

# 9

## True Experiments, Part 1: Single-Factor Designs

**A**s we saw in Chapter 6, a key concept in designing experiments is that of control. The experimenter seeks to control as many of the potential threats to validity as possible. When a sufficient number of these are under control, the study is a true experiment. A true experiment is one in which the experimenter has reason to believe that he or she has control over both the assignment of subjects to conditions and the presentation of conditions to subjects. When a study does not meet the requirements of a true experiment, it is called a quasi experiment. The word *quasi* means *as if* or *to a degree*. Thus, a quasi experiment is one that resembles an experiment but lacks at least one of its defining characteristics.

We will begin this chapter by defining true experiments and quasi experiments and discussing their differences. Then we will define the basic elements of a valid experimental design, mention some designs that should be avoided, and finally, describe some representative experimental designs.

### TRUE EXPERIMENTS VERSUS QUASI EXPERIMENTS

**true experiment**  
research procedure  
in which the scien-  
tist has complete  
control over all  
aspects

In a **true experiment**, the experimenter has complete control over the experiment: the who, what, when, where, and how. Control over the who of the experiment means that the experimenter can assign subjects to conditions randomly. Recall that random assignment is preferred because it allows one to conclude that any other variable could be confounded with the

**quasi experiment**  
 research procedure  
 in which the scien-  
 tist must select sub-  
 jects for different  
 conditions from  
 preexisting groups

independent variable only by chance. No other method of assignment of subjects to conditions permits such a conclusion. Control over the what, when, where, and how of the experiment means that the experimenter has complete control over the way the experiment is to be conducted.

A **quasi experiment**, on the other hand, is an experiment in which the investigator lacks the degree of control over the conditions that is possible in a true experiment. The most important difference is that whereas it is possible to *assign* subjects to conditions in a true experiment, in a quasi experiment it is necessary to *select* subjects for the different conditions from previously existing groups.

For example, you may wish to study the effect of number of food pellets on the rate at which rats learn a maze. This situation would permit the design of a true experiment, because you could arbitrarily assign some rats to the large-reward condition and others to the small-reward condition. Assume that before the experiment the rats belonged to a homogeneous population of rats. For experimental purposes, you assign the rats to groups that you create according to your needs.

On the other hand, if you were interested in sex differences in detecting hidden figures, you would have to conduct a quasi experiment, because you cannot assign subjects to the two conditions, male and female. Here the researcher cannot create groups of males and females but instead selects members from preexisting groups.

Quasi experiments are sometimes called *ex post facto*, or after the fact, experiments because the experiment is conducted after the groups have been formed. In the case of a quasi experiment with sex as the independent variable, the experiment takes place long after the subjects become males or females. If you performed an experiment using preexisting classes of students, the two classes would be an *ex post facto* variable because the classes were formed before you did the experiment.

Another way to look at the difference between true experiments and quasi experiments is to note that in true experiments we *manipulate* variables, whereas in quasi experiments we *observe* categories of subjects. When we take two preexisting groups and consider a difference between them to be the independent variable, we are not manipulating a variable but simply labeling groups according to what we think is the important difference between them. The true difference between them for our experiment may be quite different from what we think it is. If we find that two different socioeconomic groups differ on some measure, the difference may be caused not by the socioeconomic difference itself but by cultural differences between the two socioeconomic groups. By calling the difference socioeconomic, we may obscure the fact that the difference is actually a difference in need achievement or perceived helplessness or religion.

When we present some independent variable to two preexisting groups, more is involved than measuring their behavior. We have the additional

problem of not knowing whether the difference in behavior was caused by differences between the groups or by the independent variable. If we studied the effects of two different teaching methods on learning in two preexisting classes, we would not be sure whether any differences in learning resulted from the teaching methods or from preexisting differences between the classes.

A true experiment, then, permits the experimenter the greatest degree of control in ruling out alternative hypotheses, or alternative independent variables, as being the cause of the difference between two groups or conditions. This permits the most powerful use of Mill's method of differences (discussed in Chapter 2, and also in Chapter 5), because all other potential independent variables have been eliminated by randomly assigning subjects to conditions. A quasi experiment leaves open the possibility that other differences exist between the experimental and control conditions, and thus permits other potential differences to remain.

It is possible to have one experimental variable and one quasi-experimental variable in an experiment. For example, in studying the effects of two different teaching methods on classroom learning, we might be interested in whether slow learners differ from fast learners in their response to the teaching methods. The two teaching methods would constitute a true experimental variable, assuming we assigned students to sections, and the classification into slow and fast learners would constitute a quasi-experimental variable.

The degree of control that is possible varies from one type of investigation to another. The fact that we are discussing true experiments first does not mean that true experiments are necessarily better than other types of investigations. Rather, it reflects the fact that as the degree of control that the researcher can exercise decreases, the threats to the validity of the conclusions increase. Other things being equal, one would choose a true experiment over a quasi experiment and a quasi experiment over a nonexperimental method. Things are seldom equal, though. Many social-psychological phenomena are difficult to bring into the laboratory in a realistic fashion. Therefore, a field study may be preferable to an experiment, because the advantage of realism outweighs the loss of control. Nevertheless, you should try first to design a true experimental study and use the other designs only when you believe that the gain in validity will be worth the loss of control.

## **FACTORS, LEVELS, CONDITIONS, AND TREATMENTS**

Up to this point, we have simply used the term *independent variable* to talk about what the experimenter manipulates in an experiment. Now we need to introduce some other terms that are often used in discussing independent variables.

**factors**

the independent variables of an experiment

The independent variables of an experiment are often called the **factors** of the experiment. Suppose we are doing an experiment on the effect of human handling and of the type of cage in which they are raised on the emotionality of rats. Rats would either be handled or not handled and raised either in a standard laboratory cage or in a large cage with lots of added items for the rats to play with. We would say that there were two independent variables: handling and type of cage. Handling and cage type, therefore, would be the two factors of this experiment. An experiment always has at least one factor, or independent variable; otherwise it wouldn't be an experiment! In order to have an experiment, it is necessary to vary some independent variable, or some factor.

**level**

in an experiment, a particular value of an independent variable

Each of the two independent variables in our example would have two **levels**. A level is a particular value of an independent variable. An independent variable always has at least two levels, because if it didn't it wouldn't be a variable! The two levels of handling in this example would be handling versus no handling, and the two levels of cage type would be normal versus enriched. It is possible for an independent variable to have any number of levels, of course: The rats could be handled either 0, 10, 20, 30, or any other number of minutes per day, giving that many different levels of the independent variable of handling.

**condition**

a group or treatment in an experiment

The term **condition** is the broadest of the terms used to discuss independent variables. It refers to a particular way in which subjects are treated. In a between-subjects experiment, such as the present example, the experimental conditions are the same as the groups. Any one rat in this example was either handled or not handled, and was in one of two cage types. In other words, each rat experienced only one treatment or condition. In the present example, we might speak of a particular rat as being in the enriched, nonhandled condition or group. When it is possible for each subject in an experiment to experience every condition, then we speak only of various conditions, not groups, because there is only one group of subjects and that group experiences all conditions. There would be as many conditions as there are different ways in which subjects are treated. Thus, there could be many conditions in a complicated experiment.

**treatment**

another word for a condition of an experiment

**Treatment** is just another word for condition. You should be aware of this usage, however, as you may run across the term *treatment effect* in statistics. It is a statistical test of the effect of various conditions of the experiment.

## SOME DESIGNS TO AVOID

In this chapter we will consider a number of examples of designs that are appropriate to the particular problem they address. However, before we consider good designs, let us look at some designs that should be avoided. These undesirable designs are weak in that they do not control for various

	TREATMENT	TEST
Single group	Yes	Yes

alternative explanations of the results. They all fail to control for certain threats to validity discussed in Chapter 5.

## ■ The One-Group Posttest-Only Design

### **one-group posttest-only design**

research design that measures the behavior of a single group of subjects after the treatment only

The **one-group posttest-only design** is a simple one in which a group of subjects is given a treatment and then tested on some dependent variable (see Table 9.1).

Suppose you wanted to test the effectiveness of est (Erhard Seminars Training), a program in which people attend a series of lectures and group activities, some of which are humiliating and exhausting. In order to evaluate the effect of the training, you decide to survey the participants. You find that most of them say that the experience was worthwhile and that they feel better about themselves than they did before the training. After a little reflection, you realize that the results of your survey are nearly worthless. Although the people report that they feel better than they did before the training, you have no measures of how they felt before. Therefore, you cannot determine if they changed. Furthermore, even if they did change, you have no assurance that the training itself caused this change. Perhaps the interruption in their routine caused the change. Or perhaps they would have felt better if they had gone to the movies for several evenings and then gone camping for a weekend.

Such a one-group posttest-only design leaves so many threats to validity uncontrolled that it is nearly worthless. Nevertheless, you can certainly recall people who have recommended a product or practice to you on the basis of their experience. Many people in everyday life, as well as some scientists, have used this design.

## ■ The Posttest-Only Design with Nonequivalent Control Groups

Suppose that you wanted to improve on the study of the effectiveness of est by comparing people who had taken the training with a control group who had not. You might try to find a group of people who matched the est group on as many variables as possible: age, income, education, and so forth. Table 9.2 illustrates this design. Although the design is an improvement over the one-group posttest-only design, it still has a serious flaw. The flaw is that the

**TABLE 9.2** POSTTEST-ONLY DESIGN WITH NONEQUIVALENT CONTROL GROUPS

	ALLOCATION OF SUBJECTS AND GROUPS	TREATMENT	TEST
Group 1	Any method that is <i>not</i> random	Yes (or $A_1$ )	Yes
Group 2		No (or $A_2$ )	Yes

**nonequivalent control group**

a group of subjects that is not randomly selected from the same population as the experimental group

control group is not equivalent in every way to those who took the training. The most important difference is that the test persons had selected themselves for the training and the control groups had not. Thus, we have a **nonequivalent control group** because the two groups were not randomly constituted from the same population.

A nonequivalent control group is better than no control group, but you would have to consider this study a quasi experiment at best, because the subjects were not randomly assigned to groups. The only way you could construct a control group that was equivalent to the est group would be to ask the est organization to provide a list of all people who applied for the training. Then you would randomly place half of them into a control group that would not be allowed to take the training.

## ■ The One-Group Pretest/Posttest Design

**one-group pretest/posttest design**

research design that measures the behavior of a single group of subjects both before and after treatment

Another way of improving on the one-group posttest-only design is to take a measure of behavior before the treatment that can be compared with behavior after the treatment. This approach is called the **one-group pretest/posttest design**. In the example of the est study, you might obtain responses of the participants before they took the training to compare with responses after training. Such a design is illustrated in Table 9.3. If you were to use this design in the est study, you would probably find a change in the subjects' reports of their moods, feelings of self-worth, and so forth. You would still have the problem of determining what caused the changes—the est training or some unrelated event. Even if the est procedure did cause the changes, you wouldn't know what aspect of it was responsible. You still wouldn't know if going to the movies for several nights followed by a camping trip might have been equally beneficial. The following example discusses these problems in the context of a different situation.

Suppose a company introduced a new work schedule whereby its employees put in four 10-hour days a week instead of five 8-hour days. If output increased, management would probably credit the new schedule. This conclusion represents an improvement over the one-group posttest-only design, because you know that a behavior change did follow the treatment. However, you have not considered other potential causes of the increase in output, and hence other threats to validity. Workers may have responded to

	PRETEST	TREATMENT	POSTTEST
Single group	Yes	Yes	Yes

the attention paid them by management when the change was initiated. Or any number of events may have led to increased productivity: favorable weather conditions that allowed the workers to get to work on time, a change in seasons that made the plant more comfortable, or a favorable response to a new supervisor. These occurrences represent threats to internal validity: The change was caused by a variable other than the one management thought to be responsible. Threats to external validity could arise from the possibility that these workers are young and like long hours, whereas older workers might have preferred shorter days.

This study would have been better designed by forming two groups through random allocation of workers to different schedules so that one group would remain on the old schedule as a control group. This control would have eliminated the threats to internal validity. Random allocation of workers to two groups may not be possible, however. In that event, if the company had two plants, one could be switched to the four-day week while the other is kept on the five-day week. Productivity in the two plants could then be compared. This example is a nonequivalent-control-group design. Differences between the workers at the two plants or in the plants themselves may account for the results instead of the work schedule. The addition of a nonequivalent control group to a pretest/posttest design improves the control sufficiently that the design may be considered a quasi experiment. Chapter 12 will discuss a number of quasi-experimental designs.

## THE BASIC ELEMENTS OF A VALID EXPERIMENTAL DESIGN

Chapter 5 discussed types of validity and the many threats to validity that exist, and Chapter 6 discussed methods of control that are available to improve the validity of an experiment. In this chapter, we begin to consider some specific experimental designs as examples of ways of achieving control over threats to validity. You should keep in mind a point made previously, that designing an experiment is an exercise in problem solving. When threats to validity are adequately controlled for, the experiment has been designed. At the same time you should realize that no design can rule out all threats to validity for all time. For example, as we said in Chapter 5, societal changes since the “dirty word” study was conducted have reduced the external validity of that experiment.

	ALLOCATION	TREATMENT	TEST
Condition 1	<i>Either random allocation of subjects to conditions or all subjects experience both conditions</i>	Yes (or $A_1$ )	Yes
Condition 2		No (or $A_2$ )	Yes

Even though there can be no perfect experiment, two particular elements of design provide control over so many different threats to validity that they are basic to all good experimental designs. They are (1) the existence of a control group or a control condition and (2) the random allocation of subjects to groups. Both of these methods of control were discussed in Chapter 6. (If the experiment is a within-subjects design, each subject experiences all conditions, so random allocation of subjects to conditions is not applicable. In such experiments, the subjects should experience the conditions either in random order or in counterbalanced order.)

These two basic elements of good experimental design are illustrated in Table 9.4, which represents a simple experiment with two conditions. If this is taken to represent a between-subjects experiment, then different subjects would be randomly allocated to the two conditions. This allocation assures that the groups will be equal in all respects, except as they may differ by chance. If this is a within-subjects experiment, then all subjects experience both conditions: A subject's behavior in one condition is compared with his or her behavior in another condition. Either way, we have reason to believe that the subjects in both conditions were equal to begin with, which enables us to compare their performance between experimental and control conditions. Any difference in behavior can be attributed to differences between the two conditions.

## WITHIN-SUBJECTS DESIGNS

Recall that Chapter 6 discussed, as a specific strategy of achieving control, using a subject as his or her own control. Recall also that this strategy is desirable when the effect of one condition will not carry over to, or contaminate, the other condition or conditions of the experiment to a serious degree. The first group of designs we will discuss in this chapter makes use of this strategy. Because the same subjects experience all conditions in within-subjects designs, it is often necessary to exercise some ingenuity in controlling for possible carry-over effects.

## ■ Controlling for Order and Sequence Effects

In within-subjects experiments, because a subject experiences more than one experimental condition, the possibility exists that some variable may influence the data as a result of the repeated testing. Some of these possible variables are related to the subjects; others are related to the conditions of testing. The subjects may get fatigued during the session, or the first condition may be tested before lunch when subjects are hungry and the second after lunch when they are sleepy. Ordinarily, experimenters avoid within-subjects designs if they believe that order or sequence effects will be substantial. Then a between-subjects design is probably more appropriate.

Before we discuss ways of controlling for these problems, let us note the distinction between order effects and sequence effects. **Order effects** are those that result from the (ordinal) position in which the condition appears in an experiment, regardless of the specific condition that is experienced. The best example of an order effect is the warm-up or practice effect that often occurs in experiments on learning. Whichever condition is presented first will show poorer performance than later conditions simply because the subjects have not warmed up to the task. **Sequence effects**, on the other hand, depend on an interaction between the specific conditions of the experiment. For example, in an experiment on judging the heaviness of lifted weights, there is likely to be a contrast effect such that a light weight will feel even lighter if it follows a heavy one, and vice versa. Order effects are more general and result from warm-up, learning, fatigue, and the like. Sequence effects are interactions among the conditions themselves.

The difference between order and sequence effects can be seen by referring ahead to page 231, where the six possible ways of ordering three different conditions, A, B, and C, are presented. Note that Subjects 1 and 3 experience condition C in the same ordinal position, namely third, and so have the same order effect for C. Subjects 2 and 3, however, both experience condition C following A, and so have the same sequence effect for C (with respect to A).

In general, one controls for order effects by arranging that each condition occur equally often in each ordinal position—first, second, third, and so on. Sequence effects are generally controlled for by arranging that each condition follow every other condition equally often. Note that these controls are not the same. You should also note that the various methods of controlling for order and sequence effects are effective only under certain conditions, which we will discuss shortly. However, if you are using a within-subjects design, you are wise to control as effectively as possible for order and sequence effects.

Two basic strategies are available for controlling order and sequence effects. The first, and preferable, one is to arrange the order of conditions in such a way that order and sequence effects are controlled *within subjects*. When this is not possible, you must control for order and sequence effects between subjects.

### **order effects**

changes in a subject's performance resulting from the position in which a condition appears in an experiment

### **sequence effects**

changes in a subject's performance resulting from interactions among the conditions themselves

## Within-Subjects Control of Order and Sequence

Controlling for order and sequence effects within subjects is possible when each subject receives each condition. Randomization can be used when each condition is given several times to each subject or when a sufficient number of subjects will be tested so that one particular sequence is unlikely to have much influence on the outcome. Experiments in learning or perception typically involve presenting each stimulus many times to the subject. The best procedure is to randomize the order of conditions for each subject. Although you might prefer to be told a magic number of subjects or repetitions that are sufficient for randomizing to be effective, this determination remains a matter for your judgment.

A useful variation on randomizing to control for order and sequence effects is **block randomization**. Block randomization means that the order of conditions is randomized, with the restriction that each condition is presented once before any condition is repeated. If there are four conditions and each one is to be represented twice, block randomization might give you the following sequence: BCAD, ADCB. Here each of the four conditions is presented once in random order within each of two blocks. Thus there is less chance that unwanted sequence effects would be produced by orders of the following type: AABDBCCD. Block randomization is particularly useful if you want to present each condition twice and your experiment requires two sessions.

When relatively few subjects will be tested and you have several conditions that can be presented only a few times, you must begin to exercise ingenuity. A typical example is the instance in which you have three conditions, each presented twice. In this situation, it is common to use **reverse counterbalancing** to control for order effects. The three conditions are presented in order the first time and then in reverse order. This technique is known as ABCCBA sequence, or ABBA for short. Counterbalancing works well when you suspect that the possible confounding variables will act in a linear manner over conditions. Figure 9.1 gives an example of ABCCBA order in which there is a linear effect. The order effect produces a large increase in the dependent variable that, since the effect is linear, averages out. The counterbalancing has done its job.

On the other hand, suppose a variable has a large effect in the early part of the experiment but a smaller effect later on. The best example is a warm-up, or practice, effect that may occur in the early part of an experiment and be less important later. Figure 9.2 shows an example in which there is a large practice effect. Here you can easily see that counterbalancing has not been effective in eliminating the order effect. One way of improving such an experiment is to provide enough practice beforehand that the practice effect is eliminated.

Note that counterbalancing may do an incomplete job of controlling sequence effects in an ABCDDCBA experiment: The B condition follows A once and C once but never follows D or itself.

**block randomization**  
control procedure in which the order of conditions is randomized, but with each condition being presented once before any condition is repeated

**reverse counterbalancing**  
method of control in which conditions are presented in order the first time and then in reverse order

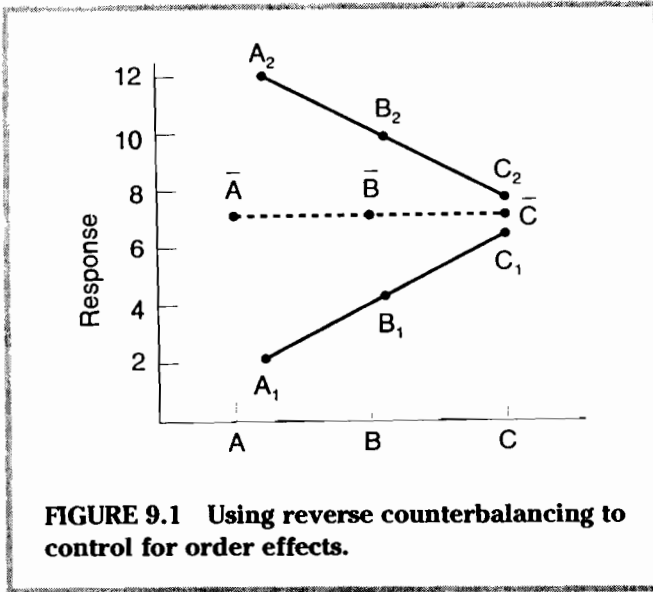


FIGURE 9.1 Using reverse counterbalancing to control for order effects.

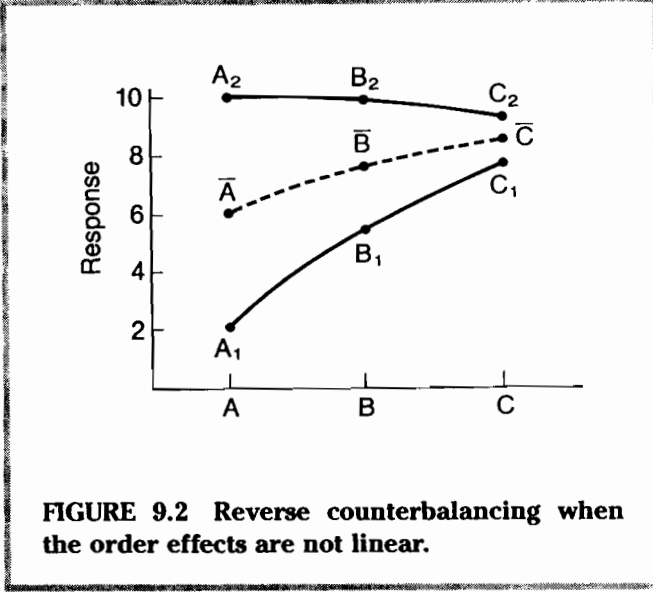


FIGURE 9.2 Reverse counterbalancing when the order effects are not linear.

**Within-Groups Control of Order and Sequence**

If presenting each condition enough times to randomize the order is not possible, or if counterbalancing within subjects does not seem appropriate, you must leave order and sequence confounded with condition *within subjects* and you must control for order and sequence *within groups*. For example, suppose you have three conditions and each one is to be presented only once to each subject. Then if you have six subjects, or 6*N* subjects, you can control for order and sequence in the following way:

SUBJECT	ORDER
1	ABC
2	ACB
3	BAC
4	BCA
5	CAB
6	CBA

Both order and sequence are completely counterbalanced within the group, because each condition occurs an equal number of times in each rank-order position and follows every other condition an equal number of times. Thus, you have controlled for order and sequence within a group of subjects, even though every subject experiences a biased sequence. The disadvantage of this method of counterbalancing is that as the number of conditions increases, the number of orders required increases geometrically. You have 2 possible orders of 2 conditions (AB and BA), 6 orders of 3 conditions (as in the previous example), 24 orders of 4 conditions, and 120

orders of 5 conditions! Even for only 4 conditions you would need 24 subjects to control for order and sequence.

It is possible to control for the order in which each condition occurs with fewer subjects than would be required by complete counterbalancing if you give up the requirement that each condition follow every other condition an equal number of times. You would be controlling for order, but not for sequence, of conditions. This type of incomplete counterbalancing is called the **Latin square** technique, after the ancient puzzle of finding ways to arrange a number of letters in a matrix such that each letter occurs once and only once in each row and column:

**Latin square**

control procedure in which each subject experiences each condition in a different order from other subjects

SUBJECT	RANK ORDER			
	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	C

If the letters represent conditions, the columns represent order, and the rows represent subjects, you are controlling for order effects with the Latin square counterbalancing technique.

A disadvantage of the Latin square technique is that sequence is not controlled for. Notice in the previous example that B always follows A, C always follows B, and so forth. Thus sequence is always perfectly confounded with order in this particular Latin square. If there were a contrast effect between conditions, this design would not control for it. However, you can control for sequence effects of the immediately preceding condition by using particular sets of Latin squares known as *balanced squares*. In the balanced Latin square, each condition is immediately preceded once by every other condition (W.A. Wagenaar, 1969), as in the following example:

SUBJECT	RANK ORDER			
	1	2	3	4
1	A	B	C	D
2	B	D	A	C
3	C	A	D	B
4	D	C	B	A

When you can assume that the contrast effects are primarily between pairs of conditions, the balanced Latin square will be effective.

The advantage of the Latin square technique over complete counterbalancing is that it permits greater flexibility in choosing the number of subjects to be tested. Instead of needing 24 or 48 subjects in a four-condition

experiment, for example, you can use only 4 or 8. This advantage is great enough to outweigh the disadvantage of leaving small sequence effects uncontrolled.

We turn now to some representative within-subjects designs that offer typical solutions to the problems of validity and control discussed in Chapters 5 and 6. The designs presented here do not constitute all that are possible, because an indefinite number of designs exists. These designs are simply the most common solutions to common experimental problems.

## ■ Two Conditions, Tested within Subjects

### two-conditions design

the simplest research design, involving only two conditions

The **two-conditions design** is the simplest possible true-experiment design, because it has only two conditions and each subject serves as its own control. This design is illustrated in Table 9.5. The two conditions are labeled Condition 1 and Condition 2, although one of them may be considered the experimental condition and the other the control condition. All subjects experience both conditions in counterbalanced order. In spite of its simplicity, this design is not used as often as one might expect, for two reasons. First, many experiments involve more than two conditions. Second, there is the possibility of carry-over effects from one condition to the other.

An experiment that serves as an example of this design is one by John Marshall and Philip Teitelbaum (1974) on the phenomenon of sensory neglect. Research has shown that damage to both sides of a small part of the brain known as the lateral hypothalamus produces a severe impairment of feeding and drinking. In addition to these motivational effects, there is interference with the ability to respond to sensory stimulation after this damage. Marshall and Teitelbaum were interested in the contribution of the sensory neglect to the motivational deficit.

Recall that the brain is approximately bilaterally symmetrical; that is, the left half is the mirror image of the right. Marshall and Teitelbaum destroyed the lateral hypothalamus on one side of the brains of 12 rats. When

**TABLE 9.5** TWO-CONDITIONS DESIGN, TESTED WITHIN SUBJECTS

	ALLOCATION	TREATMENT	TEST
Condition 1 (or experimental)	All subjects experience both conditions in counterbalanced order	Condition 1 (or experimental treatment)	Yes
Condition 2 (or control)		Condition 2 (or control treatment)	Yes

the rats had recovered from the operation, they were tested for their responsiveness to stimuli that were presented to one side or the other of their bodies. Visual, tactile, and olfactory stimuli were used. In all cases, the rats responded only when the stimuli were presented to the same side that had received brain damage. (Each side of the brain controls the opposite side of the body.) When the rats were stimulated on the opposite side, they did not respond.

Thus, from the design point of view, the unilateral damage allowed each animal to be used as its own control. One side of the brain was a control for the other, ruling out differences between subjects as a possible source of error. For example, one of the tests used was whether the rat would attack a mouse presented to the brain-damaged side. (Some rats will kill mice, and others will not.) By using each rat as its own control, the experimenters were able to say that all rats that were mouse killers attacked mice presented to the undamaged side of the brain but did not react to those presented to the damaged side.

## ■ Multiple Conditions, Tested within Subjects

Psychology experiments generally employ more than two conditions. The first reason researchers choose a **multiple-conditions design** is that seldom do they wish to ask a simple yes-or-no question. Usually they want to compare several variables or treatments for effectiveness. For example, the question may be which of three different types of psychotherapy is most effective.

A second reason for conducting multiple-conditions experiments is to determine the shape of the function that relates the independent and dependent variables. When experimenters want to know how brightness increases with intensity of a light, they present each of several intensities of the light to a group of subjects. From the responses to the various intensities, they can plot the relation between intensity and brightness. Each intensity is a condition of the experiment.

A third reason for doing multiple-conditions experiments is the presence of more than one rival hypothesis that must be ruled out. Suppose a child has a favorite toy that is fuzzy, colorful, and noisy. If you want to find out which of the three characteristics is responsible for the child's attachment to the toy, you could make up three versions of the same toy, as follows:

- A FUZZY, not colorful, not noisy
- B not fuzzy, COLORFUL, not noisy
- C not fuzzy, not colorful, NOISY

Toy A tests for fuzziness, B for colorfulness, and C for noisiness as the independent variable causing attachment. Because toys A and B are not noisy, they control for noisiness. Similarly, A and C control for colorfulness, and B and C control for fuzziness. To accommodate the three hypotheses, you

### multiple-conditions design

research design that involves more than two conditions

would need three conditions (toys). Each toy serves as a partial control (condition) for the other hypotheses. In this example, we have varied three independent variables in a single experiment.

Most multiple-conditions experiments are between-subjects experiments, because it is often impossible or inappropriate to expose all subjects to the various conditions. Within-subjects experiments are fairly common, however.

One such experiment is a classic study by Fergus Craik and Endel Tulving (1975), which examined whether different strategies of processing words would affect memory. They flashed words on a screen. Before each word appeared, they asked the subject a question: "Is the word in capital letters?" or "Does the word rhyme with train?" or "Does the word fit in this sentence: 'The girl put the \_\_\_\_ on the table?'" Each of these questions was designed to induce the subjects to adopt a different strategy of processing the word. The first strategy focused on the visual properties of the word, the second on the acoustic properties, and the third on the semantic properties. Craik and Tulving theorized that each successive type of processing would induce greater "depth of processing." Their theory predicted that increasing depth of processing increases memory for words.

Each subject in their study experienced all three types of questions, making this a within-subjects design. The experimenters believed that subjects could adopt different strategies of processing on different trials. The various questions were the independent variables in the study. The questions were randomly varied for each trial. After the words were all presented, the experimenters unexpectedly gave the subjects a list that contained all of the words they had presented along with an equal number of words that they had not presented. They asked the subjects to indicate which words they recognized from the list. The percentage of words recognized varied as a function of the depth of processing induced by the questions. The subjects recognized only 18% of the visually processed words, but they recognized 78% of the acoustically processed and 96% of the semantically processed words.

Table 9.6 indicates schematically the general design of a multiple-conditions, within-subjects experiment.

	ALLOCATION	TREATMENT	TEST
Condition 1	All subjects experience all conditions, in either random or counterbalanced order	1	Yes
Condition 2		2	Yes
Condition 3		3	Yes

## BETWEEN-SUBJECTS DESIGNS

As indicated previously, there are many situations in which subjects cannot be used as their own controls because of the possibility of carry-over effects. As with within-subjects designs, between-subjects experiments may have two conditions, or more than two.

### ■ Two Conditions, Tested between Subjects

The experiment by Marshall and Teitelbaum described previously was a within-subjects experiment, because each rat served as its own control. This design may not be desirable when the possibility of large order or sequence effects is present. Such was the case in an experiment on the effect of anxiety on affiliation conducted by Stanley Schachter (1959). He hypothesized that inducing anxiety in people would cause an increase in their tendency to seek the company of others. He induced anxiety by telling subjects that they were to be connected to an apparatus that would deliver painful electric shock. After a lecture on the purpose of the research, subjects were instructed that they were to wait 10 minutes before receiving the shock. They filled out a questionnaire that asked whether they wished to wait alone or with other subjects. Subjects in the control condition were told they would receive mild, nonpainful shock. Otherwise, they were treated the same as the experimental subjects. Of the low-stress subjects, 33% indicated that they wished to wait with others, whereas 63% of the high-stress subjects preferred to wait with other subjects. Analysis of these data using chi-square is shown in Box 9.1.

You can see that this experiment had to be conducted as a between-subjects experiment. Once subjects had experienced one of the conditions, they would no longer be naive about the situation. In actuality, none of the subjects in either condition received shock. If they had been asked to serve again in the other condition, the instructions would not have had the same effect. Even if they had been shocked and then asked to participate again, they would have noticed the difference in instructions between conditions and would have begun to suspect the experiment's purpose. Only by having a separate group in each condition could Schachter test the effect of anxiety on the dependent variable.

### ■ Multiple Conditions, Tested between Subjects

The design of a multiple-conditions, between-subjects experiment is indicated in Table 9.7. An interesting example of such an experiment was conducted by Andrew Baum and Glenn Davis (1980). Their experiment is unusual in that it was a true experiment, with the experimenters having complete control over the independent variables, even though it was conducted partly as a field experiment and partly as a laboratory experiment. (Often field experiments must be run as quasi experiments because of practical complications.)

Baum and Davis wanted to study the effect of residential crowding on stress and social behavior. Through the cooperation of a college administra-