operationalism, and why is it used?

Multiple operationalism refers to the use of two or more measures (rather than just one measure) to represent a construct. The use of multiple measures of a single construct gives you your best chance of fully representing a construct. The *worst* way to measure something is to try to measure it with a single item. For example, you certainly would not want to measure IQ with a single item, right?

8.25. What is meant by research validity in qualitative research?

In qualitative research (just like in quantitative research) we want our research findings to be trustworthy and defensible. That's what we mean by research validity in qualitative research.

8.26. Why is researcher bias a threat to validity, and what strategies are used to reduce this effect?

Researcher bias occurs when a researcher selectively notices only the results that are consistent with what he or she wants or expects to find. The researcher must be very careful to avoid this. One strategy is called reflexivity, which refers to self-reflection by the researcher on his or her biases and predispositions. The point of reflexivity is to see and attempt to minimize the influence of your personal biases. An important strategy for minimizing researcher bias (in addition to reflexivity) is to use negative-case sampling (i.e., to purposively look for and, if present, carefully examine cases that disconfirm your expectations)

8.27. What is the difference between descriptive validity, interpretive validity, and theoretical validity?

- Descriptive validity refers to the factual accuracy of the account as reported by the researcher.
- Interpretive validity means that the qualitative researcher accurately portrays the meanings given by the participants to what is being studied.
- Theoretical validity refers to the degree to which a theoretical explanation developed to explain the data actually fits the data.
- As you can see, one has to do with accurate description (descriptive validity), one has to do with getting and representing the insider's view (interpretive validity), and one has to do with the explanation or theory fitting the data (theoretical validity).

8.28. How is external validity assessed in qualitative research, and why is qualitative research typically weak on this type of validity?

You will recall that external validity refers to the degree to which you can generalize your findings. This is often weak in qualitative research because only a few cases are typically examined in qualitative research. In fact, qualitative researchers are often far less interested in obtaining external validity than in having good in-depth examination of the cases or group and the context in which it is located. (The book points out ways that generalizing still can be done even in these situations.)

I want to add one more study question. "What are the major strategies used in qualitative research to obtain trustworthy and defensible (i.e., valid) findings?"

Here is a list (Table 8.2) of the strategies that should be used in qualitative research. This is a very important list:

■ **TABLE 8.2** Strategies Used to Promote Qualitative Research Validity

Strategy	Description
Researcher-as-detective	A metaphor characterizing the qualitative researcher as he or she searches for evidence about causes and effects. The researcher develops an understanding of the data through careful consideration of potential causes and effects and by systematically eliminating rival explanations or hypotheses until the final case is made beyond a reasonable doubt. The detective can utilize any of the strategies listed here.
Extended fieldwork	To provide for both discovery and validation researchers should collect data in the field over an extended time period.
Low-inference descriptors	The use of description phrased very close to the participants' accounts and researchers' field notes. Verbatims (i.e., direct quotations) are a commonly used type of low-inference descriptors.
Triangulation	Cross-checking information and conclusions through the use of multiple procedures or sources. When the different procedures or sources are in agreement you have corroboration.
Data triangulation	The use of multiple data sources to help understand a phenomenon.
Methods triangulation	The use of multiple research methods to study a phenomenon.
Investigator triangulation	The use of multiple investigators (i.e., multiple researchers) in collecting, analyzing, and interpreting the data.
Theory triangulation	The use of multiple theories and perspectives to help interpret and explain the data.
Participant feedback	The feedback and discussion of the researcher's interpretations and conclusions with the actual participants and other members of the participant community for verification and insight.
Peer review	Discussion of the researcher's interpretations and conclusions with other people. This includes discussion with a disinterested peer, (e.g., with another researcher not directly involved). This peer should be skeptical and play the devil's advocate, challenging the researcher to provide solid evidence for any interpretations or conclusions. Discussion with peers who are familiar with the research can also help provide useful challenges and insights.
External audit	Using outside experts to assess the study quality.
Negative-case sampling	Locating and examining cases that disconfirm the researcher's expectations and tentative explanation.
Reflexivity	Involves self-awareness and critical self-reflection by the researcher on his or her potential biases and predispositions as these may affect the research process and conclusions.
Pattern matching	Predicting a series of results that form a distinctive pattern and then determining the degree to which the actual results fit the predicted pattern or "fingerprint."