

There is some potential for confusion about the use of the word *validity*. In Chapter 3, we introduced the concept of validity as it applies to measurement; the validity of a measurement procedure refers to whether the procedure actually measures the variable that it claims to measure. Here, however, we introduce the concept of validity as it applies to an entire research study. Specifically, we examine the quality of the research process and the accuracy of the results. The same word, *validity*, applies to both contexts. Therefore, we are careful to distinguish between the validity of a research study and the validity of measurement, and you should be careful to separate the two concepts in your own mind.

Any researcher's goal is to be able to summarize a research study by stating, "This is what happened, and this is what it means." Any factor that raises doubts about the limits of research results or about the interpretation of the results is a threat to the validity of the study. The validity of a research study is usually defined in terms of external validity and internal validity.

External Validity

Every research study is a unique event, conducted at a specific time and place with specific participants, instructions, measurement techniques, and procedures. Despite the unique nature of the study itself, researchers usually assume that the obtained results are not unique but can be generalized beyond that study. External validity concerns the extent to which the results obtained in a research study hold true outside that specific study. Can the results of the study be generalized to other populations, other settings, or other measurements? For example, Strack, Martin, and Stepper (1988) conducted a study showing that people rate cartoons as funnier when holding a pen in their teeth (which forced them to smile) than when holding a pen in their lips (which forced them to frown). Although this study was done in 1988 using undergraduate students from the University of Illinois, it seems reasonable to assume that the results are still valid today. That is, if the same study were conducted with today's undergraduate students from a different university, it would be reasonable to expect essentially the same results.

External validity focuses on any unique characteristics of the study that may raise questions about whether the same results would be obtained under different conditions. Any factor that limits the ability to generalize the results from a research study is a **threat to external validity**. For example, the results obtained from a group of 50-year-old males do not necessarily generalize to females or to other age groups. In this case, the limited range of participant characteristics is a threat to the external validity of the study.

DEFINITIONS

External validity refers to the extent to which we can generalize the results of a research study to people, settings, times, measures, and characteristics other than those used in that study.

A **threat to external validity** is any characteristic of a study that limits the ability to generalize the results from a research study.

There are at least three different kinds of generalization, and each can involve threats to external validity.

1. *Generalization from a sample to the general population.* Most research questions concern a large group of individuals known as a population. For example, a researcher may be interested in preschool children or adults with an eating disorder. In each case, the population contains millions of individuals. However, the actual research study is conducted with a relatively small group of individuals known as a sample. For example, a researcher may select a sample of 50 preschool children to participate in a study. One concern for external validity is that the sample is representative of the population so that the results obtained for the sample can be generalized to the entire population. If, for example, a researcher finds that television violence influences the behavior of preschool children in a sample, the researcher would like to conclude that television violence affects the behavior of preschool children in general.
2. *Generalization from one research study to another.* As we noted earlier, each research study is a unique event, conducted at a specific time and place using specific procedures with a specific group of individuals. One concern for external validity is that the results obtained in one specific study will also be obtained in another similar study. For example, if I conduct a study with a specific group of 25 college students, will I obtain the same (or similar) results if I repeat the study 2 years later with a different group of students? If I do my study in New York, will another researcher using the same procedures obtain the same results in California? If I measure IQ scores with the Stanford Binet Intelligence Scales, will another researcher get the same results measuring IQ with the Wechsler Adult Intelligence Scale–IV(WAIS–IV)?
3. *Generalization from a research study to a real-world situation.* Most research is conducted under relatively controlled conditions with individuals who know that they are participating in a research study. One concern for external validity is whether the results obtained in a relatively sterile research environment will also be obtained out in the real world. For example, a researcher may find that a new computer program is very effective for teaching mathematics to third-grade children. However, will the results obtained in the laboratory study also be found in a real third-grade classroom?

Internal Validity

For research studies using the experimental strategy, the goal is to obtain a cause-and-effect explanation for the relationship between two variables, and many other research studies hope to produce some support for a cause-and-effect explanation. For example, consider the following research questions:

- Does increased exercise cause a decrease in cholesterol level?
- Does this particular therapy cause a reduction in depression?
- Does this particular teaching technique cause an improvement in students' academic performance?

In each case, a valid research study would have to demonstrate that changes in one variable (for example, the amount of exercise) are followed by changes

in the other variable (cholesterol level), and that no other variable provides an alternative explanation for the results. This kind of validity is called **internal validity**. Internal validity is concerned with factors in the research study that raise doubts or questions about the interpretation of the results. A research study is said to have internal validity if it allows one and only one explanation of the results. Any factor that allows an alternative explanation for the results is a **threat to internal validity**.

For example, suppose a clinician obtains a group of depressed clients and measures the level of depression for each individual. The clinician then begins therapy with the clients and measures depression again after 3 weeks. If there is a substantial decline in depression, the therapist would like to conclude that the therapy caused a reduction in depression. However, suppose that the weather was cold and miserable when the study began, and changed to bright and sunny when the study ended 3 weeks later. In this case, the weather provides an alternative explanation for the results. Specifically, it is possible that the improved weather caused the reduction in depression. In this example, the weather is a threat to the internal validity of the research study.

DEFINITIONS

A research study has **internal validity** if it produces a single, unambiguous explanation for the relationship between two variables.

A **threat to internal validity** is any factor that allows for an alternative explanation.

Validity and the Quality of a Research Study

The value or quality of any research study is determined by the extent to which the study satisfies the criteria of internal and external validity. The general purpose of a research study is to answer a specific research question. A well-designed study produces results that accurately represent the variables being examined and justify a conclusion that accurately answers the original question. Any factor that generates doubts about the accuracy of the results or raises questions about the interpretation of the results is a threat to validity.

A good researcher is aware of these threats while planning a research study. Anticipating threats to validity allows a researcher to incorporate elements into a research design that eliminate or minimize threats to validity before the research is actually conducted. In this section, we identify and briefly describe some general threats to internal and external validity. In later chapters, we present a variety of different research designs and consider the specific threats to validity associated with each design. In addition, we identify methods of modifying or expanding each design to limit specific threats to validity.

One final caution: It is essentially impossible for a single research study to eliminate all threats to validity. Each researcher must decide which threats are most important for the specific study and then address those threats. Less-important threats can be ignored or treated casually. In fact, design changes that eliminate one threat may actually increase the potential for another threat; thus, each research study represents a set of decisions and compromises about validity. Although researchers typically try to make the best decisions and

produce the best possible studies, most still contain some flaws. This basic “fact of life” has two implications:

1. Research studies vary in terms of validity. Some studies have strong internal and external validity and their results and conclusions are highly respected. Other studies have only moderate validity, and some have little or no validity. Never accept a research result or conclusion as true simply because it is said to have been “scientifically demonstrated.”
2. Being aware of threats to validity can help you critically evaluate a research study. As you read research reports, mentally scan the list of threats and ask yourself whether each one applies. A major learning objective of this book is to make you an informed consumer of research, capable of making your own decisions about its validity and quality.

✓ Learning Check

1. Any factor that raises doubts about the research results or the interpretation of the results is a(n)
 - a. threat to reliability.
 - b. threat to validity.
 - c. artifact.
 - d. demand characteristic.
2. Any factor that limits the ability to generalize the results of the study is a threat to
 - a. internal validity.
 - b. external validity.
 - c. reliability.
 - d. accuracy.
3. Any factor that allows an alternative explanation for the results of a study is a threat to
 - a. internal validity.
 - b. external validity.
 - c. reliability.
 - d. accuracy.

Answers appear at end of chapter.

6.3 | THREATS TO EXTERNAL VALIDITY

■ Learning Objective

LO 5 Identify and explain the common threats to external validity and identify threats when they appear in a research report.

As discussed previously, external validity refers to the extent to which the results of the study can be generalized. That is, will the same (or similar) results be obtained with other populations, conditions, experimenters, other measurements, and so forth? When research findings can be generalized outside the confines of the specific study, the research is said to have external validity. Any characteristic of the study that limits the generality of the results is a threat to external validity. Some of the more common threats to external validity follow, grouped into three major categories.