chapter thirteen

QUASI-EXPERIMENTAL AND SINGLE-CASE EXPERIMENTAL DESIGNS

In the educational world, the environment or situation you find yourself in can be dynamic. You need look no further than within a school classroom. Suppose, for example, that a teacher gives an exam in which the average student scored a 50%. Why were the exam grades so low? Was the teacher ineffective in his or her teaching? Did the students study for the exam? Was the exam itself not fair? Was the material being studied too difficult or at a too high a level? In this example, the answer can be difficult to identify because the classroom environment is constrained by preexisting factors—the time, date, and content area of the exam; the teacher and students in the class were not assigned by a researcher but instead were determined by the parents and school administrators. Accounting for these preexisting factors is important to determine why the exam grades were low.

The above example involves a classwide issue. In other situations, the issue may involve only one or a few students. What if in the above example most students did well on the exam and only a few students were failing the course? There may also be a classroom situation when there is one child with disruptive behavior within a classroom of students. Other educational questions may involve a low-incidence population of students such as those with severe intellectual disability, autism, or speech language impairment. In these cases, we wouldn't need to observe the entire classroom to determine why the grades of the students were low or the why the one student is disruptive, so it would be advantageous to observe the behavior of only the target individuals. For example, we could observe the few students who are failing the course as the teacher implements a different instructional strategy or the student with disruptive behavior as a new positive reinforcement strategy is implemented to see if the behavior changes over time as a result of the new strategies.

In this chapter, we introduce quasi-experimental designs used in science to make observations in group settings that are constrained by preexisting factors. We also introduce methods used to assess the behavior of a single participant using single-case experimental designs, typically used when a large sample is not needed or cannot be obtained.

QUASI-EXPERIMENTAL DESIGNS

Suppose we hypothesize that high school graduates who attend college will value an education more than those who do not attend college. To test this hypothesis, we could select a sample of high school graduates from the same graduating class and divide them into two groups: those who attended college (Group College) and those who did not attend college (Group No College). We could then have all participants complete a survey in which higher scores on the survey indicate a higher value placed on obtaining an education. If the hypothesis is correct and we set up this study correctly, then participants in Group College should show higher scores on the survey than participants in Group No College.

A **quasi-experimental research design** is the use of methods and procedures to make observations in a study that is structured similar to an experiment, but the conditions and experiences of participants lack some control because the study lacks random assignment, includes a preexisting factor (i.e., a variable that is not manipulated), or does not include a comparison/control group.

Notice in this example that participants controlled which group they were assigned to—they either attended college or did not. Hence, in this example, the factor of interest (whether or not students attended college) was a quasi-independent variable. When a researcher does not manipulate a factor in a study (i.e., quasi-independent), this typically means that the study is a type of quasi-experimental research design. In this chapter, we separate the content into two major sections: quasi-experimental designs and single-case experimental designs. We begin this chapter with an introduction to the type of research design illustrated here: the quasi-experimental research design.

13.1 An Overview of Quasi-Experimental Designs

In this major section, we introduce a common type of research design called the quasi-experimental research design. The **quasi-experimental research design,** also defined in

Chapter 6, is structured similar to an experiment, except that this design does one or both of the following:

1. It includes a **quasi-independent variable** (also defined in Chapter 6).
2. It lacks an appropriate or equivalent comparison/control group.

A **quasi-independent variable** is a preexisting variable that is often a characteristic inherent to an individual, which differentiates the groups or conditions being compared in a research study. Because the levels of the variable are preexisting, it is not possible to randomly assign participants to groups.

In the example used to introduce this section, the preexisting factor was college attendance (yes, no). The researchers did not manipulate or randomly assign participants to groups. Instead, participants were assigned to Group College or No College based on whether they attended college prior to the study. In other words, the participants, not the researcher, controlled which group they were assigned to. In this way, the study described to introduce this section was a quasi-experiment—the study was structured like an experiment in that differences in how students value college were compared between groups, but it lacked a manipulation (of the groups: whether students attended or did not attend college) and randomization (of assigning participants to each group).

Hence, a quasi-experiment is not an experiment because, as illustrated in Figure 13.1, the design does not meet all three requirements for demonstrating cause. In the college attendance study, for example, additional unique characteristics of participants, other than whether or not they attended college, could also be different between groups and therefore could also be causing differences between groups. For example, levels of motivation and academic ability may also be different between people who attend and do not attend college. When other possible causes cannot be ruled out, the design does not demonstrate cause.

In this major section, we introduce four categories of quasi-experimental research designs used in the behavioral sciences:

- One-group designs (posttest only and pretest-posttest)
- Nonequivalent control group designs (posttest only and pretest-posttest)
- Time-series designs (basic, interrupted, and control)
- Developmental designs (longitudinal, cross-sectional, and cohort-sequential)

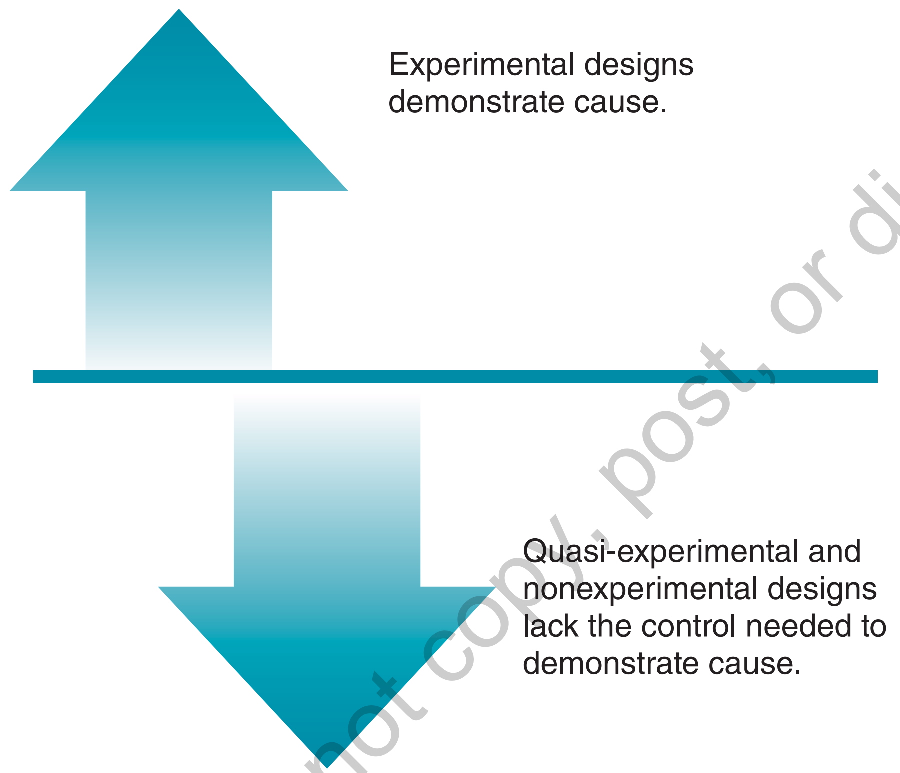
13.2 One-Group Designs

A quasi-experiment resembles an experiment but includes a quasi-independent variable and/or lacks a control group.

In some situations, researchers ask questions that require the observation of a single group. When only one group is observed, the study lacks a comparison group and so does not demonstrate cause. These designs may also be referred to as "preexperimental" designs. Two types of one-group experiments are the following:

- One-group posttest-only design
- One-group pretest-posttest design

Figure 13.1 A Simplified Distinction Between Experiments, Quasi-Experiments, and Nonexperiments



Note: The line represents the requirements for demonstrating cause: randomization, manipulation, and comparison/control. A quasi-experiment lacks at least one of these requirements and so fails to demonstrate cause.

One-Group Posttest-Only Design

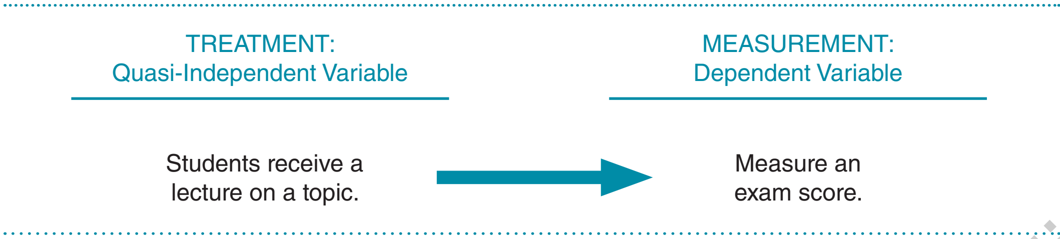
One-group research designs lack a comparison/control group.

The type of quasi-experiment most susceptible to threats to internal validity is the **one-group posttest-only design,** which is also called the *one-shot case study* (Campbell & Stanley, 1966). Using the one-group posttest-only design, a researcher measures a dependent variable for one group of participants following a treatment. For example, as illustrated in Figure 13.2, after a teacher provides instruction on the steps of long division (the treatment), she or he may record the number of division problems solved correctly on a practice worksheet (the dependent variable) to test their learning.

A **one-group posttest-only design** is a quasi-experimental research design in which a dependent variable is measured for one group of participants following a treatment.

The major limitation of this design is that it lacks a comparison or control group. Consider, for example, the number of division problems solved correctly on the practice worksheet. If the number of problems solved correctly is high following the

Figure 13.2 The One-Group Posttest-Only Quasi-Experimental Design



instruction, can we conclude that the instruction is effective? How can we know for sure if number of correct answers would have been high even without the instruction? We cannot know this because we have nothing to compare this outcome to; we have no comparison/control group. Hence, the design is susceptible to many threats to internal validity, such as history effects (unanticipated events that can co-occur with the exam) and maturation effects (natural changes in learning). In all, these limitations make the one-group posttest-only design a poor research design.

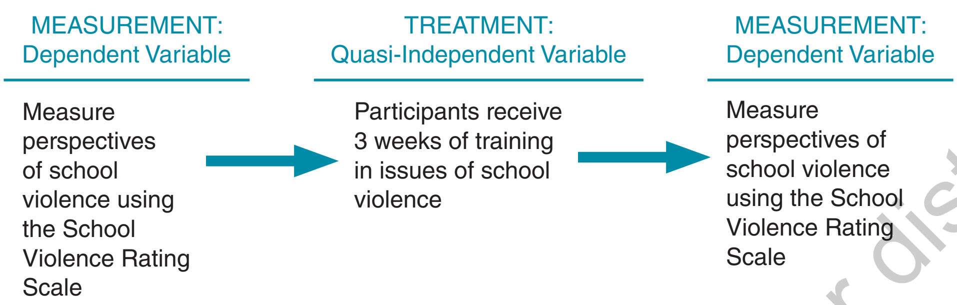
One-Group Pretest-Posttest Design

One way to minimize problems related to having no control or comparison group is to measure the same dependent variable in one group of participants before (pretest) and after (posttest) a treatment. Using this type of research design, called a **one-group pretest-posttest design,** we measure scores before and again following a treatment, then compare the difference between pretest and posttest scores. The advantage is that we can compare scores after a treatment to scores on the same measure in the same participants prior to the treatment. The disadvantage is that the one-group design does not include a no-treatment control group or a business-as-usual comparison group and therefore is still prone to many threats to internal validity, including those associated with observing the same participants over time (e.g., testing effects and regression toward the mean).

A **one-group pretest-posttest design** is a quasi-experimental research design in which the same dependent variable is measured in one group of participants before (pretest) and after (posttest) a treatment is administered.

To illustrate the one-group pretest-posttest design, we will look at the research example illustrated in Figure 13.3. McCaleb, Anderson, and Hueston (2008) measured teacher perceptions of school violence before and after a three-part workshop on school violence. Their results showed a change in perception of school violence from before to after the treatment. A limitation of this design is that participants were not randomly assigned to groups. This means that any other factors related to perception of school violence, such as previous experiences in school violence, teaching experience, or being a victim of a crime outside of school, were beyond the control of the researchers and could have also influenced the results. Also, because the study lacked a control or comparison group with teachers who did not attend the workshop, the design was susceptible to many threats to internal validity, as stated previously.

Figure 13.3 The One-Group Pretest-Posttest Quasi-Experimental Design



Note: Based on a design used by McCaleb, Anderson, and Hueston (2008).

13.3 Quasi-Experimental Design: Nonequivalent Control Group Designs

A **nonequivalent control group** is a control group that is matched upon certain preexisting characteristics similar to those observed in a treatment group but to which participants are not randomly assigned. In a quasi-experiment, a dependent variable measured in a treatment group is compared to that in the nonequivalent control group.

Selection differences are any differences, which are not controlled by the researcher, between individuals who are selected from preexisting groups or groups to which the researcher does not randomly assign participants.

In some cases, researchers can use nonequivalent control/comparison groups when it is not possible to randomly assign participants to groups. A **nonequivalent control group** is a type of control/comparison group that is matched upon certain preexisting characteristics similar to those observed in a treatment group but to which participants are not randomly assigned. For example, suppose a teacher provides instruction using cooperative learning groups for one U.S. history class and provides traditional whole-group instruction in another U.S. history class, then compares grades on the U.S. history topic. The classes are matched on certain characteristics: Both classes are on the same topic (U.S. history), offered at the same school, and taught by the same teacher. However, the class taught using the traditional method is a nonequivalent comparison group because students in that class were not randomly assigned to that class. Any preexisting differences between students in the two classes, called **selection differences,** could therefore explain any differences observed between the two classes. Two types of nonequivalent control group quasi-experiments are the following:

Nonequivalent control group quasi-experimental research designs include a comparison/control group that is nonequivalent.

- Nonequivalent control group posttest-only design
- Nonequivalent control group pretest-posttest design

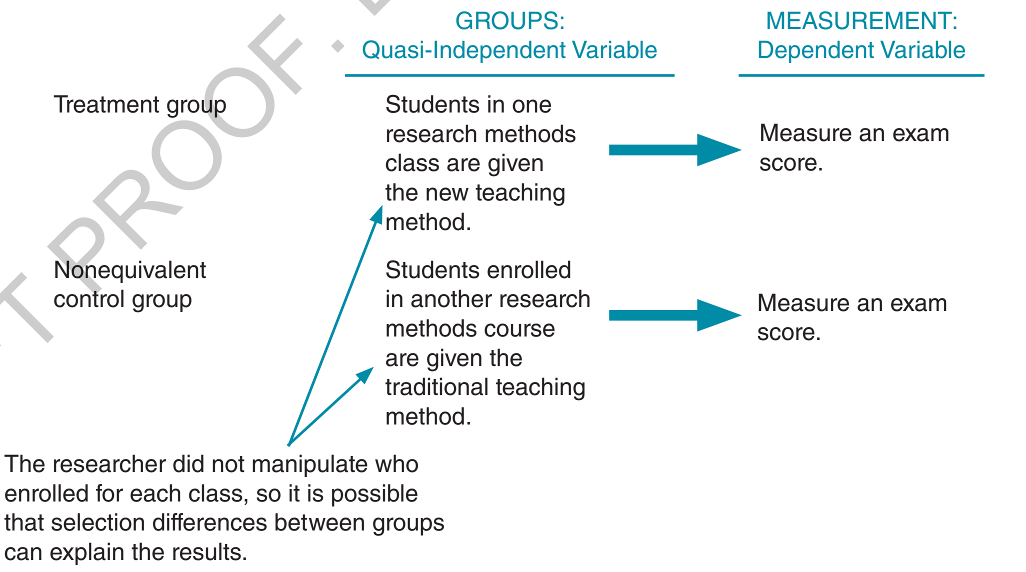
Nonequivalent Control Group Posttest-Only Design

Using the **nonequivalent control group posttest-only design,** a researcher measures a dependent variable following a treatment in one group and compares that measure to a nonequivalent control/comparison group that does not receive the treatment. The nonequivalent control/comparison group will have characteristics similar to the treatment group, but participants will not be randomly assigned to this group, typically because it is not possible to do so. For example, as illustrated in Figure 13.4, suppose a teacher provides a new teaching method in a high school biology class and gives a traditional method in another biology class, then tests all students on the material taught. In this example, the nonequivalent control group was selected because it matched characteristics in the treatment group (e.g., all students were taking a biology class). Students, however, were not randomly assigned to the classes, so the comparison is a nonequivalent control group.

> A **nonequivalent control group posttest-only design** is a quasi-experimental research design in which a dependent variable is measured following a treatment in one group and also in a nonequivalent control group that does not receive the treatment.

A key limitation of this research design is that it is particularly susceptible to the threat of selection differences. In the example illustrated in Figure 13.4, because the high school students registered for the biology class were assigned to a specific class by a school administrator, the researcher did not control which class they enrolled in. Therefore, any preexisting differences between students in the two classes, such as how busy the students' daily schedules are or how much they study, are actually causing differences in grades between classes. For this reason, the nonequivalent control group posttest-only design demonstrates only that a treatment is associated with differences between groups and not that a treatment caused differences between groups, if any were observed.

Figure 13.4 The Nonequivalent Control Group Posttest-Only Quasi-Experimental Design



Nonequivalent Control Group Pretest-Posttest Design

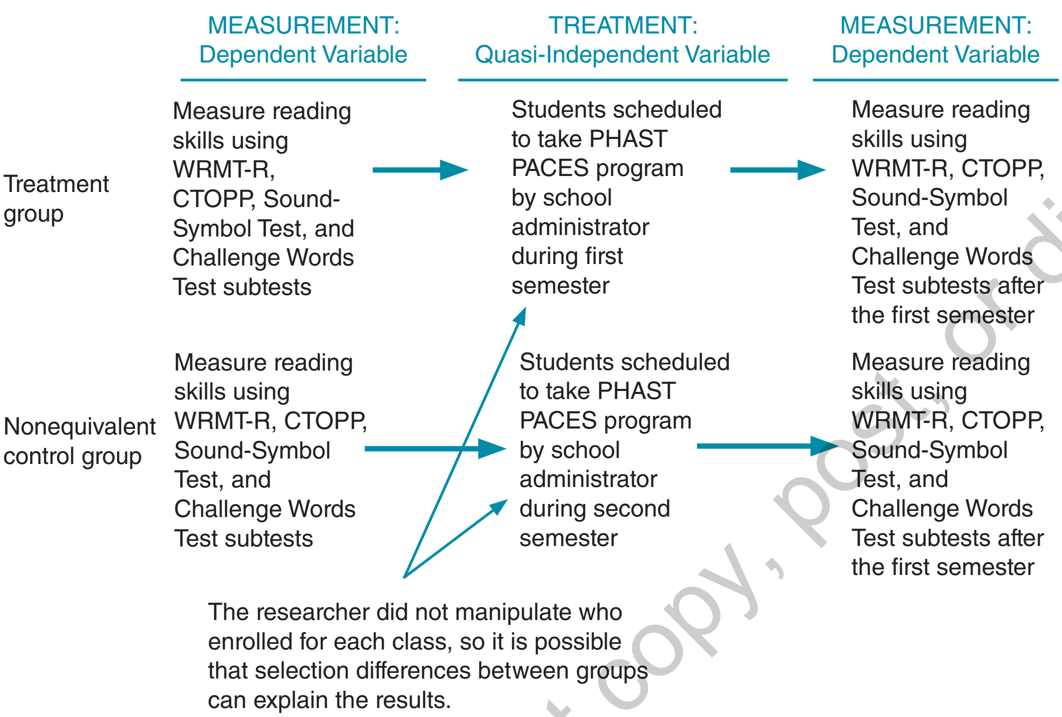
One way to minimize problems related to not having a comparison group is to measure a dependent variable in one group of participants observed before (pretest) and after (posttest) a treatment and also measure that same dependent variable at pretest and posttest in another nonequivalent control group that does not receive the treatment. This type of design is called the **nonequivalent control group pretest-posttest design.** The advantage of this design is that we can compare scores before and after a treatment in a group that receives the treatment and also in a nonequivalent control group that does not receive the treatment. While the nonequivalent control group will have characteristics similar to the treatment group, participants are not randomly assigned to this group, typically because it is not possible to do so. Hence, selection differences still can possibly explain observations made using this research design.

A **nonequivalent control group pretest-posttest design** is a quasi-experimental research design in which a dependent variable is measured in one group of participants before (pretest) and after (posttest) a treatment and that same dependent variable is also measured at pretest and posttest in another nonequivalent control group that does not receive the treatment.

To illustrate the nonequivalent control group pretest-posttest design, we will look at the research example in Figure 13.5. Lovett, Lacerenza, DePalma, and Frijters (2012) used several measures of reading (i.e., word attack, word reading, and passage comprehension) to measure the reading skills of high school students who were identified as struggling readers. These researchers hypothesized that a reading intervention called PHAST PACES that teaches word identification strategies, knowledge of text structures, and reading comprehension strategies would increase reading skills among the struggling readers. To test this hypothesis, struggling readers were assigned by school administrators to the PHAST PACES course in the first semester (the treatment group) or second semester (the nonequivalent control group) of high school. All of the struggling readers were given the battery of reading tests before and after the implementation of PHAST PACES in the first semester. As shown in Figure 13.6, students who took the PHAST PACES course in the first semester (the treatment group) showed a larger change or increase in reading skills compared with students in the nonequivalent control group who did not take the PHAST PACES course in the first semester.

A key limitation of this research design is that it is particularly susceptible to the threat of selection differences. In the example illustrated in Figure 13.5, because students were assigned to the first semester or second semester course by school administrators, the school administrators, and not the researcher, controlled what semester course they were in. Any preexisting differences between students could also be causing differences classes in the reading skills. For example, while students were identified as being struggling readers on a battery of pretests, the level and type of reading difficulties varied among the students. Some students struggled more than others in one or more areas of reading. Because the school administrators decided which students took the PHAST PACES course in the first or second semester, it is possible that level and type of reading difficulties also varied between the groups and therefore could be the cause or reason for the differences observed. Hence, the nonequivalent control group pretest-posttest design, like the posttest-only design,

Figure 13.5 The Nonequivalent Control Group Pretest-Posttest Quasi-Experimental Design



Source: Adapted from Lovett et al. (2012).

demonstrates only that a treatment is associated with differences between groups and not that a treatment caused differences between groups, if any were observed.

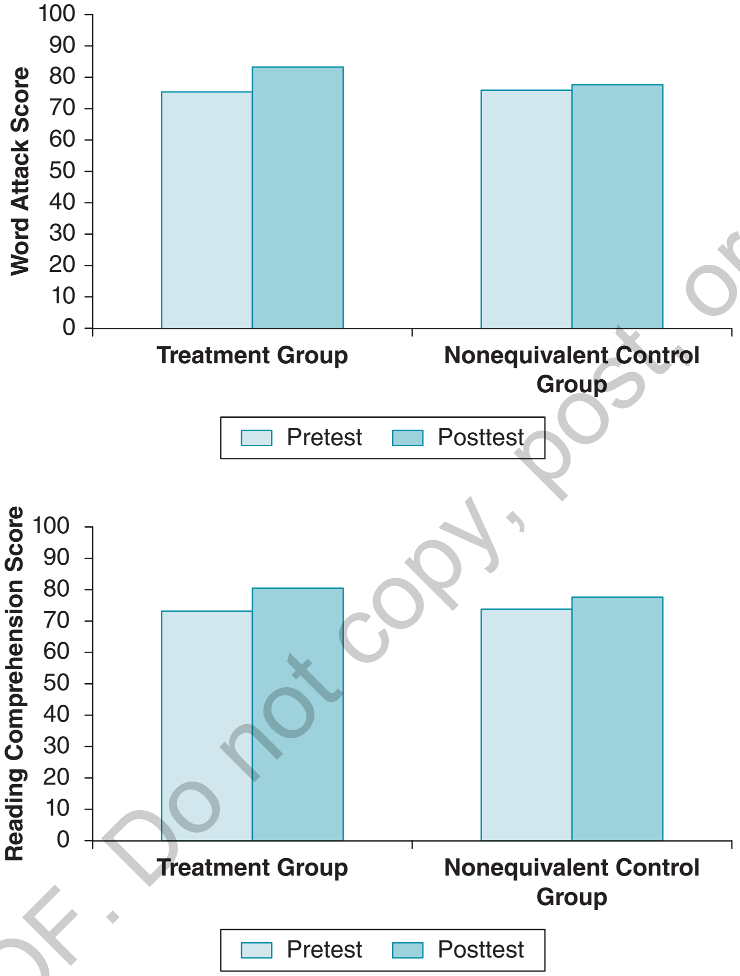
13.4 Quasi-Experimental Design: Time-Series Designs

In some situations, researchers observe one or two preexisting groups at many points in time before and after a treatment, and not just at one time, using designs called the time-series quasi-experimental designs. Using these types of designs, we compare the pattern of change over time from before to following a treatment. Three types of time-series quasi-experimental designs are as follows:

Time-series quasi-experimental research designs involve many observations made before and after a treatment.

- Basic time-series design
- Interrupted time-series design
- Control time-series design

Figure 13.6 The Word Attack and Passage Comprehension Scores for Each Group, Where Higher Scores Indicate Greater Reading Skill



Data are adapted from those reported by Lovett, Lacerenza, DePalma, and Frijters (2012). Reproduced with permission.

Basic Time-Series Design

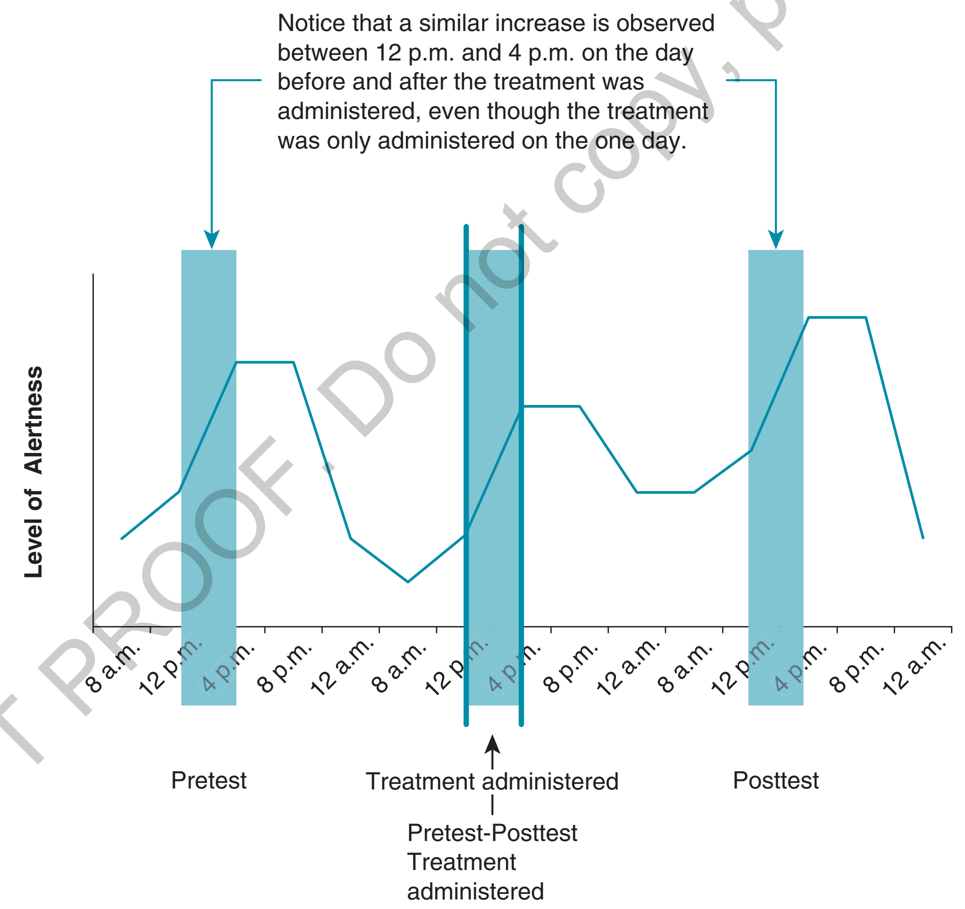
When researchers manipulate the treatment, they use a **basic time-series design** to make a series of observations over time before and after a treatment. The advantage of measuring a dependent variable at multiple times before and after a treatment is that it eliminates the problem associated with only having a snapshot of behavior. To illustrate, suppose we test a treatment for improving alertness during the day. To use the basic time-series design, we record alertness at

A **basic time-series design** is a quasi-experimental research design in which a dependent variable is measured at many different points in time in one group before and after a treatment that is manipulated by the researcher is administered.

multiple times before and after we give participants the treatment, as illustrated in Figure 13.7. Notice in the figure that a pretest (at 12 p.m.) and posttest (at 4 p.m.) measure can be misleading because the pattern observed before and after the treatment recurred without the treatment at the same time the day before and the day after the treatment was given. The basic time-series design allows us to uniquely see this pattern by making a series of observations over time.

Using the basic time-series design, the researcher manipulates or controls when the treatment will occur. The advantage of this design is that we can identify if the pattern of change in a dependent variable before and after the treatment occurs only during that period of time and not during other periods of time when the treatment is not administered. The disadvantage of this design is that only one group is observed, so we cannot compare the results in the treatment group to a group that never received the treatment.

Figure 13.7 The Time-Series Quasi-Experimental Design



Note: A time-series design is used to compare the pattern of behavior before and after the treatment. In this example, the pattern that occurs before and after the treatment recurs at the same time of day, even without the treatment.

Interrupted and Control Time-Series Designs

A basic time-series design involves a manipulated treatment, whereas the interrupted time-series design involves a naturally occurring treatment.

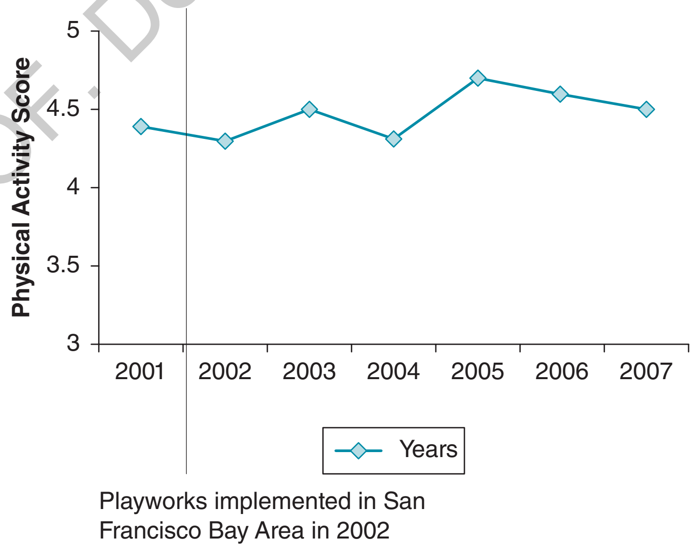
In some situations, educational researchers will measure a dependent variable multiple times before and after a naturally occurring treatment or event. Examples of a naturally occurring treatment or event in education include changes in educational policy such as class size and curriculum adoptions. These events occur beyond the control of the researcher, so the researcher loses control over the timing of the manipulation. In these situations, when multiple measurements are taken before and after a naturally occurring treatment, researchers use the **interrupted time-series design.**

As an example of the interrupted time-series design, Madsen, Hicks, and Thompson (2011) measured physical activity reports from the California Healthy Kids Survey 1 year before and 6 years after implementing a Playworks. For this study, the line in Figure 13.8 shows that the days per week of physical activity (exercise, dance, or play sports) increased in the San Francisco Bay Area, 6 years following the implementation of the Playworks curriculum.

An advantage of the interrupted time-series design is that we can identify if the pattern of change in a dependent variable changes from before to following a naturally occurring treatment or

An **interrupted time-series design** is a quasi-experimental research design in which a dependent variable is measured at many different points in time in one group before and after a treatment that naturally occurred.

Figure 13.8 Interrupted Time-Series Design: Physical Activity Reports 1 Year Before and 6 Years After Implementing Playworks



Note: Data are adapted from those reported by Madsen, Hicks, and Thompson (2011). Reproduced with permission by Wiley.

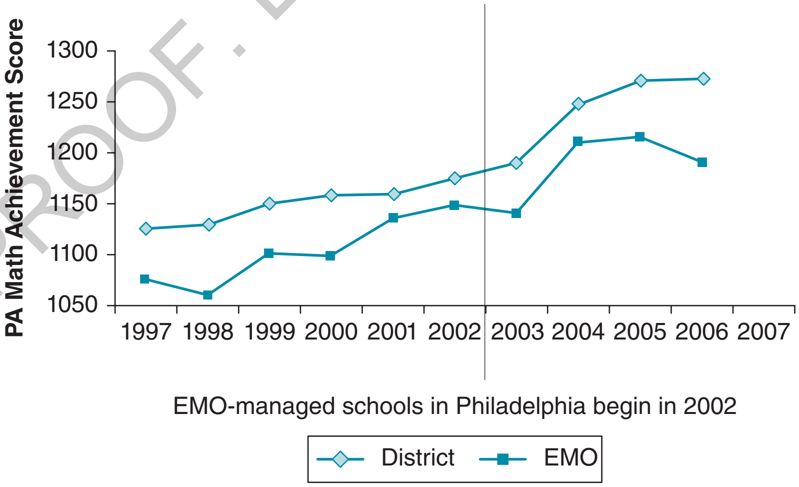
event. The disadvantage of this design, like that for the basic time-series design, is that only one group is observed, so we cannot compare the results in the treatment group to a group that never received a treatment. To address this disadvantage, we can include a matched or nonequivalent control group.

A basic or interrupted time-series design that includes a matched or nonequivalent control group is called a **control time-series design.** Byrnes (2009) examined the achievement scores of middle grade schools in Pennsylvania that were privatized to an education management organization 6 years before and 5 years after the privatization. He included a control group of schools from Pennsylvania that were not privatized by also recording achievement scores during the same period of time. As shown in Figure 13.9, achievement scores in the nonprivatized schools were greater than those of the privatized schools. The addition of this control group can increase how confident we are in the effect of privatizing schools with education management organizations.

A **control time-series design** is a basic or interrupted time-series quasi-experimental research design that also includes a nonequivalent control group that is observed during the same period of time as a treatment group but does not receive the treatment.

As a caution, keep in mind that the students in each of the schools are preexisting groups in that the researcher did not assign students to the schools or which schools would be privatized. It is therefore possible, like for all other designs that use a nonequivalent control group, that selection differences, such as differences in attendance rates or student demographics (e.g., free and reduced-price lunch eligibility or percentage of minority students) between students in each of the schools, could have caused the different observed pattern of achievement scores and not the privatization of

Figure 13.9 Control Time-Series Design: Pennsylvania Achievement Scores 6 Years Before and 5 Years After Privatizing Schools to an Education Management Organization



Note: Data are adapted from Byrnes, V. (2009). Getting a feel for the market: The use of privatized school management in Philadelphia. *American Journal of Education, 115*, 437–455. Reproduced with permission by the University of Chicago Press.

Table 13.1 The Quasi-Experimental Research Designs

Type of Quasi-Experimental Design	Description	Key Limitation
One-group posttest only	Observe one group after (posttest) a treatment	No control group for comparison
One-group pretest-posttest	Observe one group before (pretest) and after (posttest) a treatment	No control group for comparison
Nonequivalent control group posttest only	Observe treatment and nonequivalent control groups after (posttest) a treatment	No random assignment between groups
Nonequivalent control group pretest-posttest	Observe treatment and nonequivalent control groups before (pretest) and after (posttest) a treatment	No random assignment between groups
Basic time-series design	Make many observations over a period of time before and after a treatment manipulated by the researcher	No control group for comparison
Interrupted time-series design	Make many observations over a period of time before and after a naturally occurring treatment	No control group for comparison
Control series design	A time-series design with a matched or nonequivalent control group	No random assignment between groups

school management (the treatment). For this reason, we conclude that the privatization was associated with reduced achievement and not that the privatization caused the difference in achievement scores.

Table 13.1 summarizes each quasi-experimental research design described in this chapter. In the next section, we will introduce a special case of quasi-experiments used in developmental research.

LEARNING CHECK 1 ✓

1. The quasi-experimental research design is structured similar to an experiment, except [complete the sentence].
2. State the type of quasi-experimental research design described in each of the following examples:
   A. A researcher records the time (in seconds) it takes a group of students to complete a computer-based task following an online "how-to" course.

B. A researcher records the rate of school attendance at a school for 2 years before and 2 years after a reduced school attendance reward program was implemented.

C. A researcher records teacher satisfaction for 3 months before and 3 months after a training seminar. He compares satisfaction scores for teachers at one school compared to the satisfaction scores for teachers at another school who did not receive the seminar.

Answers: 1. the research design includes a quasi-independent variable and/or lacks an appropriate or equivalent control group; 2. A. one-group posttest-only design, B. interrupted time-series design, C. nonequivalent control group pretest-posttest design.

MAKING SENSE—IDENTIFYING QUASI-EXPERIMENTAL DESIGNS

While reading a journal article, you might find a thorough description of the steps of the study, but the specific type of design is not named. Diagramming the steps of the study can help identify the design. We use a notation system to diagram studies. The system uses these notations:

X = exposure to the independent variable

O = observation (or data collection) of the dependent variable(s)

A, B, C, etc. = groups of participants

Diagrams for the quasi-experimental designs are as follows:

One-group posttest only	X	O						
One-group pretest-posttest	O	X	O					
Nonequivalent control group posttest only	A	X	O					
	B	X	O					
Nonequivalent control group pretest posttest	A	O	X	O				
	B	O	X	O				
Basic and interrupted time series	O	O	O	X	O	O	O	
Control time series	A	O	O	X	O	O		
	B	O	O	X	O	O		

Diagramming the number of groups, as well as timing the measurement of the independent and dependent variable(s), can help you identify the type of design if it is not specifically named in the description of the study.

CONNECTING TO THE CLASSROOM

Much of the research in education involves the identification of effective educational programs, policies, and practices. An effective educational program, policy, or practice either improves student outcomes such as achievement, engagement, or motivation or decreases behaviors that interfere with educational attainment such as disruptive behavior at school or truancy. This research, which involves testing a potentially effective program, policy, or practice and evaluating its effect on students, must follow rigorous research methods. A panel of educational research experts convened by the U.S. Department of Education, Institute for Education Sciences created the *What Works Clearinghouse Procedures and Standards Handbook* (U.S. Department of Education, 2010) that outlines the rigorous methods that must be followed to be considered a high-quality study.

Most of the research designed to identify effective educational programs, policies, and practices are group designs. The nonequivalent group quasi-experimental designs discussed in this chapter lack one of the most important elements of a high-quality study that can be used to identify effective practices—random assignment to groups. Nonequivalent group designs rely on intact groups such as a group of students within classrooms where random assignment to the treatment and control/comparison groups may not be possible. One way educational researchers can overcome this problem of lack of random assignment to groups is to provide some evidence of baseline (preintervention) equivalence of the groups. To provide this evidence, researchers examine the differences in important characteristics of the groups that may affect the outcomes, such as gender, ethnicity, or performance on the dependent measure using statistics. A second way to provide equivalent groups is to make statistical adjustments to account for the any preexisting group differences when conducting the statistical analysis of the results of the study.

When reading research in consideration for implementation as an effective practice to improve student educational outcomes, if the study uses a quasi-experimental design, look for information that either (a) demonstrates that the groups are equivalent before the study is conducted or (b) accounts for group differences during analysis once the study is completed.

SINGLE-CASE EXPERIMENTAL DESIGNS

In this section, we begin by identifying a new research design to test the following research hypothesis: Giving positive reinforcement to a student who is disruptive in class only while he or she stays on task will increase on-task behavior in the classroom. To answer this hypothesis, we could measure the time (in minutes) that the disruptive student stays on task. We could observe the student for a few days with no positive reinforcement. Then we could observe the student for a few days with positive reinforcement given as he or she works on the task. Then we could again observe the student for a few more days with no positive reinforcement. If the hypothesis is correct and we set up this study correctly, then we should expect to find that the time (in minutes) spent on task was high when the positive reinforcement was given but low during the observation periods before and after when no positive reinforcement was given. The unique feature of this design is that only one participant was observed.

In this final section, we introduce the research design that was illustrated here: the single-case experimental design.

13.5 An Overview of Single-Case Designs

In some cases, educational researchers want to observe and analyze the behavior of an individual case. An individual case can be a single participant or a single cluster of participants such as a classroom. We can study individual cases using a research design called the **single-case experimental design.** A single-case design is unique in that the individual case serves as its own control compared to at least two conditions or phases, without and with an intervention (Kazdin, 2011). In addition, the dependent variable is repeatedly measured in a single-case design across conditions for each individual case and is not averaged to compare groups. Single-case designs are useful in education to evaluate the effectiveness of interventions when applied to individual cases, as illustrated above, rather than groups. Although it can be applied to individual clusters, single case is most often applied with individual students in educational settings. In contrast, all other experimental research designs, introduced in Chapters 10 to 12, are grouped designs.

A **single-case experimental design** is an experimental research design in which an individual case serves as his, her, or its own control, and the dependent variable measured is analyzed for each individual case.

Single-case designs have three main characteristics. These characteristics include an individual case, manipulation, and repeated measurements.

The single-case design, which is also called the *single-subject, single-participant*, or *small n design*, is most often used in applied areas of special education.

1. As described above, the single-case design involves the analysis of an individual case whereby the individual case is compared to itself. However, to establish experimental control, the control must be repeated at least three times (Horner, Swaminathan, Sugai, & Smolkowski, 2012). These demonstrations of control can be across three cases (such as across individual people or classrooms), settings (such as across different times of day or different places), or materials (such as reading from different books or academic content areas).
2. *Manipulation* involves control over the absence and delivery of the independent variable across the phases or treatments that are experienced by each case. The researcher must control when the independent variable or treatment is delivered. If there is more than one independent variable or treatments, then the researcher is in control over when each is delivered across the different phases of the study.

The single-case design is characterized by an individual case that serves as its own control with repeated measurements across phases of the study.

3. Repeated measurement involves the frequent measurement of the dependent variable across all phases of the study. The dependent variable is measured at least three times in each phase of the study.

An advantage of analyzing the data one participant at a time is that it allows for the critical analysis of each individual measure, whereas averaging scores across groups can give a spurious appearance of orderly change. To illustrate this advantage, suppose that a researcher measures the effect of an academic intervention for incarcerated adolescents. The hypothetical data, shown in Table 13.2, show that the adolescents as a group gained on average 25 points on the assessment. However, Student C scored the same. An analysis of each individual student could be used to explain this outlier; a grouped design would often disregard this outlier as "error" so long as weight loss was large enough on average.

The single-case designs described in this chapter include reversal design (AB design), multiple-baseline design, changing-criterion design, and alternating treatment design (ABC design).

13.6 Single-Case Baseline-Phase Designs

Single-case designs are typically structured by alternating baseline and treatment phases over many trials or observations. In this major section, we will introduce three types of single-case experimental research designs:

Table 13.2 The Value of an Individual Analysis

Participant	Baseline assessment score	Assessment score following academic intervention	Assessment Increase
Adolescent A	70	85	15
Adolescent B	50	80	30
Adolescent C	66	66	0
Adolescent D	58	83	25

Average assessment gain: 17.5 points

Adolescent C was the only participant to not gain points on the assessment. An individual analysis would investigate why, whereas a group analysis would mostly disregard this anomaly as long as average assessment increase was large enough.

Note: In this example, an individual analysis could be used to explain why Student C was the only student not to improve his or her test score.

- Reversal design
- Multiple-baseline design
- Changing-criterion design

Reversal Design

One type of single-case design, called the **reversal design,** involves observing a single case prior to (A), during (B), and following (A) a treatment or manipulation. The reversal design is structured into **phases,** represented alphabetically with an A or a B. Each phase consists of many observations or trials. The researcher begins with a **baseline phase (A),** in which no treatment is given, then applies a treatment in a second phase (B), and again returns to a baseline phase (A) in which the treatment is removed. This type of research design can be represented as follows:

A **reversal design,** or **ABA design,** is a single-case experimental design in which a single participant is observed before (A), during (B), and after (A) a treatment or manipulation.

A **phase** is a series of trials or observations made in one condition.

The **baseline phase (A)** is a phase in which a treatment or manipulation is absent.

A (baseline phase) → B (treatment phase) → A (baseline phase)

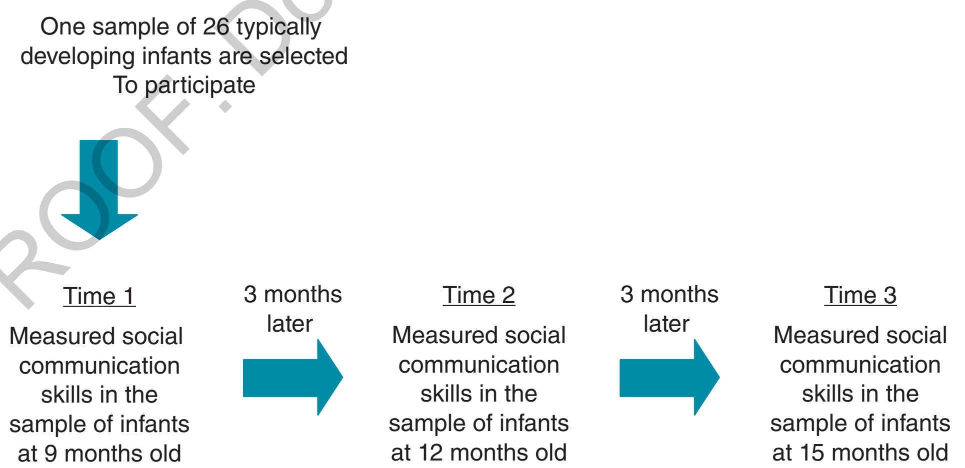
If the treatment in Phase B causes a change in the dependent variable, then the dependent variable should change from baseline to treatment, then return to baseline levels when the treatment is removed. For example, we opened this section with the hypothesis that giving positive reinforcement to a disruptive student while he or she is on task will increase the amount of on-task behavior in the classroom. To test this hypothesis, we measured the time in minutes that the disruptive student spent on task in a class with no positive reinforcement (baseline, A) for a few trials, then with positive reinforcement (treatment, B) for a few trials, and again with no positive reinforcement (baseline, B) for a few more trials. If the positive reinforcement (the treatment) was successful, then the time (in minutes) spent on task would be higher when the positive reinforcement was given but lower during the observation periods before and after when no positive reinforcement was given. The second baseline phase minimizes the possibility of threats to internal validity. Adding another B and A phase would further minimize the possibility of threats to internal validity because the pattern of change would be repeated using multiple treatment phases.

A visual inspection of the data, and not inferential statistics, is used to analyze the data when only a single participant is observed. When visually inspecting data in single-case studies, we are looking to identify a functional relationship between the independent and dependent variables. To analyze the data in this way, we look for three types of patterns that indicate that a treatment caused an observed change within a phase and three types of patterns across phases (*What Works Clearinghouse Procedures and Standards Handbook;* U.S. Department of Education, 2010). Figure 13.10 illustrates the three types of patterns to look for within a phase:

- A *change in level* is displayed graphically, as shown in Figure 13.10 (Graph A), when the level of the dependent variable within the baseline phases is obviously less than or greater than the level of the dependent variable within the treatment phase.
- A *change in trend* is displayed graphically, as shown in Figure 13.10 (Graph B), when the direction or pattern of change within the baseline phases is different from the pattern of change within the treatment phase. In the typical case, a dependent variable gradually increases or decreases in the treatment phase but is stable or does not change in the baseline phases.
- A change in *variability* is displayed graphically, as shown in Figure 13.10 (Graph C), when the pattern of data points within the baseline phases is different from the pattern of data points within the treatment phase. Typically, the pattern of data points in the treatment phase will be less variable than the pattern of the data points in the baseline phase.

Overlap (Graph D), immediacy of the effect (Graph E), and the consistency of data in similar phases (Graph F) make it possible to infer that some treatment is causing an effect or a change in behavior.

Figure 13.10 Three Ways to Identify If a Treatment Caused Changes in a Dependent Variable Within the Phase



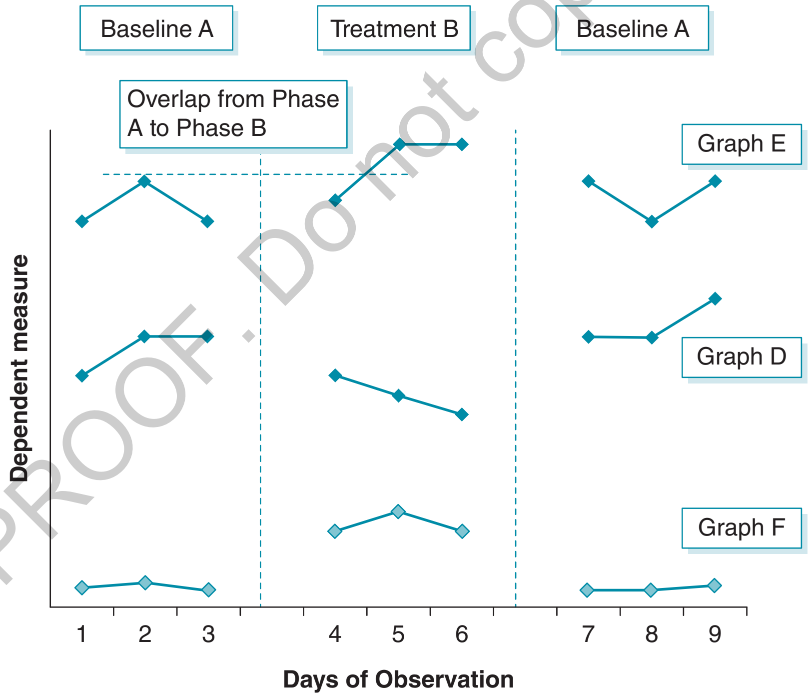
Source: Adapted from Wu & Chiang (2014).

Note: A change in level within a phase (Graph A), a change in trend (Graph B), and the variability (Graph C) make it possible to infer that some treatment is causing an effect or a change in behavior.

Visual inspection of single-case data also includes looking at the pattern of data across phases. We look for three additional patterns of the data across phases, as illustrated in Figure 13.11.

- *Overlap* is displayed in Figure 13.11 (Graph D) when the pattern of change in one phase overlaps the data in the other phase. The smaller the amount of overlap across baseline and treatment phases is also indicative of a functional relationship between the independent and dependent variables.
- *Immediacy of the effect* is displayed in Figure 13.11 (Graph E) when the direction or pattern of change in the phases occurs immediately following the implementation or withdrawal of the treatment in the different phases. The more immediate the change is from one phase to another, the more compelling the argument is for a functional relationship between the independent and dependent variables.
- *Consistency of data in similar phases* is displayed in Figure 13.11 (Graph F) when the pattern of change in similar phases (all baseline A or all treatment phases B)

Figure 13.11 Three Ways to Identify If a Treatment Caused Changes in a Dependent Variable Across Phases



The overlap across levels (graph D), immediacy of the effect across levels (graph E), and consistency across levels (graph F) also make it possible to infer that some treatment is causing an effect or a change in behavior.

Note: Based on a design used by Ridgers, Saint-Maurice, Welk, Siapush, and Huberty (2014). Notice that participants are grouped based on their weight and gender using the cross-sectional design.

is consistent with each other. The greater the consistency of the pattern of data across similar phases, the greater the likelihood of a functional relationship.

The reversal design is typically conducted in educational research to investigate the effectiveness of interventions that may benefit the individual participant. Often this means that researchers will be asked by ethics committees to end their study with a treatment phase (B), which was the phase that was beneficial to the participant. For this reason, many reversal designs are at least four phases, or ABAB, so as not to return to baseline to end an experiment.

A visual inspection of the data, and not inferential statistics, is used to analyze the data when only a single participant is observed.

A limitation of the reversal design is that the change in a dependent variable in a treatment phase must return to baseline levels when the treatment is removed. However, when applied to interventions that affect learning, a return to baseline is not possible. When a participant is taught a new skill, for example, it is often not possible to undo what the participant learned—as fully expected, the behavior will not return to baseline. There may also be times when returning to baseline would be detrimental to the participant such as reducing harmful behavior like self-injury. In these situations, when it is not possible for changes in a dependent variable to return to baseline, a reversal design cannot be used.

A **multiple-baseline design** is a single-case experimental design in which a treatment is successively administered over time to different participants, for educational materials, or in different settings.

Multiple-Baseline Designs

For situations in which it is not possible for changes in a dependent variable to return to baseline levels following a treatment phase, researchers can use the **multiple-baseline design.** The multiple-baseline design is a single-case design in which the treatment is successively administered over time to different participants, for different educational materials, or in different settings. This design allows researchers to systematically observe changes caused by a treatment without the need of a second baseline phase and can be represented as follows:

Case #1	Baseline	Treatment		
Case #2	Baseline		Treatment	
Case #3	Baseline			Treatment

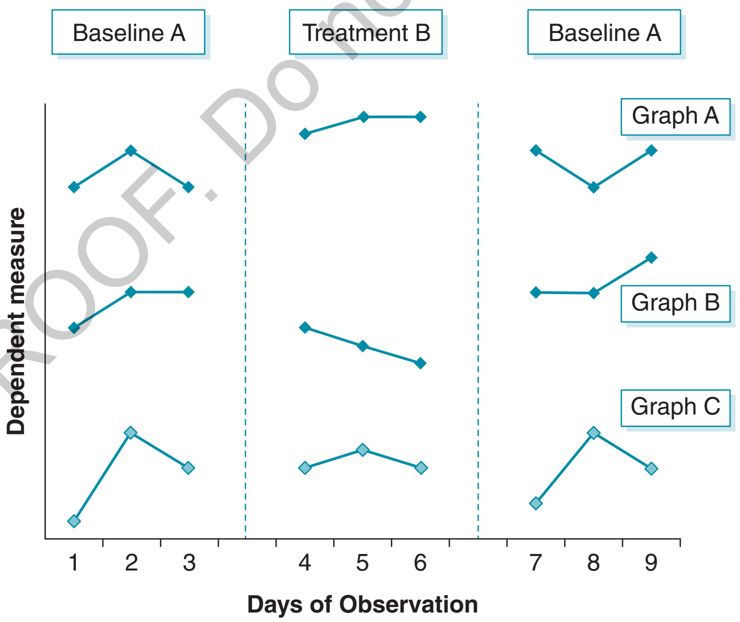
By representing the multiple-baseline design in this way, a *case* refers to a unique time, behavior, participant, or setting. Treatment phases staggered to the individual cases illustrate control over the changes in the data. While the treatment phase is implemented for Case 1, Cases 2 and 3 remain in baseline. Once Case 1 demonstrates a change in level of pattern, then Case 2 will enter treatment while Case 3 remains in baseline. Case 3 will enter the treatment phase last. If the treatment causes an effect following a baseline phase for each case, then the change in level or pattern should begin only when the baseline phase ends, which is different for each case. If this occurs, then we can be confident that the treatment is causing the observed change. This design minimizes the likelihood that something other than the treatment is causing the observed changes if the changes in a dependent variable begin only after the baseline phase ends for each case.

To illustrate the multiple-baseline design, we will look at the research example illustrated in Figure 13.12. Dukes and McGuire (2009) used a multiple-baseline design to measure the effectiveness of a sex education intervention, which they administered to multiple participants with a moderate intellectual disability. The researchers recorded participant knowledge of sexual functioning using the Sexual Consent and Education Assessment (SCEA K-Scale; Kennedy, 1993), on which higher scores indicate greater ability to make decisions about sex. Each participant was given a baseline phase for a different number of weeks. Scores on the SCEA K-Scale were low in this baseline phase. As shown in Figure 13.12 for three participants, only after the baseline period ended and the intervention was administered did scores on the scale increase. Scores also remained high for 4 weeks after the program ended. Hence, the results showed a change in level from baseline to intervention for each participant.

The start of a treatment phase varies using the multiple-baseline design to determine if the changes in a dependent variable begin only after each baseline phase ends.

Each participant in the sex education study received the intervention (or the treatment) in successive weeks: Tina (Week 11), Josh (Week 12), and Debbie (Week 13). Because the treatment was administered at different times, and changes in the dependent variable only occurred once the treatment was administered, the pattern showed that the treatment, and not other factors related to observing participants over time, caused the observed changes in SCEA K-Scale scores.

Figure 13.12 Results From a Multiple-Baseline Design for Three Participants Receiving a Sex Education Intervention



A change in level (graph A), a change in trend (graph B), and a change in variability (graph C) make it possible to infer that some treatment is causing an effect or a change in behavior.

Note: Adapted from Dukes & McGuire (2009). Reproduced with permission by Wiley.

The advantage of a multiple-baseline design is that it can be used when we expect a treatment will not return to baseline, such as when we study learning on some measure, as illustrated in Figure 13.12 for our example. One limitation to the multiple-baseline design in an educational context is the collection of data at each instructional session. For some studies that have multiple tiers of the intervention, such as across participants and settings, this can result in a lot of data collection. To make this design more efficient, the multiple-probe design is often used. Instead of collecting data at each session, data are collected at prescribed times (called probes) and in sufficient number to still provide visual evidence of the pattern of data in each of the phases. A second limitation of a multiple-baseline design is that the design is used when only a single type of treatment is administered. This same limitation applies to the reversal design. For situations when we want to administer successive treatments, then, we require a different type of single-case experimental design.

A **changing-criterion design** is a single-case experimental design in which a baseline phase is followed by successive treatment phases in which some criterion or target level of behavior is changed from one treatment phase to the next. The participant must meet the criterion of one treatment phase before the next treatment phase is administered.

Changing-Criterion Designs

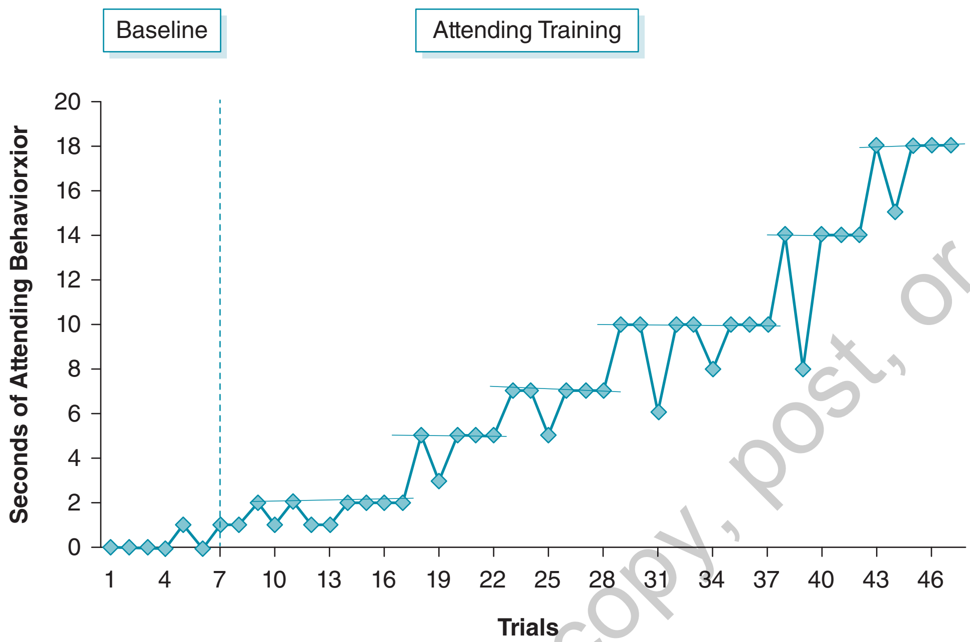
For research situations in which we want to change a criterion or treatment after the participant meets an initial criterion or responds to one particular treatment, we can use a **changing-criterion design.** Using the changing-criterion design, we begin with a baseline phase, which is followed by many successive treatment phases to determine if participants can reach different levels or criteria in each treatment phase. The criterion can be changed as often as necessary or until some final criterion is met.

A Baseline → B1 Treatment 1 → B2 Treatment 2 → B3 Treatment 3

To illustrate the changing-criterion design, we will look at the research example illustrated in Figure 13.13. Plavnick (2012) used the changing-criterion design to increase the number of seconds that a student with autism attended to a video displayed on a cellphone of same-age peers modeling how to communicate using words and picture cue cards. In a baseline phase, the student looked at the video once for 1 second out of six trials. Then a series of manipulations followed. Each time the student attended to the screen of the cellphone while the video was played, he was rewarded with a preferred edible. The initial criterion was 2 seconds. This criterion was increased over time between 2 and 4 seconds after three consecutive sessions of meeting the previous criterion. As shown in Figure 13.13, each time the criterion, or the number of seconds required to gain a reward, was increased, the student's attending behavior increased.

Two advantages of the changing-criterion design are that it does not require a reversal to baseline of an otherwise effective treatment and that it enables experimental analysis of a gradually improving behavior. A limitation of the design is that the target

Figure 13.13 A Changing-Criterion Design to Increase the Number of Seconds a Single Student With Autism Attended to a Video



Note: At baseline, Sam attended to the video once for 1 second. He began with 2 seconds, then 5 seconds, 7 seconds, 10 seconds, 14 seconds, and finally 18 seconds in order to receive the reward. The changing criterion is highlighted in the treatment phase by the dotted lines. Notice that as the criterion was increased, the student increased the number of seconds he looked at the video displayed. Data based on those presented by Plavnick (2012).

behavior must already be in the participant's repertoire. For example, the student with autism needed to look at the video on the cellphone at least once to be able to increase the amount of time spent attending to the video. Also, researchers should be cautious to not increase or decrease the criterion too soon or by too much, which may impede the natural learning rate of the participant being observed.

With the reversal, multiple-baseline, and changing-criterion designs, notice that only one treatment is being evaluated. In some cases, researchers want to compare treatments to compare the effectiveness of each. For this type of research, we use the alternating treatment or ABC design.

Alternating Treatment Design

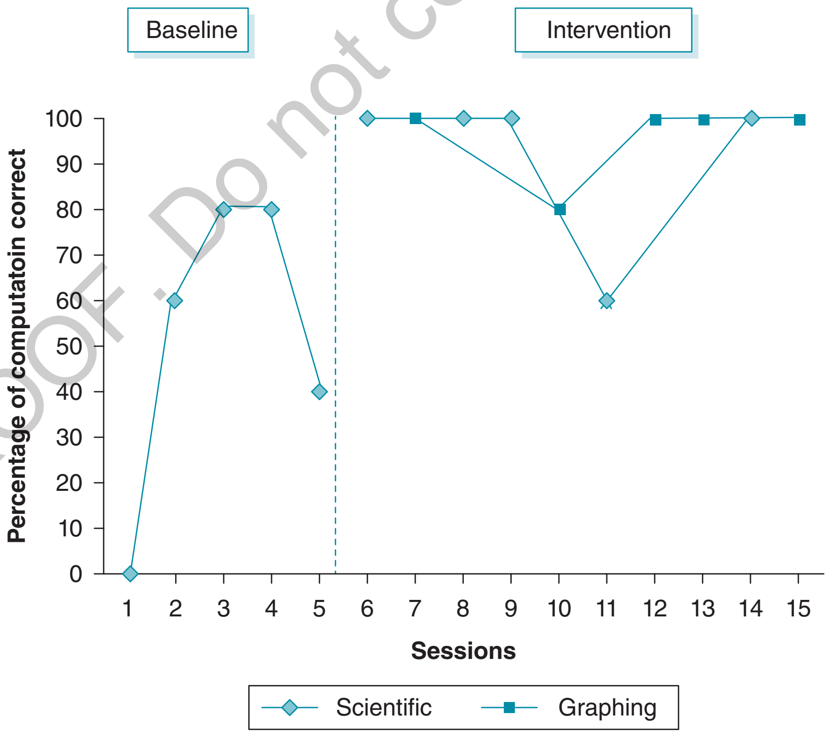
For situations where we want to compare different treatments or treatment conditions, we can use the alternating treatment design. The premise here is to alternate the treatment conditions during the treatment phase and compare how the individual performs under different conditions or treatments. If the performance of the participant does not

An **alternating treatment (ABC) design** is a single-case experimental design in which a baseline phase is followed by a treatment phase in which the conditions or treatments are alternated. The performance of the participant is compared under the different conditions or treatments.

change under the different conditions or treatments, then we can conclude that the different conditions or treatments do not influence performance. This design is also known as the **ABC design** and is illustrated in Figure 13.14.

Baseline data illustrate the accuracy of the participant prior to using the calculator. During the treatment (intervention) phase, the use of a scientific calculator (triangles) and a graphing calculator (squares) is alternated so accuracy of solving math problems can be compared. Notice at baseline, the accuracy in subtraction computations ranges from 0% to 20%. During treatment, there is an immediate increase to 100% for use of the graphing calculator. Accuracy with the use of the scientific calculator also reached 100% accuracy but took more sessions to reach that level. The accuracy of solving word problems was 0% at baseline. Performance of both types of calculators varied for the accuracy of solving word problems. Data are based on those presented by Yakubova and Bouck (2014).

Figure 13.14 An Alternating Treatment Design to Compare Two Types of Calculators on the Subtraction and Word Problem Accuracy of a Student With Mild Intellectual Disability



Source: Adapted from Yakubova & Bouck (2014). Reproduced with permission.

The advantage of the alternating treatment design is the ability to directly compare different conditions or treatments in the same study. One disadvantage is a possible carryover effect in comparing the treatment or conditions. The learning from one condition or treatment may carry over to the other condition or treatment. Therefore, the conditions or treatments must be distinctly different. Yakubova and Bouck (2014) compared two different types of calculators that operate differently to solve novel math problems. A second disadvantage to the alternating treatment design is that the alternated treatments must be able to alternate rapidly from treatment session to the next and the dependent variable must be able to change rapidly. For example, studies to compare effects of medications for attention-deficit hyperactivity disorder (ADHA) cannot be used in an alternating treatment design as many medications cannot be alternated or combined safely, and a medication may take time to take effect.

13.7 Application of Single-Case Designs in an Applied School Setting

Experimental designs have three main characteristics that enable these designs to infer cause-and-effect conclusions. These characteristics include manipulation, comparison, and randomization. Single-case designs do include manipulation of the independent variable and a comparison across individuals or conditions. *Randomization* typically implies random selection in nonexpermiental designs or random assignment to treatment/comparison groups in experimental designs. Randomization is important to the generalizability of the study to a larger population and in the application of inferential statistics. In single-case designs, randomization can include different forms of randomization depending on the type of single-case study. For example, conditions can also be randomized by randomly determining the presentation of the conditions. In single-case designs with more than one participant, participants can be randomly assigned to the order in which they will enter the treatment. In an applied educational setting where much of the educational research is conducted, such randomization is not always possible. Educational research is conducted in the confines of the schedule of school activities during the day, which may interfere with randomizing treatment conditions that need to occur rapidly. Single-case researchers also often purposefully select the order of entry of the participants based on who they believe will respond quicker to the intervention since the subsequent participants cannot enter until the preceding participant has demonstrated change. Randomly selecting a slower responding participant to enter treatment first will delay the entry of the other participants and extend the length of the study. Single-case researchers also underemphasize the need for randomization since inferential statistics cannot be used to analyze single-case data. To counteract the potential lack of generalizability, single-case research will include a more in-depth description of the participants, often providing a detailed description of each of the participants individually.

13.8 Single-Case Designs in the Identification of Effective Educational Programs, Policies, and Practices

The use of statistics is an important element in measuring student outcomes and defining effective educational programs, policies, and practices. We mentioned earlier in this chapter that analysis of single-case research is primarily through visual analysis. The data provided in single-case research do not lend themselves easily to quantitative analysis because they violate many of the assumptions needed to apply to statistics. One example of this violation is the assumption of independent data. This means that the data being analyzed need to be independent of each other. In the case of group designed research, all participants in the group are considered independent since they represent different participants. In the case of single-case research, the data collected come from the same participant over time; hence, the data are not independent. Therefore, single-case research is not currently used to identify effective educational practices by the What Works Clearinghouse (http://ies.ed.gov/ncee/wwc/). Statisticians are currently working on this issue to develop a statistical procedure so that single-case studies can be used to identify effective educational practice that can be included in the What Works Clearinghouse.

Even though single-case research is not used to identify effective educational programs, policies, and practices in the What Works Clearinghouse, it is still a very useful research tool that can be used to evaluate the effectiveness of educational work with individual participants. The *What Works Clearinghouse Procedures and Standards Handbook* (U.S. Department of Education, 2010) does supply guidelines that can be used to judge the quality of single-case research. A high-quality single-case research study will include (a) more than one assessor of the dependent variable and information regarding interrater agreement of the multiple assessors; (b) at least three demonstrations of the effect of the treatment across participants, materials, or settings; (c) at least three data points in the baseline phase; and (d) at least 20 to 30 data points across phases depending on the specific design.

LEARNING CHECK 2 ✓

1. When would a single-case design be used instead of a group design?
2. Identify whether each of the following is an example of a reversal design, a multiple-baseline design, a changing-criterion design, or an alternating treatment design:
   A. A researcher gives a child successively greater levels of positive reinforcement after an initial baseline phase to reduce how often the student shouts out during class. The successive treatments are administered until the child has reached a level where she is no longer shouting out in class.

B. A researcher records the duration of time a participant stays on task in math class 4 days before, 4 days during, and 4 days after a behavioral intervention strategy is implemented.

C. A researcher records the number of times a child raises her hand in class for 5 days before and 10 days while implementing two different types of reinforcement that are provided on alternate days.

D. A researcher records the level of engagement made by three participants. Each participant was given a treatment phase after 3, 4, or 5 days of a baseline phase; no baseline phase was given after the treatment was administered.

3. For a single-case experimental study, why would a researcher use a multiple-baseline design instead of a reversal design?

Answers: 1. When we want to study an individual case rather than a group; 2. A. changing-criterion design, B. reversal design, C. alternating treatment design, D. multiple-baseline design; 3. A multiple-baseline design would be used when it is not possible for changes in a dependent variable to return to baseline.

13.9 Validity, Stability, Magnitude, and Generality

The analysis of single-case experimental research designs is based largely on a visual inspection of the data in a graph and not based on statistical analyses that require data to be grouped across multiple participants or groups. The specific visual features in a graph that indicate the validity of an observation are described in this section.

The stability and magnitude of change across phases in a single-case design determine the extent to which a researcher has established control. The greater the control, the higher the internal validity.

Internal Validity, Stability, and Magnitude

Recall from Chapter 6 that internal validity is the extent to which we can demonstrate that a manipulation or treatment causes a change in a dependent measure. Importantly, the extent to which we establish experimental control of all other possible causes is directly related to the internal validity of a research study. The greater the control we establish, the higher the internal validity.

A single-case design requires a visual analysis of the graphical data of a single participant. The level of control and therefore the internal validity of a single-case design can be determined when the following two features are observed in a graph using this type of analysis:

- The stability in the pattern of change across phases
- The magnitude or size of the change across phases

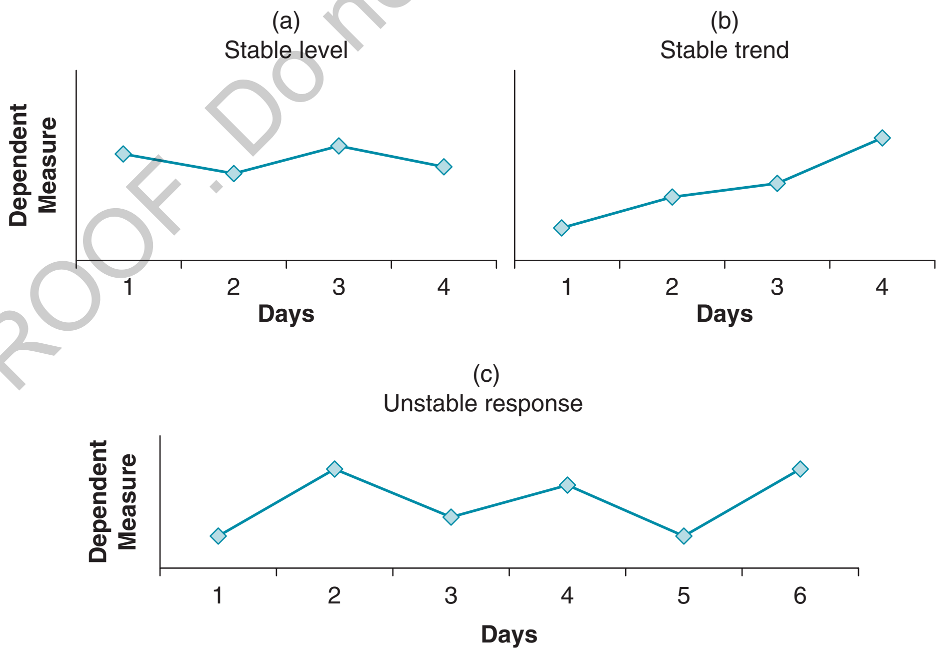
Stability is the consistency in the pattern of change in a dependent measure in each phase of a design. The more stable or consistent changes in a dependent measure are in each phase, the higher the internal validity of a research design.

In a visual inspection of a graph, the **stability** of a measure is indicated by the consistency in the pattern of change in each phase. The stability of a dependent measure is illustrated in Figure 13.15. Data in a given phase can show a stable level, as in Figure 13.15a; show a stable trend, as in Figure 13.15b; or be unstable, as in Figure 13.15c. The stability of a measure in each phase is important because when a measure is unstable, changes are occurring in a dependent variable even when the researcher is not manipulating the behavior. When a dependent measure is stable, we can be confident that any changes in level or trend were caused by the manipulation because changes only occurred between each phase and were otherwise stable or consistent within each phase. Therefore, the more stable a measure, the greater the control and the higher the internal validity in an experiment.

Another level of control can be demonstrated by the **magnitude** of change, which is the size of the change in a dependent measure observed between phases. When a measure is stable *within* each phase, we look at the magnitude of changes *between* phases. For a treatment to be causing changes in a dependent measure, we should observe immediate changes as soon as the treatment phase is administered. We can observe an immediate change in level, as shown in Figure 13.16a, or we can observe an immediate change in trend, as shown in Figure 13.16b. The greater the magnitude of changes between phases, the greater the control and the higher the internal validity in a single-case experiment.

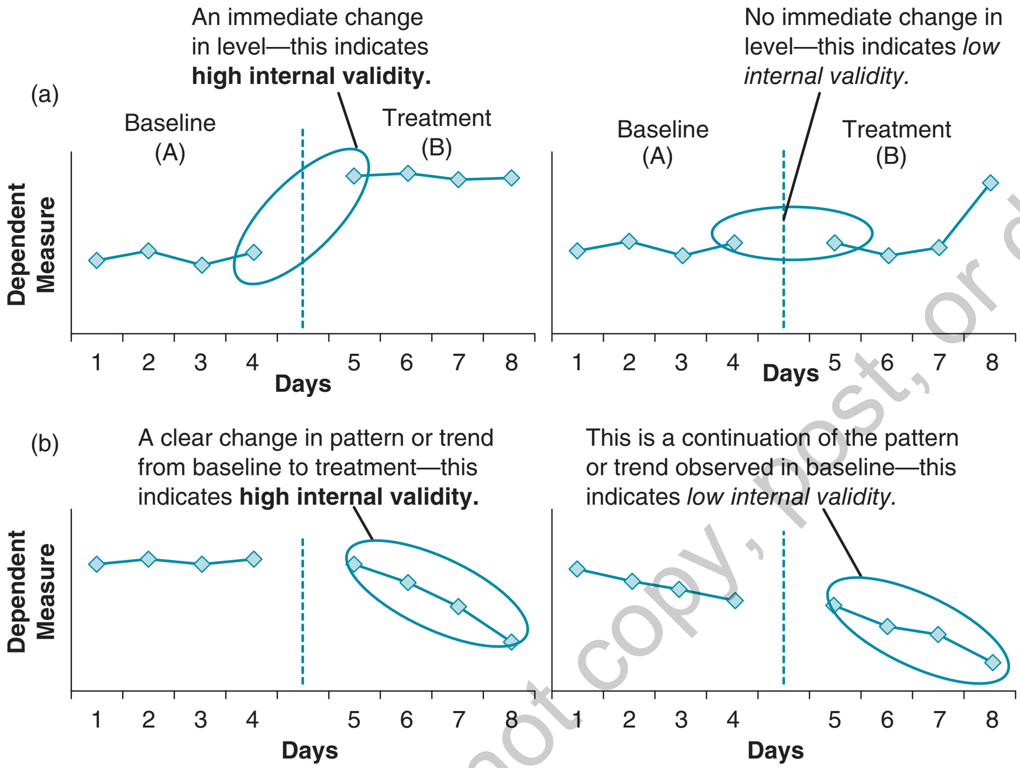
Magnitude is the size of the change in a dependent measure observed between phases of a design. The larger the magnitude of changes in a dependent measure between phases, the higher the internal validity of a research design.

Figure 13.15 A Stable Level (a), a Stable Trend (b), and an Unstable Response (c)



Note: Graphs (a) and (b) show a response that indicates high internal validity, whereas graph (c) indicates low internal validity.

Figure 13.16 Internal Validity and Control



Note: Identifying an immediate change in level (top row, a) or a change in trend (bottom row, b) that would indicate a high level of control and high internal validity.

External Validity and Generality

Recall from Chapter 6 that external validity is the extent to which observations generalize beyond the constraints of a study. A single-case design is typically associated with low population validity, which is a subcategory of external validity. In other words, it is not possible to know whether the results in the sample would also be observed in the population from which the sample was selected because single-case experimental designs are associated with very small sample sizes. However, the results in a single-case design can have high external validity in terms of generalizing across behaviors, across subjects or participants, and across settings. The following is an example of each way to generalize results to establish the external validity of a single-case experiment:

- As an example of generalizing across behaviors, a researcher may examine the extent to which an intervention for reducing aggression toward others generalizes to or reduces self-injury. In this example, the therapist generalizes across behaviors, from aggression toward others (Behavior 1) to self-injury (Behavior 2).

- As an example of generalizing across subjects or participants, a researcher may examine the effectiveness of an intervention to improve reading skills across multiple participants.
- As an example of generalizing across settings, a teacher may want to determine the extent to which an intervention to increase social exchanges with peers during recess generalizes to increases in social exchanges with peers during group activities during science. In this example, the researcher generalizes across settings, from social exchanges during recess (Setting 1) to social exchanges during science (Setting 2).

13.10 Ethics in Focus: The Ethics of Innovation

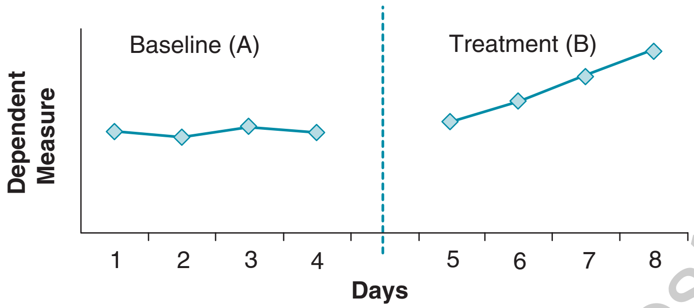
Many single-case experiments look at early treatments for behaviors needed for learning (such as attending to task or engagement) or improving academic skills. When these types of behaviors are studied using a single-case design, the treatment is typically hypothesized to have benefits, such as increasing the frequency of behaviors needed for learning, reducing the frequency of behaviors that interfere with learning, or teaching new skills. Researchers will end an experiment with the treatment phase that was most beneficial, so as to maximize the benefits that participants receive. In a reversal design, this means that researchers end the study in a B phase (e.g., ABAB). A multiple-baseline design, an alternating treatment, and a changing-criterion design already end in a treatment phase. Adding a treatment phase or otherwise adapting a single-case design is quite manageable for researchers because they observe only one or a few subjects or participants in a single-case experiment. Observing such a small sample size allows researchers the flexibility to make changes, such as when they add or omit treatments to maximize benefits to participants. Teachers can use this design to make data-based decisions regarding the efficacy of the instruction they provide in producing the desired change in the students.

Researchers often generalize observations in single-case experiments across behaviors, across subjects or participants, and across settings.

The flexibility of a single-case design also allows for greater "investigative play" (Hayes, 1981, p. 193) or greater freedom to ask innovative or new questions about treatments with unknown causes or with unknown costs or benefits. Single-case designs allow for the conduct of such innovative research to rigorously evaluate potential, yet untested, treatments with small samples, thereby testing the treatment without exposing such a treatment to large groups of participants, particularly when the potential costs of implementing such a treatment are largely unknown or untested. In this way, single-case designs can be used as an initial research design for testing some of the most innovative research in the behavioral sciences, which can then be tested in larger group studies.

LEARNING CHECK 3 ✓

1. Perform a visual inspection of the following data. Does the graph illustrate a study with high internal validity? Explain.



2. A researcher uses a single-case design to record the number of minutes spent studying in a baseline phase and a calming music treatment phase with a student who studied in a school media center and the same student who studied in a study hall room. Based on this description, can the researcher generalize across behaviors, across participants, or across settings?

3. Single-case designs allow for greater freedom to ask innovative or new questions about educational interventions with unknown costs or benefits. Why can a single-case design be an ethically appropriate research design to test the effectiveness of such treatments?

Answers: 1. Yes, because the data in baseline are stable, and a change in trend from baseline to treatment; 2. Generalize across settings; 3. Because single-case designs are used with small samples, thereby testing the treatment without exposing such a treatment to large groups of participants.

CHAPTER SUMMARY

LO 1 Define and identify a quasi-experiment and a quasi-independent variable.

- A quasi-experimental research design is structured similar to an experiment, except that this design lacks random assignment, includes a preexisting factor (i.e., a variable that is not manipulated), or does not include a comparison/control group.
- A quasi-independent variable is a preexisting variable that is often a characteristic inherent to an individual, which differentiates the groups or conditions being compared in a research study. Because the levels of the variable are preexisting, it is not possible to randomly assign participants to groups.

LO 2 Identify and describe two one-group quasi-experimental research designs: the posttest-only and pretest-posttest designs.

- The **one-group posttest-only design** is a quasi-experimental research design in which a dependent variable is measured for one group of participants following a treatment.
- The **one-group pretest-posttest design** is a quasi-experimental research design in which the same dependent variable is measured in one group of participants before and after a treatment is administered.

LO 3 Identify and describe two nonequivalent control group quasi-experimental research designs: the posttest-only and pretest-posttest designs.

- A **nonequivalent control group** is a control group that is matched upon certain preexisting characteristics similar to those observed in a treatment group but to which participants are not randomly assigned. When a nonequivalent control group is used, **selection differences** can potentially explain an observed difference between an experimental and a nonequivalent control group.
- The **nonequivalent control group posttest-only design** is a quasi-experimental research design in which a dependent variable is measured following a treatment in one group and is compared to a nonequivalent control group that does not receive the treatment.
- The **nonequivalent control group pretest-posttest design** is a quasi-experimental research design in which a dependent variable is measured in one group of participants before (pretest) and after (posttest) a treatment and that same dependent variable is also measured at pretest and posttest in a nonequivalent control group that does not receive the treatment.

LO 4 Identify and describe three time-series quasi-experimental research designs: basic, interrupted, and control designs.

- The **basic time-series design** is a quasi-experimental research design in which a dependent variable is measured at many different points in time in one group before and after a treatment that is manipulated by the researcher.
- The **interrupted time-series design** is a quasi-experimental research design in which a dependent variable is measured at many different points in time in one group before and after a treatment that naturally occurs.
- A **control time-series design** is a basic or interrupted time-series quasi-experimental research design that also includes a nonequivalent control group that is observed during the same period of time as a treatment group but does not receive the treatment.

LO 5 Identify and describe three developmental quasi-experimental research designs: longitudinal, cross-sectional, and cohort-sequential designs.

- A **longitudinal design** is a quasi-experimental research design used to study changes across the life span by observing the same participants over time and measuring the same dependent variable at each time.

- A **cross-sectional design** is a quasi-experimental research design in which participants are grouped by their age and participant characteristics are measured in each age group. Each age group is a **cohort,** so this design is prone to **cohort effects,** which occur when unique characteristics in each cohort can potentially explain an observed difference between groups.
- A **cohort-sequential design** is a quasi-experimental research design that combines longitudinal and cross-sectional techniques by observing different cohorts of participants over time at different or overlapping ages.

LO 6 Define the single-case experimental design.

- The **single-case experimental design** is an experimental research design in which an individual case serves as its own control and the dependent variable measured is analyzed for each individual case and is not averaged across groups or across participants. This design has specific requirements of the number of data points within and across phases to be considered high quality.

LO 7 Identify and describe three types of single-case research designs: the reversal, multiple-baseline, and changing-criterion designs.

- The **reversal design** is a single-case experimental design in which a single case is observed before (A), during (B), and after (A) a treatment or manipulation.
- The **multiple-baseline design** is a single-case experimental design in which a treatment is successively administered over time to different participants, for different behaviors, or in different settings.
- The **changing-criterion design** is a single-case experimental design in which a baseline phase is followed by successive treatment phases in which some criterion or target level of behavior is changed from one treatment phase to the next. The participant must meet the criterion of one treatment phase, before the next treatment phase is administered.
- The alternating treatment design is a single-case experimental design in which the baseline phase is followed by a treatment phase where at least two treatments are delivered on alternating sessions. The treatments are then compared to determine which one was more effective.

LO 8 Identify in a graph the stability and magnitude of a dependent measure, and explain how each is related to the internal validity of a single-case design.

- The **stability** of a measure is the consistency in the pattern of change in a dependent measure in each phase of a design. The more stable or consistent changes in a dependent measure are in each phase, the higher the internal validity of a research design.
- The **magnitude** of change in a measure is the size of the change in a dependent measure observed between phases of a design. A measure can have a change in level or a change in trend. The larger the magnitude of change, the greater the internal validity of a research design.

LO 9 Identify three ways that researchers can strengthen the external validity of a result using a single-case design.

- A single-case design is typically associated with low population validity (a subcategory of external validity). However, researchers can strengthen the external validity of a result using a single-case design by generalizing across behaviors, across subjects or participants, and across settings.

KEY TERMS

ABA design 358

alternating treatment design (ABC) 000

baseline phase (A) 351

basic time-series design 342

changing-criterion design 356

cohort 000

cohort effects 000

cohort-sequential design 000

control time-series design 345

cross-sectional design 000

interrupted time-series design 344

longitudinal design 000

magnitude 362

multiple-baseline design 354

nonequivalent control group 338

nonequivalent control group posttest-only design 339

nonequivalent control group pretest-posttest design 340

one-group posttest-only design 336

one-group pretest-posttest design 337

phase 351

quasi-experimental research design 334

quasi-independent variable 335

reversal design 351

selection differences 338

single-case experimental design 349

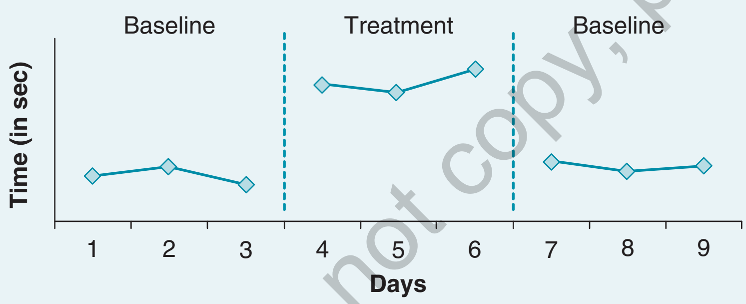
stability 361

REVIEW QUESTIONS

1. A quasi-experimental research design is structured similar to an experiment, with what two exceptions?
2. State whether each of the following factors is an example of an independent variable or a quasi-independent variable. Only state "quasi-independent variable" for participant variables that cannot be manipulated.
   A. The age of participants
   B. Time allotted for taking an exam
   C. A teacher's prior teaching experience
   D. Time of day a study is conducted

E. A participant's state of residence
F. Amount of time spent studying

3. How does a one-group pretest-posttest design improve on the posttest-only quasi-experimental design? What is the major limitation of all one-group designs?
4. What is a nonequivalent control group, and why does this type of group make it difficult to determine cause and effect using a nonequivalent control group quasi-experimental design?
5. What is the key difference between the basic and interrupted time-series quasi-experimental research designs?
6. A reversal design is used to test the hypothesis that low lighting in a room reduces how quickly students read. As shown in the figure for one student, a student reads passages of similar length in a room with normal lighting (baseline), then in the same room with dim lighting (treatment), and then again with normal lighting. Do the results shown in the figure support the hypothesis? Explain.



7. What is the most likely reason that a researcher uses a multiple-baseline design instead of a reversal design?
8. Define the changing-criterion design and explain when the design is used.
9. Describe how an alternating treatment design differs from the other single-case designs.
10. In the following scenarios, is the researcher generalizing across behaviors, participants, or settings?
    A. A researcher examines the generalizability of an educational intervention for increasing vocabulary acquisition by testing the same treatment to increase reading comprehension.
    B. A researcher examines if the effectiveness of a new learning system used in a classroom is also effective when used in a home (for homeschooled children).

ACTIVITIES

1. Use an online database, such as PsycINFO, to search scientific research articles for any topic you are interested in. Perform two searches. In the first search, enter a search term

related to your topic of interest, and enter the term *longitudinal* to find research that used this design in your area of interest. Select and print one article. In the second search, again enter a search term related to your topic of interest, and this time enter the term *cross-sectional* to find research that used this design in your area of interest. Again, select and print one article. Once your searches are complete, complete the following assignment:

A. Write a summary of each article, and explain how each research design differed.

B. Describe at least two potential threats to internal validity in each study.

C. Include the full reference information for both articles at the end of the assignment.

2. A researcher proposes that parental involvement will improve school attendance of the child. (a) Write a research plan to test this hypothesis using a single-case experimental design. (b) What is the predicted outcome or pattern if the hypothesis that parental involvement will improve attendance is correct? (c) Graph the expected results. (d) Identify the extent to which your results demonstrate high or low internal validity.