Using Designs 11, 12, and 13 in Single-Subject Studies

Reversal, alternating treatments, and multiple baseline designs can be used not only with groups but also with single individuals, in what are collectively known as **single-subject designs**. A study by Deaver, Miltenberger, and Stricker (2001) illustrates how a researcher might use two of these—reversal and multiple baseline—simultaneously. A 2-year-old girl named Tina had been referred for treatment because she often twirled her hair with her fingers so vigorously that she pulled out some of her hair. On one occasion she wrapped her hair around a finger so tightly that the finger began to turn blue and the hair had to be removed with scissors. Tina engaged in such behavior primarily when she was alone (e.g., at naptime); hence, there was no parent or other adult present to discourage it. The researchers identified a simple treatment—putting thin cotton mittens on Tina's hands—and wanted to document its effect. They videotaped Tina's behaviors when she was lying down for a nap in either of two settings, her bedroom at home or her daycare center, and two observers independently counted the number of hair twirling incidents as they watched the videotapes. Initially, the observers collected baseline data. Then, during separate time periods for the bedroom and daycare settings, they gave Tina the mittens to wear during naptime. After reversing back to baseline in both settings, they had Tina wear the mittens once again. The percentages of time that Tina twirled her hair in the two settings over the course of the study are presented in Figure 9.3.

In both the bedroom and daycare settings, the researchers alternated between baseline and treatment; this is the *reversal* aspect of the study. Furthermore, they initiated and then later reinstituted the treatment at different times in the two settings; this is the *multiple baseline* aspect of the study. Figure 9.3 consistently shows dramatic differences in hair twirling during baseline versus mittens conditions, leading us to conclude that the mittens, rather than some other factor, were almost certainly the reason for the disappearance of hair twirling.

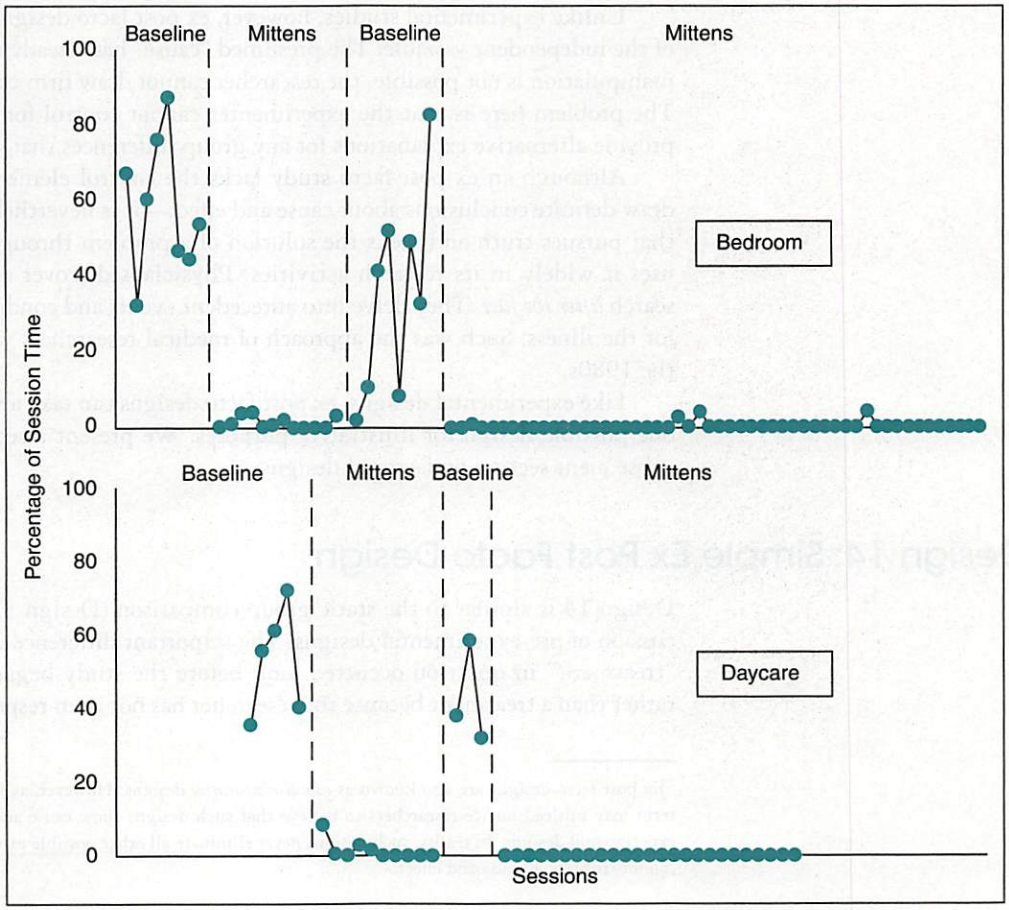

FIGURE 9.3

Percentage of session time in which hair twirling was observed both in the bedroom and at daycare

Reprinted from "Functional Analysis and Treatment of Hair Twirling in a Young Child" by C. M. Deaver, R. G. Miltenberger, & J. M. Stricker, 2001, *Journal of Applied Behavior Analysis, 34,* p. 537. Reprinted with permission of the Society for the Experimental Analysis of Behavior, Inc.

Ex Post Facto Designs

In many situations, it is either unethical or impossible to manipulate certain variables in order to investigate their potential influence on other variables. For example, a researcher cannot intentionally infect people with a potentially deadly new virus, withhold instruction, ask parents to abuse their children, or modify a person's personality to compare the effects of these factors on the dependent variables in one's research problem.

Ex post facto designs[2] (the term *ex post facto* literally means "after the fact") provide an alternative means by which a researcher can investigate the extent to which specific independent variables—perhaps involving a virus, lack of schooling, a history of family violence, or a personality trait—may possibly affect the dependent variable(s) of interest. In an ex post facto study, a researcher identifies *events that have already occurred or conditions that are already present* and then collects data to investigate a possible relationship between these factors and subsequent characteristics or behaviors. In particular, after observing that differing circumstances have prevailed for two or more different groups—such circumstances comprise the independent variable—the researcher tries to determine whether the groups differ on some other, dependent variable. For example, a researcher might identify two groups of adults with different immunization records—those who, as children, were vaccinated against measles and those who were not—and then calculate the percentage of reported cases of measles in each group. Similarly, a researcher might identify two groups of 10-year-olds—those who had extensive musical training in preschool and those whose preschools provided no such training—and compare the musical skills of the two groups of children.

Ex post facto designs are often confused with correlational or experimental designs because they share certain characteristics with each of these other design types. Like correlational research, ex post facto research involves looking at existing circumstances. But like experimental research, it has clearly identifiable independent and dependent variables.

Unlike experimental studies, however, ex post facto designs involve no direct manipulation of the independent variable: The presumed "cause" has already occurred. To the extent that such manipulation is not possible, the researcher cannot draw firm conclusions about cause and effect. The problem here is that the experimenter cannot control for confounding variables that may provide alternative explanations for any group differences that are observed.

Although an ex post facto study lacks the control element—and so does not allow us to draw definite conclusions about cause and effect—it is nevertheless a legitimate research method that pursues truth and seeks the solution of a problem through the analysis of data. Medicine uses it widely in its research activities. Physicians discover an illness and then initiate their search *after the fact.* They delve into antecedent events and conditions to discover a possible cause for the illness. Such was the approach of medical researchers when the AIDS virus emerged in the 1980s.

Like experimental designs, ex post facto designs can take a variety of forms. Here we present one possible design for illustrative purposes. We present a second ex post facto design in the subsequent section on factorial designs.

Design 14: Simple Ex Post Facto Design

Design 14 is similar to the static group comparison (Design 3), which we included in our discussion of pre-experimental designs. The important difference is one of timing: In this case, the "treatment" in question occurred long before the study began; hence, we call it an *experience* rather than a treatment because the researcher has not been responsible for imposing it. A simple

[2] Ex post facto designs are also known as *causal-comparative* designs. However, as Johnson (2001) has pointed out, the latter term may mislead novice researchers to believe that such designs show cause and effect as clearly and definitively as true experimental designs. In reality, such designs never eliminate all other possible explanations for an observed effect; thus, they cannot truly show cause and effect.

ex post facto design can be depicted as follows, where *Exp* refers to a prior experience that one group has had and another has not:

Group	Time →	
	Prior event(s)	*Investigation period*
Group 1	Exp	Obs
Group 2	—	Obs

An obvious variation on this design is one in which Group 2 has an experience as well, albeit a different experience from that of Group 1.

Such designs are common in studying the possible effects of previously occurring environmental variables such as television viewing habits, child abuse, and malnutrition. They are also used in studying the potential influences of pre-existing characteristics—perhaps those that are inherited or congenital—such as gender, mental illness, and physical disability. (In the latter instances, we might want to replace the term *experience* with a term such as *characteristic.)* The most we can conclude from these studies is that certain behaviors or other variables tend to be *associated* with certain pre-existing conditions; we can never determine that those other variables were actually caused by those conditions.

Factorial Designs

Thus far, we have been describing designs in which only one independent variable is studied. Yet in many situations, a researcher examines the effects of two or more independent variables in a single study; this approach is known as a **factorial design**.

Design 15: Two-Factor Experimental Design

In its simplest form—one involving two independent variables, which we call *Variable 1* and *Variable 2*—such a design might look something like the following:

	Group	Time →		
		Treatments related to the two variables may occur simultaneously or sequentially		
		Treatment related to Variable 1	*Treatment related to Variable 2*	
Random Assignment	Group 1	Tx_1	Tx_2	Obs
	Group 2	Tx_1	—	Obs
	Group 3	—	Tx_2	Obs
	Group 4	—	—	Obs

We can determine the effects of the first independent variable by comparing the performance of Groups 1 and 2 with that of Groups 3 and 4. We can determine the effects of the second independent variable by comparing Groups 1 and 3 with Groups 2 and 4. If you think you've seen this design before, in a way you have. This is simply a more generalized form of the Solomon four-group design (Design 5), but we are no longer limiting ourselves to having the presence or absence of a pretest be one of our independent variables.

Such a design allows us to examine not only the possible effects of two independent variables but also the possible *interaction* of the variables as they influence the dependent variable. For example, imagine that, after presenting both treatments, we find that Groups 2, 3, and 4 show similar performance but that Group 1 outperforms the other three. Such a result may indicate that neither independent variable produces a particular effect on its own—that *both* variables are necessary to bring about the effect.

Design 16: Combined Experimental and Ex Post Facto Design

In the factorial design just presented, participants are randomly assigned to groups in a true experimental study. But it is also possible to combine elements of experimental research and ex post facto research into a single factorial design. In its simplest form, such a design looks like the following:

Group	Time →				
	Prior event(s)	*Investigation period* →			
Group 1	Exp_a	Random Assign-ment	Group 1a	Tx_a	Obs
			Group 1b	Tx_b	Obs
Group 2	Exp_b	Random Assign-ment	Group 2a	Tx_a	Obs
			Group 2b	Tx_b	Obs

In this case, the researcher initially divides the sample into two groups based on the participants' previous experiences or pre-existing conditions; this is the *ex post facto* part of the study. Then the researcher randomly assigns members of each group to one of two treatment groups (or perhaps a treatment group and a control group); this is the *experimental* part of the study. The result is four groups that represent all four possible combinations of the previous experience/pre-existing characteristic and the treatment variable. Such a design enables the researcher to study how an experimental manipulation may influence a particular dependent variable *and* how a previous experience or pre-existing characteristic may possibly interact with that manipulation.

In a variation of such a design, the experimental manipulation might be a within-subjects variable rather than a between-groups variable. As an example, one of us authors once joined forces with two colleagues and a graduate student to test the hypothesis that people with different educational backgrounds interpret and remember maps differently and, more specifically, that only people with a background in geography apply general principles of geography when they interpret maps (J. E. Ormrod, Ormrod, Wagner, & McCallin, 1988). We constructed two maps to test our hypothesis. One map was arranged in accordance with the patterns of a typical city; for instance, a downtown business district was located at a point where it could be easily reached from different directions (this is typical), and factories, a lumberyard, and low-income housing were situated near railroad tracks (also typical). The second map was less "logical" in the sense that it violated basic geographic principles; for instance, a

river originated in the plains and ran *up* into a mountain range, and various transportation networks did not interconnect in ways that they normally do. The two different maps reflected one of our independent variables: logic (or lack thereof) of the spatial arrangement of features within a map.

Three groups of college professors—geographers, sociologists, and educational psychologists—provided the basis for our second independent variable: educational background. We asked each professor to study each of the two maps aloud for three two-minute intervals (we tape-recorded what they said during the study sessions) and then, after each interval, to draw as much of the map as he or she could remember.

Thus, if we call the two maps Tx_a (logical map) and Tx_b (illogical map), our design looked like the following:

Group	Time →							
Geographers	Tx_a	Obs	Obs	Obs	Tx_b	Obs	Obs	Obs
Sociologists	Tx_a	Obs	Obs	Obs	Tx_b	Obs	Obs	Obs
Educational psychologists	Tx_a	Obs	Obs	Obs	Tx_b	Obs	Obs	Obs

In this situation, one independent variable—the logic or illogic of the map presented—was a variable we directly manipulated, and we presented it to all participants in a *within-subjects* (repeated-measures) manner. The second independent variable, educational background, was a pre-existing condition and therefore something we could *not* control; this was the *ex post facto* part of the design.

The upshot of the study was that there was an *interaction* between the two independent variables, map logic and educational background. In particular, the geographers remembered more of the logical map than they did of the illogical map; in contrast, the sociologists and educational psychologists remembered each map with equal accuracy. We interpreted this result to indicate that only the geographers were applying geographic principles to study the maps and that they could use such principles effectively only with the geographically logical one. We supported our conclusion with a qualitative element in our study; that is, we used a *mixed-methods design.* In particular, we conducted content analyses of the professors' study sessions. Indeed, the content analyses revealed that the geographers had applied many geographic principles to the logical map but had trouble applying them to the illogical one. Meanwhile, the sociologists and educational psychologists studied both maps in a haphazard fashion, with few attempts to make sense of what they saw on the maps.

Table 9.1 provides a summary of the pre-experimental, experimental, quasi-experimental, ex post facto, and factorial designs described in the preceding sections. Keep in mind that, as stated earlier, this is not an exhaustive list of experimental and ex post facto designs. You can combine and expand on these designs in a number of ways—and perhaps incorporate elements of qualitative or descriptive-quantitative designs (e.g., content analysis or longitudinal data collection) as well—to more effectively address your own research question.

TABLE 9.1 Summary of Experimental and Ex Post Facto Designs

Name of the Design	Aim of the Research	Graphic Depiction	Comments on the Design
Pre-Experimental Designs			
1. One-shot experimental case study	To show that one event (a treatment) precedes another event (the observation)	Group / Time → : Group 1 — Tx — Obs	Shows a before-and-after sequence but cannot substantiate that this is a cause-and-effect relationship.
2. One group pretest–posttest design	To show that change occurs after a treatment	Group / Time → : Group 1 — Obs — Tx — Obs	Provides a measure of change but yields no conclusive results about the cause of the change.
3. Static group comparison	To show that a group receiving a treatment behaves differently than one receiving no treatment	Group / Time → : Group 1 — Tx — Obs; Group 2 — — — Obs	Fails to determine pretreatment equivalence of groups.
True Experimental Designs			
4. Pretest–posttest control group design	To show that change occurs following, but only following, a particular treatment	Group / Time → ; Random Assignment: Group 1 — Obs — Tx — Obs; Group 2 — Obs — — — Obs	Controls for many potential threats to internal validity.
5. Solomon four-group design	To investigate the possible effect of pretesting	Group / Time → ; Random Assignment: Group 1 — Obs — Tx — Obs; Group 2 — Obs — — — Obs; Group 3 — — — Tx — Obs; Group 4 — — — — Obs	Enables the researcher to determine how pretesting may affect the final outcome observed.
6. Posttest-only control group design	To determine the effects of a treatment when pretesting cannot or should not occur	Group / Time → ; Random Assignment: Group 1 — Tx — Obs; Group 2 — — — Obs	Uses the last two groups in the Solomon four-group design; random assignment to groups is critical for maximizing group equivalence.
7. Within-subjects design	To compare the relative effects of different treatments for the same participants	Group / Time → : Group 1 — Tx_a — Obs_a; Tx_b — Obs_b	Useful only when effects of each treatment are temporary and localized.
Quasi-Experimental Designs			
8. Nonrandomized control group pretest–posttest design	To show that two groups are equivalent with respect to the dependent variable prior to the treatment, thus eliminating initial group differences as an explanation for posttreatment differences	Group / Time → : Group 1 — Obs — Tx — Obs; Group 2 — Obs — — — Obs	Differs from experimental designs because test and control groups are not totally equivalent; equivalence on the pretest ensures equivalence only for variables that have specifically been measured.

TABLE 9.1 Summary of Experimental and Ex Post Facto Designs *(continued)*

Name of the Design	*Aim of the Research*	*Graphic Depiction*	*Comments on the Design*
Quasi-Experimental Designs *(continued)*			
9. Simple time-series experiment	To show that, for a single group, change occurs during a lengthy period only after the treatment has been administered	Group / Time → Group 1: Obs, Obs, Tx, Obs, Obs	Provides a stronger alternative to Design 2; external validity can be increased by repeating the experiment in different places under different conditions.
10. Control group, time-series design	To bolster the internal validity of the preceding design with the addition of a control group	Group / Time → Group 1: Obs, Obs, Tx, Obs, Obs Group 2: Obs, Obs, —, Obs, Obs	Involves conducting parallel series of observations for experimental and control groups.
11. Reversal time-series design	To show, in a single group or individual, that a treatment consistently leads to a particular effect	Group / Time → Group 1: Tx, Obs, —, Obs, Tx, Obs	Is an on-again, off-again design in which the experimental treatment is sometimes present, sometimes absent.
12. Alternating treatments design	To show, in a single group or individual, that different treatments have different effects	Group / Time → Group 1: Tx_a, Obs, —, Obs, Tx_b, Obs	Involves sequentially administering different treatments at different times and comparing their effects against the possible consequences of nontreatment.
13. Multiple baseline design	To show the effect of a treatment by initiating it at different times for different groups or individuals, or perhaps in different settings for a single individual	Group / Time → Group 1: —, Obs, Tx, Obs, Tx, Obs Group 2: —, Obs, —, Obs, Tx, Obs	Involves tracking two or more groups or individuals over time, or tracking a single individual in two or more settings, for a lengthy period of time, as well as initiating the treatment at different times for different groups, individuals, or settings.
Ex Post Facto Designs			
14. Simple ex post facto design	To show the possible effects of an experience that occurred, or a condition that was present, prior to the investigation	Group / Time → Group 1: Exp, Obs Group 2: —, Obs	May show a difference between groups but does not conclusively demonstrate that the difference is due to the prior experience/condition in question.
Factorial Designs			
15. Two-factor experimental design	To study the effects of two experimenter-manipulated variables and their possible interaction	Group / Time → Random Assignment: Group 1: Tx_1, Tx_2, Obs Group 2: Tx_1, —, Obs Group 3: —, Tx_2, Obs Group 4: —, —, Obs	Requires a larger sample size than two-group studies; random assignment to treatments is essential.
16. Combined experimental and ex post facto design	To study the possible effects of an experimenter-manipulated variable, a previously existing condition, and the interaction between the two	Group / Time → Group 1: Exp_a, Random Assignment: Group 1a: Tx_a, Obs; Group 1b: Tx_b, Obs Group 2: Exp_b, Random Assignment: Group 2a: Tx_a, Obs; Group 2b: Tx_b, Obs	Requires a larger sample size than two-group studies; random assignment to the experimenter-manipulated variable is essential.

PRACTICAL APPLICATION Determining Possible Cause-and-Effect Relationships

The research designs described in this chapter vary considerably in the degree to which they control for potential confounding variables—variables that threaten a study's internal validity—and thus they also vary in terms of the degree to which they enable a researcher to draw firm conclusions about cause-and-effect relationships. The following checklist can help you evaluate a research design with respect to its internal validity.

✔ CHECKLIST

Looking for Confounding Variables

If you are planning a study in which you hope to find one or more cause-and-effect relationships—or if, instead, you are evaluating another person's research proposal or report—scrutinize the study with the following questions in mind:

_____ 1. What are the independent and dependent variables in the study:

Independent variable(s):

_____

Dependent variable(s):

_____

_____ 2. Is every independent variable actively manipulated by the researcher? ___ Yes ___ No

_____ 3. If the researcher is manipulating one or more independent variables, what precautions is the researcher taking to ensure that the manipulation is minimizing or eliminating the potential effects of confounding variables? For example, is the researcher:

- Keeping certain other variables constant? If so, which ones?

_____

_____

- Including a control group or at least two treatment groups?

_____

- Randomizing assignment to groups?

_____

- Using a within-subjects (repeated-measures) design?

_____

- Using other appropriate strategies? If so, which ones?

_____

_____ 4. If the researcher is *not* manipulating one or more independent variables, what precautions is the researcher taking to control for confounding variables? For example, is the researcher:

- Using one or more pretests to assess before-treatment group equivalence? If so, what variables are being pretested?

_____

_____

- Identifying matched pairs? If so, on the basis of what variables?

_____

_____

- Statistically controlling for confounding variables? If so, which ones?

_____

_____ 5. If the researcher is conducting a single-group or single-subject study, is the researcher:
- Conducting a series of observations both before and after the intervention (a time-series design)? _____
- Alternating either between two or more treatments or between treatment and nontreatment, with a new observation being made after each treatment or non-treatment (a reversal design)? _____
- Beginning an intervention at different times for different individuals or different contexts (a multiple baseline design)? _____

_____ 6. What other variables in the study (either identified or not identified by the researcher) might potentially affect the dependent variable? _____

_____

_____

_____ 7. To what extent might each of the following factors threaten the study's internal validity? If any of these factors pose a potential threat, how is the researcher minimizing or eliminating its influence? (Refer to Figure 9.1.)

History:

_____

Maturation:

_____

Testing:

_____

Instrumentation:

_____

Statistical regression:

_____

Selection:

_____

Attrition:

_____

_____ 8. With your answers to the preceding questions in mind, explain whether the study's results justifiably demonstrate a cause-and-effect relationship:

_____

_____

_____

_____

Meta-Analyses

Remember, we can conclude that a cause-and-effect relationship exists between an independent variable and a dependent variable only when we have directly manipulated the independent variable and have controlled for confounding variables that might offer alternative explanations for any changes in the dependent variable. Even when we have taken such precautions, however, there is the possibility that our alleged "cause" doesn't really produce the effect we think it does—that the situation we have just observed is a one-time-in-a-million fluke.

In Chapter 4 we introduced the idea of *replication:* A research study should be repeatable. In fact, we gain greater confidence in our research findings when a study is repeated over and over again—perhaps with a different population, in a different setting, or with slight variations on the treatment implementation.

Once researchers have conducted many such replications, another researcher may come along and conduct a **meta-analysis**—that is, an analysis of the analyses. In particular, the researcher combines the results of many experimental and/or ex post facto studies to determine whether they yield consistent, predictable results. A meta-analysis is primarily a statistical technique, and so we describe this procedure in greater depth in Chapter 11.

Conducting Experiments on the Internet

In Chapter 8 we mentioned that some researchers now conduct research studies on the Internet. Although most of these studies can best be categorized as descriptive studies, we occasionally see experimental studies as well. For instance, one of us authors once visited the website "Psychological Research on the Net," which provides links to numerous sites that host online research projects.[3] To learn more about this growing approach to data collection, she became a participant in several online studies that were active at the time. Although most of the studies involved completing questionnaires and so appeared to be correlational or survey studies, one of them was clearly an experimental study. In particular, this author was asked to (a) read and study a story that was illustrated by several photographs; (b) read three additional stories, one of which was quite similar to the initial story; and (c) answer a series of questions about details in the stories. In a subsequent debriefing on the website, she learned that she had been randomly assigned to the experimental group in the second part of the study; other participants were assigned to a control group, in which all three stories were quite different from the initial story. The researcher was investigating the possible effects that a similar story in Part b might have on recall for the story in Part a.

Internet-based experimental studies don't necessarily have to be one-shot affairs. For example, in one online study (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008), researchers enticed people into participating in a three-session experiment with the promise that for every session they completed, their names would be entered into an end-of-study lottery that would award cash prizes. A total of 1,354 people completed all three sessions; they ranged in age from 18 to 72 and lived in various countries around the world. In Session 1 of the experiment, participants studied a list of 32 obscure trivia facts, such as the answer to "What European nation consumes the most spicy Mexican food?" (p. 1097), and they continued to study each fact until they could correctly recall it. After this first session, participants were divided into different treatment groups that varied in terms of the timing for Sessions 2 and 3, and they were sent e-mail messages when it was time to complete these subsequent sessions. In Session 2 (which might be as little as three minutes or as much as 105 days after Session 1), participants studied the trivia facts again, this time studying each one twice. Then, in Session 3 (which was 7, 35, 70, or 350 days after Session 2), participants were asked to remember as many of the facts as they could. The findings of the study are important for any conscientious student to note: Especially when the final test session was considerably delayed (e.g., by 2½ months or almost a year), people who spread out their studying more (i.e., those with a longer delay between Sessions 1 and 2) remembered more facts.[4] (In case you're curious, Norwegians are especially partial to spicy Mexican food.)

In some instances, an Internet-based research study might be quite suitable for your research question. Keep in mind, however, that ethical practices ensuring protection from harm, informed consent, and right to privacy are as important in online experimental research as they are in any face-to-face studies. The suggestions for ethical practices presented in Chapter 8 for online questionnaires are equally applicable to online experiments (see the Practical Application "Using the Internet to Collect Data for a Descriptive Study" on pp. 205–206).

[3]As noted in Chapter 8, you can reach the site by going to the website of the Association for Psychological Science (www.psychologicalscience.org); click on "Psychology links," scroll down to "Other Sites of Interest," and then click on "Online Psychology Experiments." Alternatively, you can go directly to the site, which, as this book goes to press, is located at psych.hanover.edu/research/exponnet.html.

[4]As you might guess, the attrition (dropout) rate was higher for participants with longer between-session delays. To determine the extent to which the differing attrition rates for different treatment groups might jeopardize the study's internal validity, the researchers collected basic demographic data at the beginning of Session 1. During the data analysis, the researchers found no significant differences in any demographic variables or in learning speed between participants who completed all three sessions and those who did not.

Remember, too, that the sample you get in an online study will hardly be representative of the overall population; for instance, it is likely to consist largely of college-educated, computer-literate people who enjoy participating in research studies. An additional problem is that you cannot observe your participants to determine whether they are accurately reporting demographic information (their age, gender, etc.) and whether they are truly following the instructions you present. Accordingly, unless you are interested in a topic such as very-long-term memory (as Cepeda and his colleagues were in their 2008 study) and can carefully control the conditions under which people are participating, we suggest that you use an Internet-based study primarily to formulate tentative hypotheses or to pilot-test experimental materials you plan to use in a more controlled and observable situation.

Testing Your Hypotheses, and Beyond

Experimental and ex post facto studies typically begin with specific research hypotheses, and subsequent statistical analyses should, of course, be conducted to test these hypotheses. Such analyses often take the form of a *t* test, analysis of variance, or analysis of covariance (we discuss these procedures in Chapter 11).

Yet one's analyses need not be restricted *only* to the testing of initially stated hypotheses. Oftentimes a study may yield additional results—-results that are unexpected yet intriguing—that merit analysis. There is no reason why the researcher can't examine these findings as well, perhaps statistically, perhaps not.

CONCEPTUAL ANALYSIS EXERCISE Identifying Research Designs

As a way of reviewing the designs described in this chapter, we offer a brief pop quiz. Following are short summaries of five research studies. The studies don't necessarily fit exactly into one of the design categories presented, but each one is definitely *experimental, quasi-experimental,* or *ex post facto* in nature. Identify the type of research that each study reflects. The answers appear after the suggested readings at the end of the chapter.

1. Two researchers want to see if a particular training program is effective in teaching horses to enter a horse trailer without misbehaving in the process—that is, without rearing, trying to turn around, or in some other way resisting entry into the trailer. Five horses (Red, Penny, Shadow, Sammy, and Fancy) go through the training, with each horse beginning training on a different day. For each horse, an observer counts the number of misbehaviors every day prior to and during training, with data being collected for a time span of at least 45 days (Ferguson & Rosales-Ruiz, 2001).
2. Two researchers wonder whether an eyewitness's memory of an event is affected by questions that he or she is asked subsequent to the event. To find out, the researchers show adults a film that depicts a car accident. Each adult is then asked one of five questions (randomly selected) about the accident:
   - About how fast were the cars going when they *contacted* each other?
   - About how fast were the cars going when they *hit* each other?
   - About how fast were the cars going when they *bumped into* each other?
   - About how fast were the cars going when they *collided into* each other?
   - About how fast were the cars going when they *smashed into* each other?

   The researchers compute the average speed given in response to each of the five questions to determine whether the questions have influenced participants' "memory" for the accident (Loftus & Palmer, 1974).
3. A researcher studies the effects of two different kinds of note-taking training (one of which is a placebo) on the kinds of notes that college students take. Her sample consists of students enrolled in two sections of an undergraduate course in educational psychology; with the flip of a coin, she randomly determines which section will be the treatment group and which will be the control group. She analyzes the content of students' class notes both before and after the training, making the prediction that the two

groups' notes will be similar before the training but qualitatively different after the training (Jackson, 1996).

4. At the request of the National Park Service, two researchers at Rocky Mountain National Park investigate the degree to which signs along hiking trails might influence hikers' behaviors. Park Service officials are concerned that the heavy traffic on one particular hiking trail, the trail to Emerald Lake, may be having a negative impact on the local environment; they would like to divert some traffic to a lesser-used trail to Lake Haiyaha, which begins at the same place as the Emerald Lake trail. One day in early summer, the researchers hide battery-operated, optic counters at key locations along the two trails to record the number of hikers. The study has four phases: (1) at the spot where the two trails originate, only signs indicating the destinations of the two trails are present; (2) a "positively worded" sign is added that describes the attractive features of the Lake Haiyaha trail and encourages hikers to use it; (3) the positively worded sign is replaced by a "negatively worded" sign that describes the crowdedness of the Emerald Lake trail and discourages its use; and (4) both the positively worded and negatively worded signs are posted. The researchers compare the frequency of hikers during each of the four phases (R. K. Ormrod & Trahan, 1982).
5. A team of researchers has a sample of elementary school boys, some of whom have been identified as having attention-deficit hyperactivity disorder (ADHD) and some of whom have not. One of the researchers asks each boy to interpret several social situations that are depicted in a series of black-and-white drawings (e.g., one sequence of drawings shows a sequence of events at a Halloween party). Some of the situations involve antisocial behavior (e.g., aggression), and other situations involve prosocial behavior (e.g., sharing). The researchers compare the interpretations that boys with ADHD make with the interpretations that boys without ADHD make with respect to both kinds of situations (Milch-Reich, Campbell, Pelham, Connelly, & Geva, 1999).

A Sample Dissertation

To illustrate how an experimental study might appear in its written form, we present excerpts from Virginia Kinnick's doctoral dissertation conducted at the University of Colorado (Kinnick, 1989). The researcher, a faculty member in the School of Nursing at another university, had considerable experience teaching nursing students the knowledge and skills they would need when working with women who were in the process of delivering a baby, and her interest lay in learning more about teaching such knowledge and skills effectively.

During a woman's labor prior to the delivery of her baby, a fetal monitor is often used to assess the baby's heart rate, and the maternity nurse must frequently check the monitor for signs that the baby might be experiencing exceptional and potentially harmful stress. The researcher wanted to determine whether a particular method of teaching concepts (one described by Tennyson and Cocchiarella) might be more effective for teaching fetal monitoring skills than the method traditionally used in nursing education programs. In her own words, the researcher's problem statement was as follows:

> This study is designed to determine if use of an instructional design model for concept attainment in teaching the critical concepts related to fetal monitoring will make a significant difference in preparation of nursing students in this skill, compared to the traditional teaching method which exists in most schools. (Kinnick, 1989, p. 8)

The research design is not one of the designs we have specifically described in this chapter. Instead, it involves administering three different instructional methods to three treatment groups (with participants assigned randomly to groups) and then observing the effects of the treatments at two different times: once immediately after instruction and then later after

students had completed the clinical rotation portion of their nursing program. Thus, the design of the study was the following:

	Group	Time →		
Random Assignment	Group 1	Tx1	Obs	Obs
	Group 2	Tx2	Obs	Obs
	Group 3	Tx3	Obs	Obs

In the following pages, we present excerpts from the methodology chapter of the researcher's dissertation. Our comments and observations appear on the right-hand side.

Dissertation ANALYSIS 6

METHODOLOGY

[After an introductory paragraph outlining the chapter's contents, the author describes the sample—students enrolled in maternity nursing courses at two universities—used in the study. Then, as she begins a discussion of her procedure, she explains that the experimental treatments were based on the Tennyson-Cocchiarella concept-teaching model (1986) and presents the key elements of the model. We pick up the methodology chapter at the point where the author describes the specific treatments used for each of the three treatment groups.]

Description of the Treatment Groups

[The author first explains that, for each of the three groups, treatment consisted of instruction in the basic concepts of fetal monitoring, plus additional instructional strategies, or "teaching variables," that differed for the groups.] . . . Starting with a basic class and adding new teaching variables to each treatment group, however, did require additional time. The length of time required for teaching the three treatment groups varied between 1 and 2 hours. These timeframes were established based on the results of the survey of baccalaureate nursing schools, in which 36% of the schools responding had less than 1 hour to teach fetal monitoring theory, and 52% had 1 to 2 hours (Kinnick, 1989).

The teaching variables for the first treatment group included labels and definitions, and presentation of best examples. According to Merrill and Tennyson, these variables usually include additional information needed to aid in the clarification and understanding of the concepts (Merrill & Tennyson, 1977, p. 100). Therefore, the design of this didactic presentation began with a very basic overview of physiology at the uterofetoplacental unit. Electronic fetal monitoring patterns are a reflection of uterofetoplacental physiology. Understanding the normal physiology and changes in the physiology that cause inadequate fetal oxygenation help the learner to identify the various patterns, and whether patterns are normal or abnormal. Understanding the physiology is also the basis to identifying appropriate nursing intervention which promotes normal physiology (reduction or even elimination of fetal distress) when abnormal patterns occur.

Comments

The author points out a possible confounding variable *in her study: The three forms of instruction took varying amounts of time.*

The survey to which the author refers was administered during a pilot study that she conducted prior to conducting the dissertation itself. She published the pilot study as a research article, which she cites here.

In this and subsequent paragraphs the author describes the treatment used for each treatment group; in a later "Procedure" section, she describes the general procedure she used to conduct the study. More often, a researcher will include a description of how each group was treated within the procedure section itself. Either approach is acceptable, however, as long as the writer makes the organization of the methodology section clear (e.g., through headings and subheadings).

When the classes were taught, the majority of students did not have any theory about the process of labor and delivery. In addition, they had not seen a fetal monitor. Methods of monitoring the fetus and a brief description and discussion of external versus internal monitoring, therefore, needed to be discussed. In addition, it was necessary to show the students a print-out of a fetal monitor as well as explain what the graphs meant. Before the basic concepts related to interpretation of the fetal heart could be taught, the student also needed to recognize critical characteristics of a contraction pattern as seen on a monitor strip. Contraction patterns can be a cause of physiological changes at the uterofetoplacental site. After these areas had been covered, the concept label, definitions, and best examples were presented . . . .

This description of what most students knew (and did not know) before instruction gives the reader greater confidence that the results observed after instruction (i.e., students' test performance) were probably due to the instructional treatments, rather than to any earlier learning experiences that the students may have had.

This 1 hour presentation included labels, definitions, best examples, and clarifying information. In the experience of this researcher, this presentation reflects closely the method for teaching fetal monitoring used in most schools of nursing, especially when the allocated time for teaching this content is limited. This treatment group is referred to as Group 1 throughout the study.

Notice that the author's notion of what is "traditional" instruction is based on her own experiences, and she says so here.

The second treatment group began with the same presentation used with the first treatment group, plus the addition of expository presentations for each major concept. An expository presentation was added after the labels, definition, and best examples of each set of coordinate concepts had been completed. For example, following the definition and display of the best examples of baseline fetal heart rate and its coordinate concepts, an expository presentation was done of the coordinate concepts. When that was completed, the concept of baseline variability was introduced and the same order of teaching variables was used. The addition of the expository presentations added approximately half an hour, so that this treatment group was scheduled for one and one-half hours. This group (labels, definitions, best examples and expository presentation) is referred to as Group 2.

After describing Group 1, the author proceeds to descriptions of Group 2 and then Group 3 in a logical and systematic fashion. The use of three subheadings (something along the lines of Treatment for Group 1 *or* Group 1 Instruction*) might have been helpful, however.*

By "expository presentation," the author means giving a short explanation or lecture about important ideas and concepts.

The design in Group 2 was chosen based on the results of Dunn's research (1984) on concept learning with college age students. . . . *[The author briefly describes Dunn's findings and their relevance for the instruction presented to Group 2.]*

A rationale for a particular experimental treatment strengthens any research report. A brief rationale can easily be incorporated into the description of procedures; a longer one should probably be presented earlier in the research report.

The treatment design for the third group used the same teaching variables as in Group 2, plus the addition of an interrogatory presentation to follow each expository presentation. This involved the addition of . . . transparencies specifically developed for the interrogatory presentation. When a fetal monitor pattern was shown on the screen, students were requested to compare it with their handout of definitions (list of critical characteristics) and best examples, and to identify the concept shown on the fetal monitoring pattern. This treatment design incorporated all of the teaching variables of the Tennyson-Cocchiarella concept-teaching model.

By "interrogatory presentation," the author means asking questions to assess students' understanding of, and ability to apply, what they have learned.

Development of the Instruments

[In this section, the author describes the tests she used to assess what participants knew about fetal monitoring following instruction, as well as a short questionnaire she used to determine the extent to which each participant knew something about fetal monitoring before instruction.]

Procedure

Prior to implementing this research, approval for the project was obtained from the Human Research Committee at the University of Colorado, and the Internal Review Board for Research at the University of Northern Colorado (Appendix E). The researcher then met with all students in each maternity nursing course during their first class to explain the research and ask their consent to participate. Consent forms were provided for each student (Appendix E). Once this process was completed, the research design was implemented.

Because the author conducted the study at two universities, she followed the necessary human research review procedures at both institutions.

Each maternity nursing course had three groups participating in the research. Students in each of the courses were randomly assigned to one of these three groups. One group received the instructional method described in the Tennyson-Cocchiarella model of concept attainment. A second group received the same instructional method with the exception of the interrogatory presentation. The third group had a didactic presentation using only labels, definitions, best examples and clarifying information. In other words, both the expository and interrogatory presentations were eliminated from the presentation for the third group. In both schools, the researcher taught all three methods. A script (or lecture) was developed for the researcher to use in all the treatment groups so that the content was the same in each group (Appendix F). The students were tested in a class session within 2 to 3 days following the class (treatment).

As noted earlier in Chapter 9, random assignment is one effective way of ruling out the possible effects of confounding variables.

The first group mentioned here ("one group") is actually Group 3, and the last ("the third group") is actually Group 1; this reversal might cause confusion for the reader.

The use of a "script" here would help the researcher teach the content similarly for all three treatment groups (except, of course, for the things she intentionally wanted to do differently for the three groups). Thus, it should help minimize any influences the researcher's hypotheses might have on her delivery of different instructional methods.

After the completion of the clinical experience of all groups in each university, a parallel form of the classification test was again administered. The sequence can be summarized as follows:

Class instruction → Posttest → Clinical Rotation → Delayed Test upon Completion of Clinical Rotation

This graphic display of the procedure used is a helpful summary for the reader.

In addition, each student was requested to keep a record of the number of contacts each of them had with fetal monitoring tracings, the context, and type of pattern (Appendix G). For example, the student may have been assigned to a labor patient who had a normal pattern. The contact, however, could have been in clinical conference where actual monitor strips of patients were discussed, or also in a prenatal clinic where a nonstress test was done on a patient. The purpose of keeping these records [was] to identify the number of interrogatory examples the students encountered clinically and the range of examples. This information [could] be compared with the post test results.

The author presumably asked students to keep such records as a way of helping her interpret any unexpected results related to the delayed (postclinical rotation) test. Keep in mind, however, that such self-reporting techniques, dependent as they are on participants' diligence and memories, will not always yield totally accurate information.

Ideally, none of the students were to have had any contact in the clinical setting before the instruction and first test were done. However, it was impossible to schedule all three treatments before students in each maternity nursing course were assigned to the clinical setting since they began their clinical experiences the second week of classes. A few students in this situation were assigned to patients with fetal monitors attached. Since they did not have any theory on fetal monitoring, they were not responsible for interpretation of fetal monitor patterns. However, staff nurses and/or clinical instructors may have demonstrated how to attach and detach the equipment and talked about tracings seen by each student on their individual patients.

Here the author points out a potential weakness in her study: Some students had additional exposure to fetal monitoring outside of the instruction she had given them in their respective treatment groups. The exposure was apparently minimal, however, and so probably did not jeopardize the quality of her study. Such honesty is essential *in any research report.*

Statistical Analysis

[The author continues with a discussion of the statistical analyses she used to compare the performance of the three groups.]

NOTE: Excerpt is from *Learning Fetal Monitoring Under Three Conditions of Concept Teaching* (pp. 58–69) by V. Kinnick, 1989, unpublished doctoral dissertation, University of Colorado, Boulder. Reprinted with permission.

For Further Reading

Antony, J. (2003). *Design of experiments for engineers and scientists.* Oxford, England: Butterworth-Heineman/Elsevier.

Barlow, D. H., Nock, M. K., & Hersen, M. (2008). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston: Allyn & Bacon.

Bausell, R. B. (1994). *Conducting meaningful experiments: Forty steps to becoming a scientist.* Thousand Oaks, CA: Sage.

Canavos, G. C., & Koutrouvelis, I. A. (2009). *An introduction to the design and analysis of experiments.* Upper Saddle River, NJ: Prentice Hall/Pearson Education.

Deming, S. N., & Morgan, S. L. (1993). *Experimental design* (2nd ed.). New York: Elsevier Science.

Friedman, D., & Sunder, S. (1994). *Experimental methods: A primer for economists.* New York: Cambridge University Press.

Glass, D. J. (2007). *Experimental design for biologists.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Glass, G. V. (1988). Quasi experiments: The case of interrupted time series. In R. M. Jaeger (Ed.), *Complementary methods for research in education* (pp. 445–464). Washington, DC: American Educational Research Association.

Heath, D. (1995). *An introduction to experimental design and statistics for biology.* Bristol, PA: Taylor & Francis.

Kirk, R. E. (1994). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Montgomery, D. C. (2009). *Design and analysis of experiments* (7th ed.). New York: Wiley.

Morgan, D. L., & Morgan, R. K. (2001). Single participant research design: Bringing science to managed care. *American Psychologist, 56,* 119–127.

Phillips, D. C. (1981). Toward an evaluation of the experiment in educational contexts. *Educational Researcher, 10*(6), 13–20.

Pukelsheim, F. (2006). *Optimal design of experiments (Classics in applied mathematics).* New York: Wiley.

Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists.* Cambridge, England: Cambridge University Press.

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs.* Washington, DC: American Educational Research Association.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Solso, R. S., & MacLin, K. (2008). *Experimental psychology: A case approach* (8th ed.). Boston: Pearson/Allyn & Bacon.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Identifying Research Designs" on page 251:

1. This is a *quasi-experimental* study. In particular, it involves a *multiple baseline* design: Each of the horses begins training on a different day. In the section of the chapter "Using Designs 11, 12, and 13 in Single-Subject Studies," a multiple baseline study involving one 2-year-old girl is described. In this example, however, we see the approach being used with five horses, each of which is treated identically except for the date on which training begins.
2. This is an *experimental* study in which the researchers randomly assign participants to one of five groups, each of which is asked a different question.
3. Don't let the random selection of treatment and control groups fool you. This is a *quasi-experimental* study because the participants are not randomly assigned as *individuals* to the treatment and control groups. More specifically, the study is a *nonrandomized control group pretest–posttest* design (Design 8).
4. This, too, is a *quasi-experimental* study. It is a *time-series* design in which the effects of no intervention (Phase 1) are compared to the effects of two different interventions (the two new signs) imposed either singly or in combination. Of the designs described in this

chapter, it is probably most similar to Design 12. Note, however, that no phase of the study is repeated; this omission is a decided weakness in the design.

5. This is a *combined experimental and ex post facto factorial* design with two independent variables, one of which is a *within-subjects* variable. One independent variable is the presence or absence of attention-deficit hyperactivity disorder, which the researchers do not (and *cannot)* manipulate; this is the ex post facto component of the design. The other independent variable is the content of the drawings (aggression vs. prosocial behavior); this is the experimental, within-subjects component of the design.

Now go to MyEducationLab at www.myeducationlab.com to take assessments to evaluate your mastery of chapter concepts. Assignments and Activities exercises and Building Skills in reading Research exercises are also available to help you master concepts and skills. Feedback for many of these exercises is provided so that you can see why your answers are correct or incorrect.