PREVIEW

Does experiencing violence in video games have an effect on the players' behavior? One study suggests that the answer is yes and no. Bartholow and Anderson (2002) randomly assigned male and female undergraduate students to play a violent video game or a nonviolent game. After the game, each participant was asked to take part in a competitive reaction time game with another student who was actually part of the research team (a confederate). Both students were instructed to respond as quickly as possible to a stimulus tone and, on each trial, the loser was punished with a blast of white noise delivered through headphones. Part of the instructions allowed the participant to set the level of intensity for the punishment noise and the level selected was used as a measure of aggressive behavior for that participant, with higher levels indicating more aggressive behavior. The results of the study showed that the level of violence in the video game had essentially no effect on the behavior of the female participants but the males were significantly more aggressive after playing the violent game compared to the nonviolent game.

The Bartholow and Anderson study is an example of research that involves two independent variables in the same study. The independent variables are:

1. Level of violence in the video game (high or low)
2. Gender (male or female)

The results of the study indicate that the effect of one variable (violence) *depends on* another variable (gender).

You should realize that it is quite common to have two variables that interact in this way. For example, a drug may have profound effects on some patients and have no effect whatsoever on others. Some children survive abusive environments and live normal, productive lives, while others show serious difficulties. To observe how one variable interacts with another, it is necessary to study both variables simultaneously in one study. However, the analysis of variance (ANOVA) procedures introduced in Chapters 12 and 13 are limited to evaluating mean differences produced by one independent variable and are not appropriate for mean differences involving two (or more) independent variables.

Fortunately, ANOVA is a very flexible hypothesis testing procedure and can be modified again to evaluate the mean differences produced in a research study with two (or more) independent variables. In this chapter we introduce the *two-factor* ANOVA, which tests the significance of each independent variable acting alone as well as the interaction between variables.

14.1 An Overview of the Two-Factor, Independent-Measures, ANOVA: Main Effects and Interactions

LEARNING OBJECTIVES

1. Describe the structure of a factorial research design, especially a two-factor independent-measures design, using the terms factor and level.
2. Define a main effect and an interaction and identify the patterns of data that produce main effects and interactions.
3. Identify the three *F*-ratios for a two-factor ANVOA and explain how they are related to each other.

In most research situations, the goal is to examine the relationship between two variables. Typically, the research study attempts to isolate the two variables to eliminate or reduce the influence of any outside variables that may distort the relationship being studied. A typical experiment, for example, focuses on one independent variable (which is expected to influence behavior) and one dependent variable (which is a measure of the behavior). In real life, however, variables rarely exist in isolation. That is, behavior usually is influenced by a variety of different variables acting and interacting simultaneously. To examine these more complex, real-life situations, researchers often design research studies that include more than one independent variable. Thus, researchers systematically change two (or more) variables and then observe how the changes influence another (dependent) variable.

An independent variable is a manipulated variable in an experiment. A quasi-independent variable is not manipulated but defines the groups of scores in a nonexperimental study.

In Chapters 12 and 13, we examined ANOVA for *single-factor* research designs—that is, designs that included only one independent variable or only one quasi-independent variable. When a research study involves more than one factor, it is called a *factorial design*. In this chapter, we consider the simplest version of a factorial design. Specifically, we examine ANOVA as it applies to research studies with exactly two factors. In addition, we limit our discussion to studies that use a separate sample for each treatment condition—that is, independent-measures designs. Finally, we consider only research designs for which the sample size (n) is the same for all treatment conditions. In the terminology of ANOVA, this chapter examines *two-factor, independent-measures, equal* n *designs.*

We will use the Bartholow and Anderson video game violence study described in the Chapter Preview to introduce the two-factor research design. Table 14.1 shows the structure of the study. Note that the study involves two separate factors: one factor is manipulated by the researcher, changing from a violent to a nonviolent game, and the second factor is gender, which varies from male to female. The two factors are used to create a *matrix* with the different genders defining the rows and the different levels of violence defining the columns. The resulting two-by-two matrix shows four different combinations of the variables, producing four different conditions. Thus, the research study would require four separate samples, one for each *cell*, or box, in the matrix. The dependent variable for the study is the level of aggressive behavior for the participants in each of the four conditions.

TABLE 14.1
The structure of a two-factor experiment presented as a matrix. The two factors are gender and level of violence in a video game, with two levels for each factor.

		Factor B: Level of Violence	
		Nonviolent	Violent
Factor A: Gender	Male	Scores for a group of males who play a nonviolent video game	Scores for a group of males who play a violent video game
	Female	Scores for a group of females who play a nonviolent video game	Scores for a group of females who play a violent video game

The two-factor ANOVA tests for mean differences in research studies that are structured like the gender-and-video-violence study in Table 14.1. For this example, the two-factor ANOVA evaluates three separate sets of mean differences.

1. What happens to the level of aggressive behavior when violence is added or taken away from the game?
2. Is there a difference in the aggressive behavior for male participants compared to females?
3. Is aggressive behavior affected by specific combinations of game violence and gender? (For example, a violent game may have a large effect on males but only a small effect for females.)

Thus, the two-factor ANOVA allows us to examine three types of mean differences within one analysis. In particular, we conduct three separate hypotheses tests for the same data, with a separate F-ratio for each test. The three F-ratios have the same basic structure:

$$F = \frac{\text{variance (differences) between treatments}}{\text{variance (differences) expected if there is no treatment effect}}$$

In each case, the numerator of the F-ratio measures the actual mean differences in the data, and the denominator measures the differences that would be expected if there is no treatment effect. As always, a large value for the F-ratio indicates that the sample mean differences are greater than would be expected by chance alone, and therefore provides evidence of a treatment effect. To determine whether the obtained F-ratios are *significant*, we need to compare each F-ratio with the critical values found in the F-distribution table in Appendix B.

Main Effects and Interactions

As noted in the previous section, a two-factor ANOVA actually involves three distinct hypothesis tests. In this section, we examine these three tests in more detail.

Traditionally, the two independent variables in a two-factor experiment are identified as factor A and factor B. For the study presented in Table 14.1, gender is factor A, and the level of violence in the game is factor B. The goal of the study is to evaluate the mean differences that may be produced by either of these factors acting independently or by the two factors acting together.

Main Effects

One purpose of the study is to determine whether differences in gender (factor A) result in differences in behavior. To answer this question, we compare the mean score for all the males with the mean for the females. Note that this process evaluates the mean difference between the top row and the bottom row in Table 14.1.

To make this process more concrete, we present a set of hypothetical data in Table 14.2. The table shows the mean score for each of the treatment conditions (cells) as well as the overall mean for each column (each level of violence) and the overall mean for each row (each gender group). These data indicate that the male participants (the top row) had an overall mean of $M = 8$. This overall mean was obtained by computing the average of the two means in the top row. In contrast, the female participants had an overall mean of $M = 4$ (the mean for the bottom row). The difference between these means constitutes what is called the *main effect* for gender, or the *main effect for factor* A.

TABLE 14.2 Hypothetical data for an experiment examining the effect of violence in a video game on the aggressive behavior of males and females.

	Nonviolent Game	Violent Game	
Males	$M = 7$	$M = 9$	$M = 8$
Females	$M = 3$	$M = 5$	$M = 4$
	$M = 5$	$M = 7$	

Similarly, the main effect for factor B (level of violence) is defined by the mean difference between the columns of the matrix. For the data in Table 14.2, the two groups of participants who played a nonviolent game had an overall mean score of $M = 5$. Participants who played a violent game had an overall average score of $M = 7$. The difference between these means constitutes the *main effect* for the level of game violence, or the *main effect for factor* B.

DEFINITION

The mean differences among the levels of one factor are referred to as the **main effect** of that factor. When the design of the research study is represented as a matrix with one factor determining the rows and the second factor determining the columns, then the mean differences among the rows describe the main effect of one factor, and the mean differences among the columns describe the main effect for the second factor.

The mean differences between columns or rows simply *describe* the main effects for a two-factor study. As we have observed in earlier chapters, the existence of sample mean differences does not necessarily imply that the differences are *statistically significant*. In general, two samples are not expected to have exactly the same means. There will always be small differences from one sample to another, and you should not automatically assume that these differences are an indication of a systematic treatment effect. In the case of a two-factor study, any main effects that are observed in the data must be evaluated with a hypothesis test to determine whether they are statistically significant effects. Unless the hypothesis test demonstrates that the main effects are significant, you must conclude that the observed mean differences are simply the result of sampling error.

The evaluation of main effects accounts for two of the three hypothesis tests in a two-factor ANOVA. We state hypotheses concerning the main effect of factor A and the main effect of factor B and then calculate two separate F-ratios to evaluate the hypotheses.

For the example we are considering, factor A involves the comparison of two different genders. The null hypothesis would state that there is no difference between the two levels; that is, gender has no effect on aggressive behavior. In symbols,

$$H_0: \mu_{A_1} = \mu_{A_2}$$

The alternative hypothesis is that the two genders do produce different aggression scores:

$$H_1: \mu_{A_1} \neq \mu_{A_2}$$

To evaluate these hypotheses, we compute an F-ratio that compares the actual mean differences between the two genders vs. the amount of difference that would be expected without any systematic difference.

$$F = \frac{\text{variance (differences) between the means for factor } A}{\text{variance (differences) expected if there is no treatment effect}}$$

$$F = \frac{\text{variance (differences) between the row means}}{\text{variance (differences) expected if there is no treatment effect}}$$

Similarly, factor B involves the comparison of the two different violence conditions. The null hypothesis states that there is no difference in the mean level of aggression between the two conditions. In symbols,

$$H_0: \mu_{B_1} = \mu_{B_2}$$

As always, the alternative hypothesis states that the means are different:

$$H_1: \mu_{B_1} \neq \mu_{B_2}$$

Again, the F-ratio compares the obtained mean difference between the two violence conditions vs. the amount of difference that would be expected if there is no systematic treatment effect.

$$F = \frac{\text{variance (differences) between the means for factor } B}{\text{variance (differences) expected if there is no treatment effect}}$$

$$F = \frac{\text{variance (differences) between the column means}}{\text{variance (differences) expected if there is no treatment effect}}$$

Interactions

In addition to evaluating the main effect of each factor individually, the two-factor ANOVA allows you to evaluate other mean differences that may result from unique combinations of the two factors. For example, specific combinations of game violence and gender acting together may have effects that are different from the effects of gender or game violence acting alone. Any "extra" mean differences that are not explained by the main effects are called an *interaction*, or an *interaction between factors*. The real advantage of combining two factors within the same study is the ability to examine the unique effects caused by an interaction.

DEFINITION

An **interaction** between two factors occurs whenever the mean differences between individual treatment conditions, or cells, are different from what would be predicted from the overall main effects of the factors.

To make the concept of an interaction more concrete, we reexamine the data shown in Table 14.2. For these data, there is no interaction; that is, there are no extra mean differences that are not explained by the main effects. For example, within each violence condition (each column of the matrix) the average level of aggression for the male participants is 4 points higher than the average for the female participants. This 4-point mean difference is exactly what is predicted by the overall main effect for gender.

The data in Table 14.3 show the same pattern of results that was obtained in the Bartholow and Anderson research study.

Now consider a different set of data shown in Table 14.3. These new data show exactly the same main effects that existed in Table 14.2 (the column means and the row means have not been changed). There is still a 4-point mean difference between the two rows (the main effect for gender) and a 2-point mean difference between the two columns (the main effect for violence). But now there is an interaction between the two factors. For example, for the male participants (top row), there is a 4-point difference in the level of aggression after a violent game vs. a nonviolent game. This 4-point difference cannot be explained by the 2-point main effect for the violence factor. Also, for the female participants (bottom row), the data show no difference between the two game violence conditions. Again, the zero difference is not what would be expected based on the 2-point main effect for the game violence factor. Mean differences that are not explained by the main effects are an indication of an interaction between the two factors.

TABLE 14.3

Hypothetical data for an experiment examining the effect of violence in a video game on the aggressive behavior of males and females. The data show the same main effects as the values in Table 14.2 but the individual treatment means have been modified to create an interaction.

	Nonviolent Game	Violent Game	
Male	$M = 6$	$M = 10$	$M = 8$
Female	$M = 4$	$M = 4$	$M = 4$
	$M = 5$	$M = 7$	

To evaluate the interaction, the two-factor ANOVA first identifies mean differences that are not explained by the main effects. The extra mean differences are then evaluated by an F-ratio with the following structure:

$$F = \frac{\text{variance (mean differences) not explained by main effects}}{\text{variance (differences) expected if there is no treatment effects}}$$

The null hypothesis for this F-ratio simply states that there is no interaction:

H_0: There is no interaction between factors A and B. The mean differences between treatment conditions are explained by the main effects of the two factors.

The alternative hypothesis is that there is an interaction between the two factors:

H_1: There is an interaction between factors. The mean differences between treatment conditions are not what would be predicted from the overall main effects of the two factors.

■ More about Interactions

In the previous section, we introduced the concept of an interaction as the unique effect produced by two factors working together. This section presents two alternative definitions of an interaction. These alternatives are intended to help you understand the concept of an interaction and to help you identify an interaction when you encounter one in a set of data. You should realize that the new definitions are equivalent to the original and simply present slightly different perspectives on the same concept.

The first new perspective on the concept of an interaction focuses on the notion of independence for the two factors. More specifically, if the two factors are independent, so that one factor does not influence the effect of the other, then there is no interaction. On the other hand, when the two factors are not independent, so that the effect of one factor *depends on* the other, then there is an interaction. The notion of dependence between factors is consistent with our earlier discussion of interactions. If one factor influences the effect of the other, then unique combinations of the factors produce unique effects.

DEFINITION

> When the effect of one factor depends on the different levels of a second factor, then there is an **interaction** between the factors.

This definition of an interaction should be familiar in the context of a "drug interaction." Your doctor and pharmacist are always concerned that the effect of one medication may be altered or distorted by a second medication that is being taken at the same time. Thus, the effect of one drug (factor A) depends on a second drug (factor B), and you have an interaction between the two drugs.

Returning to Table 14.2, you will notice that the size of the game-violence effect (first column vs. second column) *does not depend* on the gender of the participants. For these data, adding violence produces the same 2-point increase in aggressive behavior for both groups of participants. Thus, the effect of game violence does not depend on gender and there is no interaction. Now consider the data in Table 14.3. This time, the effect of adding violence *depends on* the gender of the participants. For example, there is a 4-point increase in aggressive behavior for the males but adding violence has no effect on aggression for the females. Thus, the effect of game violence depends on gender, which means that there is an interaction between the two factors.

The second alternative definition of an interaction is obtained when the results of a two-factor study are presented in a graph. In this case, the concept of an interaction can be defined in terms of the pattern displayed in the graph. Figure 14.1 shows the two sets of data we have been considering. The original data from Table 14.2, where there is no interaction, are presented in Figure 14.1(a). To construct this figure, we selected one of the factors to be displayed on the horizontal axis; in this case, the different levels of game violence are displayed. The dependent variable, the level of aggressive behavior, is shown on the vertical axis. Note that the figure actually contains two separate graphs: The top line shows the relationship between game violence and aggression for the males, and the bottom line shows the relationship for the females. In general, the picture in

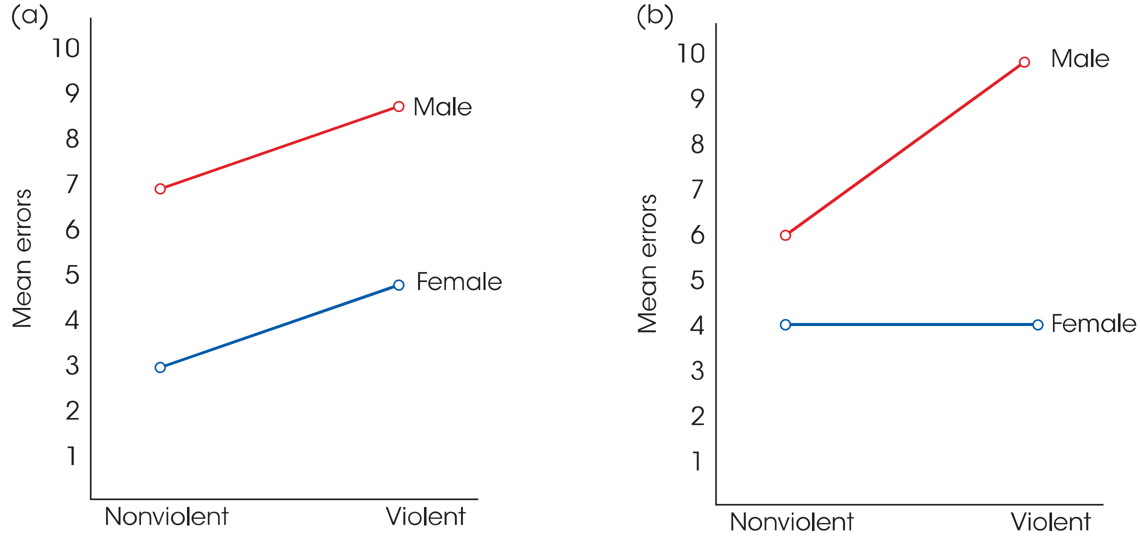

FIGURE 14.1
(a) Graph showing the treatment means for Table 14.2, for which there is no interaction. (b) Graph for Table 14.3, for which there is an interaction.

the graph matches the structure of the data matrix; the columns of the matrix appear as values along the *X*-axis, and the rows of the matrix appear as separate lines in the graph (see Box 14.1).

For the original set of data, Figure 14.1(a), note that the two lines are parallel; that is, the distance between lines is constant. In this case, the distance between lines reflects the 2-point difference in the mean aggression scores for males and females, and this 2-point difference is the same for both game violence conditions.

Now look at a graph that is obtained when there is an interaction in the data. Figure 14.1(b) shows the data from Table 14.3. This time, note that the lines in the graph are not parallel. The distance between the lines changes as you scan from left to right. For these data, the distance between the lines corresponds to the gender effect—that is, the mean difference in aggression for male vs. female participants. The fact that this difference depends on the level of game violence is an indication of an interaction between the two factors.

DEFINITION

When the results of a two-factor study are presented in a graph, the existence of nonparallel lines (lines that cross or converge) indicates an **interaction** between the two factors.

For many students, the concept of an interaction is easiest to understand using the perspective of interdependency; that is, an interaction exists when the effects of one variable *depend* on another factor. However, the easiest way to identify an interaction within a set of data is to draw a graph showing the treatment means. The presence of nonparallel lines is an easy way to spot an interaction.

BOX 14.1 Graphing Results from a Two-Factor Design

One of the best ways to get a quick overview of the results from a two-factor study is to present the data in a graph. Because the graph must display the means obtained for *two* independent variables (two factors), constructing the graph can be a bit more complicated than constructing the single-factor graphs we presented in Chapter 3 (pp. 90–91).

Figure 14.2 shows two possible graphs presenting the results from a two-factor study with 2 levels of factor *A* and 3 levels of factor *B*. With a 2 × 3 design, there are a total of 6 different treatment means that are shown in the following matrix.

		Factor *B*		
		*B*1	*B*2	*B*3
Factor *A*	A_1	$M = 10$	$M = 40$	$M = 20$
	A_2	$M = 30$	$M = 50$	$M = 30$

In the two graphs, note that values for the dependent variable (the treatment means) are shown on the vertical axis. Also note that the levels for one factor (we selected factor B) are displayed on the horizontal axis. Either factor can be used, but it is better to select one for which the different levels are measured on an interval or ratio scale. The reason for this suggestion is that an interval or ratio scale will permit you to construct a line graph as in Figure 14.2(a). If neither of the two factors is measured on an interval or ratio scale, you should use a bar graph as in Figure 14.2(b).

Line graphs: In Figure 14.2(a), we have assumed that factor *B* is an interval or a ratio variable, and the three levels for this factor are listed on the horizontal axis. Directly above the B_1 value on the horizontal axis, we have placed two dots corresponding to the two means in the B_1 column of the data matrix. Similarly, we have placed two dots above B_2 and another two dots above B_3. Finally, we have drawn a line connecting the three dots corresponding to level 1 of factor *A* (the three means in the top row of the data matrix). We have also drawn a second line that connects the three dots corresponding to level 2 of factor *A*. These lines are labeled A_1 and A_2 in the figure.

Bar graphs: Figure 14.2(b) also shows the three levels of factor *B* displayed on the horizontal axis. This time, however, we assume that factor *B* is measured on a nominal or ordinal scale, and the result is a bar graph. Directly above the B_1 value, we have drawn two bars so that the heights of the bars correspond to the two means in the B_1 column of the data matrix. Similarly, we have drawn two bars above B_2 and two more bars above B_3. Finally, the three bars corresponding to level 1 of factor *A* (the top row of the data matrix) are all colored (or shaded) to differentiate them from the three bars for level 2 of factor *A*.

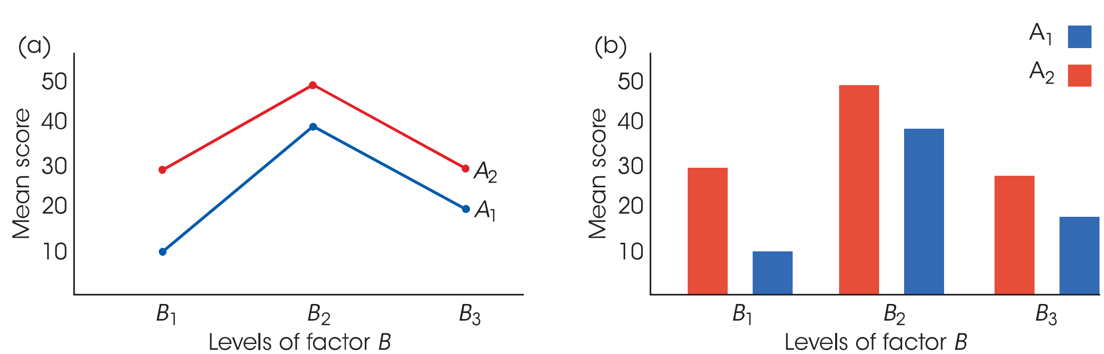

FIGURE 14.2
Two graphs showing the results from a two-factor study. A line graph is shown in (a) and a bar graph (b).

The $A \times B$ interaction typically is called "A by B" interaction. If there is an interaction between video game violence and gender, it may be called the "violence by gender" interaction.

Independence of Main Effects and Interactions

The two-factor ANOVA consists of three hypothesis tests, each evaluating specific mean differences: the A effect, the B effect, and the $A \times B$ interaction. As we have noted, these are three *separate* tests, but you should also realize that the three tests are *independent*. That is, the outcome for any one of the three tests is totally unrelated to the outcome for either of the other two. Thus, it is possible for data from a two-factor study to display any possible combination of significant and/or not significant main effects and interactions. The data sets in Table 14.4 show several possibilities.

Table 14.4(a) shows data with mean differences between levels of factor A (an A effect) but no mean differences for factor B and no interaction. To identify the A effect, notice that the overall mean for A_1 (the top row) is 10 points higher than the overall mean for A_2 (the bottom row). This 10-point difference is the main effect for factor A. To evaluate the B effect, notice that both columns have exactly the same overall mean, indicating no difference between levels of factor B; hence, there is no B effect. Finally, the absence of an interaction is indicated by the fact that the overall A effect (the 10-point difference) is constant within each column; that is, the A effect *does not depend* on the levels of factor B. (Alternatively, the data indicate that the overall B effect is constant within each row.)

TABLE 14.4
Three sets of data showing different combinations of main effects and interaction for a two-factor study. (The numerical value in each cell of the matrices represents the mean value obtained for the sample in that treatment condition.)

(a) Data showing a main effect for factor A but no B effect and no interaction

	B_1	B_2		
A_1	20	20	A_1 mean = 20	10-point difference
A_2	10	10	A_2 mean = 10	
	B_1 mean = 15	B_2 mean = 15		
	No difference			

(b) Data showing main effects for both factor A and factor B but no interaction

	B_1	B_2		
A_1	10	30	A_1 mean = 20	10-point difference
A_2	20	40	A_2 mean = 30	
	B_1 mean = 15	B_2 mean = 35		
	20-point difference			

(c) Data showing no main effect for either factor but an interaction

	B_1	B_2		
A_1	10	20	A_1 mean = 15	No difference
A_2	20	10	A_2 mean = 15	
	B_1 mean = 15	B_2 mean = 15		
	No difference			

Table 14.4(b) shows data with an *A* effect and a *B* effect but no interaction. For these data, the *A* effect is indicated by the 10-point mean difference between rows, and the *B* effect is indicated by the 20-point mean difference between columns. The fact that the 10-point *A* effect is constant within each column indicates no interaction.

Finally, Table 14.4(c) shows data that display an interaction but no main effect for factor *A* or for factor *B*. For these data, there is no mean difference between rows (no *A* effect) and no mean difference between columns (no *B* effect). However, within each row (or within each column), there are mean differences. The "extra" mean differences within the rows and columns cannot be explained by the overall main effects and therefore indicate an interaction.

The following example is an opportunity to test your understanding of main effects and interactions.

EXAMPLE 14.1

The following matrix represents the outcome of a two-factor experiment. Describe the main effect for factor *A* and the main effect for factor *B*. Does there appear to be an interaction between the two factors?

Experiment I

	B_1	B_2
A_1	$M = 10$	$M = 20$
A_2	$M = 30$	$M = 40$

You should conclude that there is a main effect for factor *A* (the scores in A_2 average 20 points higher than in A_1) and there is a main effect for factor *B* (the scores in B_2 average 10 points higher than in B_1) but there is no interaction; there is a constant 20-point difference between A_1 and A_2 that does not depend on the levels of factor *B*. ■

LEARNING CHECK

1. How many separate samples would be needed for a two-factor, independent-measures research study with 2 levels of factor *A* and 3 levels of factor *B*?

a. 2
b. 3
c. 5
d. 6

2. In a two-factor experiment with 2 levels of factor *A* and 2 levels of factor *B*, three of the treatment means are essentially identical and one is substantially different from the others. What result(s) would be produced by this pattern of treatment means?

a. a main effect for factor *A*
b. a main effect for factor *B*
c. an interaction between *A* and *B*
d. The pattern would produce main effects for both *A* and *B*, and an interaction.

3. In a two-factor ANOVA, what is the implication of a significant $A \times B$ interaction?

a. At least one of the main effects must also be significant.
b. Both of the main effects must also be significant.
c. Neither of the two main effects can be significant.
d. The significance of the interaction has no implications for the main effects.

ANSWERS

1. D, **2.** D, **3.** D

14.2 An Example of the Two-Factor ANOVA and Effect Size

LEARNING OBJECTIVES

4. Describe the two-stage structure of a two-factor ANOVA and explain what happens in each stage.
5. Compute the *SS*, *df*, and *MS* values needed for a two-factor ANOVA and explain the relationships among them.
6. Conduct a two-factor ANOVA including measures of effect size for both main effects and the interaction.

The two-factor ANOVA is composed of three distinct hypothesis tests:

1. The main effect of factor A (often called the A-effect). Assuming that factor A is used to define the rows of the matrix, the main effect of factor A evaluates the mean differences between rows.
2. The main effect of factor B (called the B-effect). Assuming that factor B is used to define the columns of the matrix, the main effect of factor B evaluates the mean differences between columns.
3. The interaction (called the $A \times B$ interaction). The interaction evaluates mean differences between treatment conditions that are not predicted from the overall main effects from factor A or factor B.

For each of these three tests, we are looking for mean differences between treatments that are larger than would be expected if there are no treatment effects. In each case, the significance of the treatment effect is evaluated by an F-ratio. All three F-ratios have the same basic structure:

$$F = \frac{\text{variance (mean differences) between treatments}}{\text{variance (mean differences) expected if there are no treatment effects}} \tag{14.1}$$

The general structure of the two-factor ANOVA is shown in Figure 14.3. Note that the overall analysis is divided into two stages. In the first stage, the total variability is separated into two components: between-treatments variability and within-treatments variability. This first stage is identical to the single-factor ANOVA introduced in Chapter 12 with each cell in the two-factor matrix viewed as a separate treatment condition. The within-treatments variability

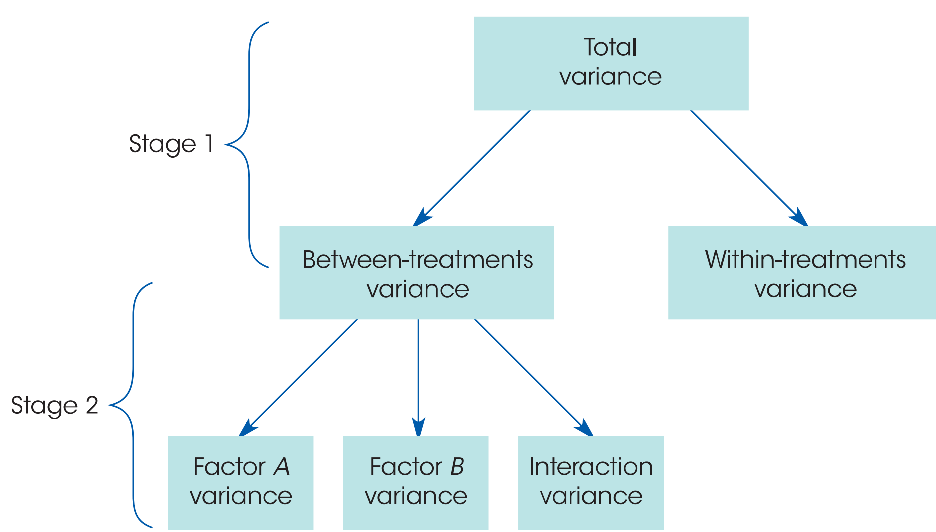

FIGURE 14.3
Structure of the analysis for a two-factor ANOVA.

that is obtained in stage 1 of the analysis is used as the denominator for the F-ratios. As we noted in Chapter 12, within each treatment, all of the participants are treated exactly the same. Thus, any differences that exist within the treatments cannot be caused by treatment effects. As a result, the within-treatments variability provides a measure of the differences that exist when there are no systematic treatment effects influencing the scores (see Equation 14.1).

The between-treatments variability obtained in stage 1 of the analysis combines all the mean differences produced by factor A, factor B, and the interaction. The purpose of the second stage is to partition the differences into three separate components: differences attributed to factor A, differences attributed to factor B, and any remaining mean differences that define the interaction. These three components form the numerators for the three F-ratios in the analysis.

The goal of this analysis is to compute the variance values needed for the three F-ratios. We need three between-treatments variances (one for factor A, one for factor B, and one for the interaction), and we need a within-treatments variance. Each of these variances (or mean squares) is determined by a sum of squares value (SS) and a degree of freedom value (df):

Remember that in ANOVA a variance is called a mean square, or *MS*.

$$\text{mean square} = MS = \frac{SS}{df}$$

EXAMPLE 14.2

To demonstrate the two-factor ANOVA, we will use a research study based on previous work by Ackerman and Goldsmith (2011). Their study compared learning performance by students who studied text either from printed pages or from a computer screen. The results from the study indicate that students do much better studying from printed pages if their study time is self-regulated. However, when the researchers fixed the time spent studying, there was no difference between the two conditions. Apparently, students are less accurate predicting their learning performance or have trouble regulating study time when working with a computer screen compared to working with paper. Table 14.5 shows data from a two-factor study replicating the Ackerman and Goldsmith experiment. The two factors are mode of presentation (paper or computer screen) and time control (self-regulated or fixed). A separate group of $n = 5$ students was tested in each of the four conditions. The dependent variable is student performance on a quiz covering the text that was studied.

TABLE 14.5
Data for a two-factor study comparing two levels of time control (self-regulated or fixed by the researchers) and two levels of text presentation (paper and computer screen). The dependent variable is performance on a quiz covering the text that was presented. The study involves four treatment conditions with $n = 5$ participants in each treatment.

Factor A Time Control	Factor B: Text Presentation Mode: Paper	Computer Screen		
Self-regulated	11 8 9 10 7 $M = 9$ $T = 45$ $SS = 10$	4 4 8 5 4 $M = 5$ $T = 25$ $SS = 12$	$T_{row} = 70$	$N = 20$ $G = 155$ $\Sigma X^2 = 1303$
Fixed	10 7 10 6 7 $M = 8$ $T = 40$ $SS = 14$	10 6 10 10 9 $M = 9$ $T = 45$ $SS = 12$	$T_{row} = 85$	
	$T_{col} = 85$	$T_{col} = 70$		

The data are displayed in a matrix with the two levels of time control (factor A) making up the rows and the two levels of presentation mode (factor B) making up the columns. Note that the data matrix has a total of four *cells* or treatment conditions with a separate sample of $n = 5$ participants in each condition. Most of the notation should be familiar from the single-factor ANOVA presented in Chapter 12. Specifically, the treatment totals are identified by T values, the total number of scores in the entire study is $N = 20$, and the grand total (sum) of all 20 scores is $G = 155$. In addition to these familiar values, we have included the totals for each row and for each column in the matrix. The goal of the ANOVA is to determine whether the mean differences observed in the data are significantly greater than would be expected if there were no treatment effects.

Stage 1 of the Two-Factor Analysis

The first stage of the two-factor analysis separates the total variability into two components: between-treatments and within-treatments. The formulas for this stage are identical to the formulas used in the single-factor ANOVA in Chapter 12 with the provision that each cell in the two-factor matrix is treated as a separate treatment condition. The formulas and the calculations for the data in Table 14.5 are as follows:

Total Variability

$$SS_{\text{total}} = \Sigma X^2 - \frac{G^2}{N} \tag{14.2}$$

For these data,

$$\begin{aligned} SS_{\text{total}} &= 1303 - \frac{155^2}{20} \\ &= 1303 - 1201.25 \\ &= 101.75 \end{aligned}$$

This SS value measures the variability for all $N = 20$ scores and has degrees of freedom given by

$$df_{\text{total}} = N - 1 \tag{14.3}$$

For the data in Table 14.5, $df_{\text{total}} = 19$.

Within-Treatments Variability To compute the variance within treatments, we first compute SS and $df = n - 1$ for each of the individual treatment conditions. Then the within-treatments SS is defined as

$$SS_{\text{within treatments}} = \Sigma SS_{\text{each treatment}} \tag{14.4}$$

And the within-treatments df is defined as

$$df_{\text{within treatments}} = \Sigma df_{\text{each treatment}} \tag{14.5}$$

For the four treatment conditions in Table 14.5,

$$SS_{\text{within treatments}} = 10 + 12 + 14 + 12 = 48$$

$$df_{\text{within treatments}} = 4 + 4 + 4 + 4 = 16$$

Between-Treatments Variability Because the two components in stage 1 must add up to the total, the easiest way to find $SS_{\text{between treatments}}$ is by subtraction.

$$SS_{\text{between treatments}} = SS_{\text{total}} - SS_{\text{within}} \tag{14.6}$$

For the data in Table 14.5, we obtain

$$SS_{\text{between treatments}} = 101.75 - 48 = 53.75$$

However, you can also use the computational formula to calculate $SS_{\text{between treatments}}$ directly.

$$\text{SS}_{\text{between treatments}} = \Sigma\frac{T^2}{n} - \frac{G^2}{N} \qquad \textbf{(14.7)}$$

For the data in Table 14.5, there are four treatments (four T values), each with $n = 5$ scores, and the between-treatments SS is

$$SS_{\text{between treatments}} = \frac{45^2}{5} + \frac{25^2}{5} + \frac{40^2}{5} + \frac{45^2}{5} - \frac{155^2}{20}$$
$$= 405 + 125 + 320 + 405 - 1201.25$$
$$= 53.75$$

The between-treatments df value is determined by the number of treatments (or the number of T values) minus one. For a two-factor study, the number of treatments is equal to the number of cells in the matrix. Thus,

$$df_{\text{between treatments}} = \text{number of cells} - 1 \qquad \textbf{(14.8)}$$

For these data, $df_{\text{between treatments}} = 3$.

This completes the first stage of the analysis. Note that the two components add to equal the total for both SS values and df values.

$$SS_{\text{between treatments}} + SS_{\text{within treatments}} = SS_{\text{total}}$$
$$53.75 + 48 = 101.75$$
$$df_{\text{between treatments}} + df_{\text{within treatments}} = df_{\text{total}}$$
$$3 + 16 = 19$$

Stage 2 of the Two-Factor Analysis

The second stage of the analysis determines the numerators for the three F-ratios. Specifically, this stage determines the between-treatments variance for factor A, factor B, and the interaction.

1. **Factor *A*** The main effect for factor A evaluates the mean differences between the levels of factor A. For this example, factor A defines the rows of the matrix, so we are evaluating the mean differences between rows. To compute the SS for factor A, we calculate a between-treatment SS using the row totals exactly the same as we computed $SS_{\text{between treatments}}$ using the treatment totals (T values) earlier. For factor A, the row totals are 70 and 85, and each total was obtained by adding 10 scores.

   Therefore,

$$SS_A = \Sigma\frac{T^2_{\text{ROW}}}{n_{\text{ROW}}} - \frac{G^2}{N} \qquad \textbf{(14.9)}$$

   For our data,

$$SS_A = \frac{70^2}{10} + \frac{85^2}{10} - \frac{155^2}{20}$$
$$= 490 + 722.5 - 1201.25$$
$$= 11.25$$

Factor A involves two treatments (or two rows), easy and difficult, so the df value is

$$\begin{aligned} df_A &= \text{number of rows} - 1 \\ &= 2 - 1 \\ &= 1 \end{aligned} \tag{14.10}$$

2. **Factor *B*** The calculations for factor B follow exactly the same pattern that was used for factor A, except for substituting columns in place of rows. The main effect for factor B evaluates the mean differences between the levels of factor B, which define the columns of the matrix.

$$SS_B = \Sigma\frac{T^2_{\text{COL}}}{n_{\text{COL}}} - \frac{G^2}{N} \tag{14.11}$$

For our data, the column totals are 85 and 70, and each total was obtained by adding 10 scores. Thus,

$$\begin{aligned} SS_B &= \frac{85^2}{10} + \frac{70^2}{10} - \frac{155^2}{20} \\ &= 722.5 + 490 - 1201.25 \\ &= 11.25 \end{aligned}$$

$$\begin{aligned} df_B &= \text{number of columns} - 1 \\ &= 2 - 1 \\ &= 1 \end{aligned} \tag{14.12}$$

3. **The *A* × *B* Interaction** The $A \times B$ interaction is defined as the "extra" mean differences not accounted for by the main effects of the two factors. We use this definition to find the SS and df values for the interaction by simple subtraction. Specifically, the between-treatments variability is partitioned into three parts: the A effect, the B effect, and the interaction (see Figure 14.3). We have already computed the SS and df values for A and B, so we can find the interaction values by subtracting to find out how much is left. Thus,

$$SS_{A\times B} = SS_{\text{between treatments}} - SS_A - SS_B \tag{14.13}$$

For our data,

$$\begin{aligned} SS_{A\times B} &= 53.75 - 11.25 - 11.25 \\ &= 31.25 \end{aligned}$$

Similarly,

$$\begin{aligned} df_{A\times B} &= df_{\text{between treatments}} - df_A - df_B \\ &= 3 - 1 - 1 \\ &= 1 \end{aligned} \tag{14.14}$$

An easy to remember alternative formula for $df_{A\times B}$ is

$$\begin{aligned} df_{A\times B} &= df_A \times df_B \\ &= 1 \times 1 = 1 \end{aligned} \tag{14.15}$$

■ Mean Squares and *F*-Ratios for the Two-Factor ANOVA

The two-factor ANOVA consists of three separate hypothesis tests with three separate *F*-ratios. The denominator for each *F*-ratio is intended to measure the variance (differences) that would be expected if there are no treatment effects. As we saw in Chapter 12, the within-treatments variance is the appropriate denominator for an independent-measures design (see p. 373). The within-treatments variance is called a *mean square*, or *MS*, and is computed as follows:

$$MS_{\text{within treatments}} = \frac{SS_{\text{within treatments}}}{df_{\text{within treatments}}}$$

For the data in Table 14.5,

$$MS_{\text{within treatments}} = \frac{48}{16} = 3$$

This value forms the denominator for all three *F*-ratios.

The numerators of the three *F*-ratios all measured variance or differences between treatments: differences between levels of factor *A*, differences between levels of factor *B*, and extra differences that are attributed to the $A \times B$ interaction. These three variances are computed as follows:

$$MS_A = \frac{SS_A}{df_A} \quad MS_B = \frac{SS_B}{df_B} \quad MS_{A\times B} = \frac{SS_{A\times B}}{df_{A\times B}}$$

For the data in Table 14.4, the three *MS* values are

$$MS_A = \frac{11.25}{1} = 11.25 \quad MS_B = \frac{11.25}{1} = 11.25 \quad MS_{A\times B} = \frac{31.25}{1} = 31.25$$

Finally, the three *F*-ratios are

$$F_A = \frac{MS_A}{MS_{\text{within treatments}}} = \frac{11.25}{3} = 3.75$$

$$F_B = \frac{MS_B}{MS_{\text{within treatments}}} = \frac{11.25}{3} = 3.75$$

$$F_{A\times B} = \frac{MS_{A\times B}}{MS_{\text{within treatments}}} = \frac{31.25}{3} = 10.42$$

To determine the significance of each *F*-ratio, we must consult the *F* distribution table using the *df* values for each of the individual *F*-ratios. For this example, all three *F*-ratios have $df = 1$ for the numerator and $df = 16$ for the denominator. Checking the table with $df = 1, 16$, we find a critical value of 4.49 for $\alpha = .05$ and a critical value of 8.53 for $\alpha = .01$. For both main effects, we obtained $F = 3.75$, so neither of the main effects is significant. For the interaction, we obtained $F = 10.42$, which exceeds both of the critical values, so we conclude that there is a significant interaction between the two factors. That is, the difference between the two modes of presentation depends on how studying time is controlled. ■

Table 14.6 is a summary table for the complete two-factor ANOVA from Example 14.2. Although these tables are no longer commonly used in research reports, they provide a concise format for displaying all of the elements of the analysis.

TABLE 14.6
A summary table for the two-factor ANOVA for the data from Example 14.2.

Source	*SS*	*df*	*MS*	*F*
Between treatments	53.75	3		
Factor *A* (time control)	11.25	1	11.25	$F(1, 16) = 3.75$
Factor *B* (paper/screen)	11.25	1	11.25	$F(1, 16) = 3.75$
$A \times B$	31.25	1	31.25	$F(1, 16) = 10.42$
Within treatments	48	16	3	
Total	101.75	19		

The following example is an opportunity to test your understanding of the calculations required for a two-factor ANOVA.

EXAMPLE 14.3

The following data summarize the results from a two-factor independent-measures experiment:

		Factor *B*		
		B_1	B_2	B_3
Factor *A*	A_1	$n = 10$ $T = 0$ $SS = 30$	$n = 10$ $T = 10$ $SS = 40$	$n = 10$ $T = 20$ $SS = 50$
	A_2	$n = 10$ $T = 40$ $SS = 60$	$n = 10$ $T = 30$ $SS = 50$	$n = 10$ $T = 20$ $SS = 40$

Calculate the total for each level of factor *A* and compute *SS* for factor *A*, then calculate the totals for factor *B*, and compute *SS* for this factor. You should find that the totals for factor *A* are 30 and 90, and $SS_A = 60$. All three totals for factor *B* are equal to 40. Because they are all the same, there is no variability, and $SS_B = 0$. ■

■ Measuring Effect Size for the Two-Factor ANOVA

The general technique for measuring effect size with an ANOVA is to compute a value for η^2, the percentage of variance that is explained by the treatment effects. For a two-factor ANOVA, we compute three separate values for eta squared: one measuring how much of the variance is explained by the main effect for factor *A*, one for factor *B*, and a third for the interaction. As we did with the repeated-measures ANOVA (p. 427) we remove any variability that can be explained by other sources before we calculate the percentage for each of the three specific treatment effects. Thus, for example, before we compute the η^2 for factor *A*, we remove the variability that is explained by factor *B* and the variability explained by the interaction. The resulting equation is

$$\text{for factor } A,\ \eta^2 = \frac{SS_A}{SS_{\text{total}} - SS_B - SS_{A\times B}} \tag{14.16}$$

Note that the denominator of Equation 14.15 consists of the variability that is explained by factor *A* and the other *unexplained* variability. Thus, an equivalent version of the equation is,

$$\text{for factor } A,\ \eta^2 = \frac{SS_A}{SS_A + SS_{\text{within treatments}}} \tag{14.17}$$