# Chapter 13
## Two-Way Factorial ANOVA:
## Using More than One Independent Variable

### 13.1 Introduction to Two-Way ANOVA

In the previous two chapters we introduced **One-Way ANOVA** designs that involve multiple levels of one independent variable (or factor). The present chapter will introduce a more sophisticated ANOVA model that allows the researcher to test the effectives of **two** independent variables; hence, this procedure is called **Two-Way ANOVA** (sometimes also called **Factorial ANOVA**). That is, factorial ANOVA improves on one-way ANOVA in that the researcher can simultaneously assess the effects of *two (or more) independent variables* on a single dependent variable within the same analysis. Thus, factorial ANOVA yields the same information that two one-way ANOVA's would, but it does so in one analysis.

But that's not all. Factorial ANOVA also allows the investigator to determine the possible *combined* effects of the independent variables. That is, it also assesses the ways in which these variables *interact* with one another to influence scores on the dependent variable. Although understanding such **interaction effects** can be a complex and difficult task, it is essential to the progress of science, since in the real world many variables interact with one another to determine behavior. In this chapter we provide a basic introduction to factorial ANOVA and the SPSS program that performs this powerful statistical analysis.

The conceptual basis of factorial ANOVA is essentially the same as that of one-way ANOVA, and the interpretation of the resulting F-values is also based on the same logic as in the one-way ANOVA. The difference is that where one-way ANOVA only generates one F-value, two-way ANOVA generates **three** F-values: one to test the **main effects** of each factor, and a third to test the **interaction effect** (i.e., the combined effect of the two factors).

#### 13.1a The need for a *factorial combination* of Independent Variables

A basic requirement for factorial experimental design is that the levels of the two independent variables have been completely crossed in a factorial combination. This means that each level of the first independent variable must be combined with each level of the other independent variable. In the simplest two-way ANOVA (a **2 x 2 design**), four different groups of participants would be needed. If the first factor (Factor A) is **GENDER** (where level **A1** is male employees and level **A2** is female employees), and the second factor (Factor B) is **MASC** (where level **B1** is low-masculine employees and level **B2** is high-masculine employees), four combinations would be required to permit a factorial ANOVA. Each unique combination is referred to as a **cell** (see Table

13.1). In this table, the four cells (in **boldface**) represent these four possible combinations.

| Table 13.1 | | | | |
|---|---|---|---|---|
| | | **Gender (Factor A)** | | |
| | | Males (A1) | Females (A2) | |
| **Masculinity (Factor B)** | Low (B1) | **Low-masc Males (A1B1)** | **Low-masc Females (A2B1)** | |
| | High (B2) | **High-masc Males (A1B2)** | **High-masc Females (A2B2)** | |

Thus, the four cells of the factorial combination shown in Table 13.1 for a **2 (Gender) x 2 (Masculinity)** factorial design would require us to divide our participants into the following four groups:

(1) Low-masculine males (A1B1)
(2) High-masculine males (A1B2)
(3) Low-masculine females (A2B1)
(4) High-masculine females (A2B2)

This factorial combination will allow us to compare the scores of men vs. women (**GENDER**) and low-masculine vs. high-masculine employees (**MASC**) on a given dependent variable. This would be the same as if we did two separate studies and conducted two t-tests (one comparing the male vs. female scores, and one comparing low- vs. high-masculine employees' scores). But it would be more economical and efficient, because we would get the same information from one study and one analysis (the 2 x 2 ANOVA). What is crucial to the factorial combination of these two independent variables is that we are also able to assess the possible interaction effect of the two independent variables combined.

## 13.2 Main Effects & Interactions

For this chapter's example we will examine differences in overall masculinity (i.e., **MASCTOT**, the dependent variable) in relation to EZ employees' **sex (GENDER)** and **masculinity category (MASC)**, the two independent variables. Recall from the two previous chapters that all ANOVA's are conceptually based upon a comparison of variance attributed to the independent variable (called *between-groups*, or *treatment* variance) to the variance due to random fluctuation (called *within-groups*, or *error* variance). Factorial ANOVA is conceptually based on the same type of ratio computations. The difference is that a 2 x 2 factorial ANOVA yields three F-ratios, instead of only one as in the one-way ANOVA. In the

following we will briefly review the reasons for this and show how to interpret these three F-values.

## 13.2a Sources of Variance in Factorial ANOVA

The reason that **three F-ratios** are computed is that conducting factorial ANOVA is like performing two one-way ANOVA's to assess the simple **Main Effects** of each of the two independent variables (**Factors A** and **B**), and a third analysis to assess the **A x B interaction effect** on the dependent variable (**MASCTOT**). Specifically, in a between-groups design, between-groups variance can be separated into three sources:

- variance due to the **main effect of Factor A** (**GENDER**)
  - masculinity differences between the two groups of men vs. women
  - evaluated by comparing the *column* **marginal means**
- variance due to the **main effect of Factor B** (**MASC**)
  - masculinity differences between the two groups of low vs. high-masculinity categories
  - evaluated by comparing the *row* **marginal means**
- variance due to the **A x B interaction effect** (**GENDER x MASC**)
  - masculinity differences between the four groups of combinations of low and high masculine men and women
  - evaluated by comparing the **cell means** of the factorial combination.

Three separate F-ratios are computed to determine how much of the variance in the dependent variable can be attributed to each of these three effects. Each F-value represents the ratio of the variance from that particular effect relative to random error variance. Thus, the three sources of between-group variance (Factor A main effect, Factor B main effect, and A x B interaction) result in three F-values. As in the one-way ANOVA, the significance of each effect is decided by looking at the probability associated with each F-value (i.e., if $p < .05$, the effect is significant).

## 13.2b Interpreting the Results of Factorial ANOVA

The **interpretation of main effects** from a 2 x 2 factorial ANOVA is straightforward. If $p < .05$ for the main effect of a particular factor, then there is a significant effect for that factor. All we have to do is examine the marginal means for the levels of the factor to determine which group is significantly higher (or lower) than the other, then state in words what the differences are.

The **interpretation of the interaction effect** is more complex. If the F-value of the interaction effect is *not* significant (i.e., $p > .05$) , then our conclusion would be that gender differences in overall masculinity scores *did not depend* on the level of masculinity (that is, the same gender difference would be seen for both the low-masculinity group and the high-masculinity group).

If the F-value for the interaction *is* significant ($p <$ .05), then we would conclude that gender differences in masculinity *depend on* the level of masculinity. For example, it might turn out that when comparing *high-masculine* men to *high-masculine* women, men have higher overall masculinity scores than do women, but that when comparing *low-masculine* men and women, we might see no significant difference in overall masculinity scores between these two groups.

The above example of an interpretation of a significant interaction would require a comparison of the four cell means relevant to the interaction effect (see our discussion of multiple comparisons in Chapters 11 and 12). In our example of the effects of **GENDER** and **MASC** on overall masculinity (**MASCTOT**), this would involve a comparison of the means of each of the four groups described earlier (high-masculine men, low-masculine men, high-masculine women, and low-masculine women). Unfortunately SPSS does not have an option to compute multiple comparisons for factorial ANOVA, so the researcher must do these analyses by hand. For now, suffice it to say that these comparisons would permit us to draw specific conclusions regarding interaction effects, such as those described above.

## 13.3 Running Two-way ANOVA

Open your **ezdata.sav** file and select **Analyze, General Linear Model, Univariate...** from the menu (Figure 13.1).
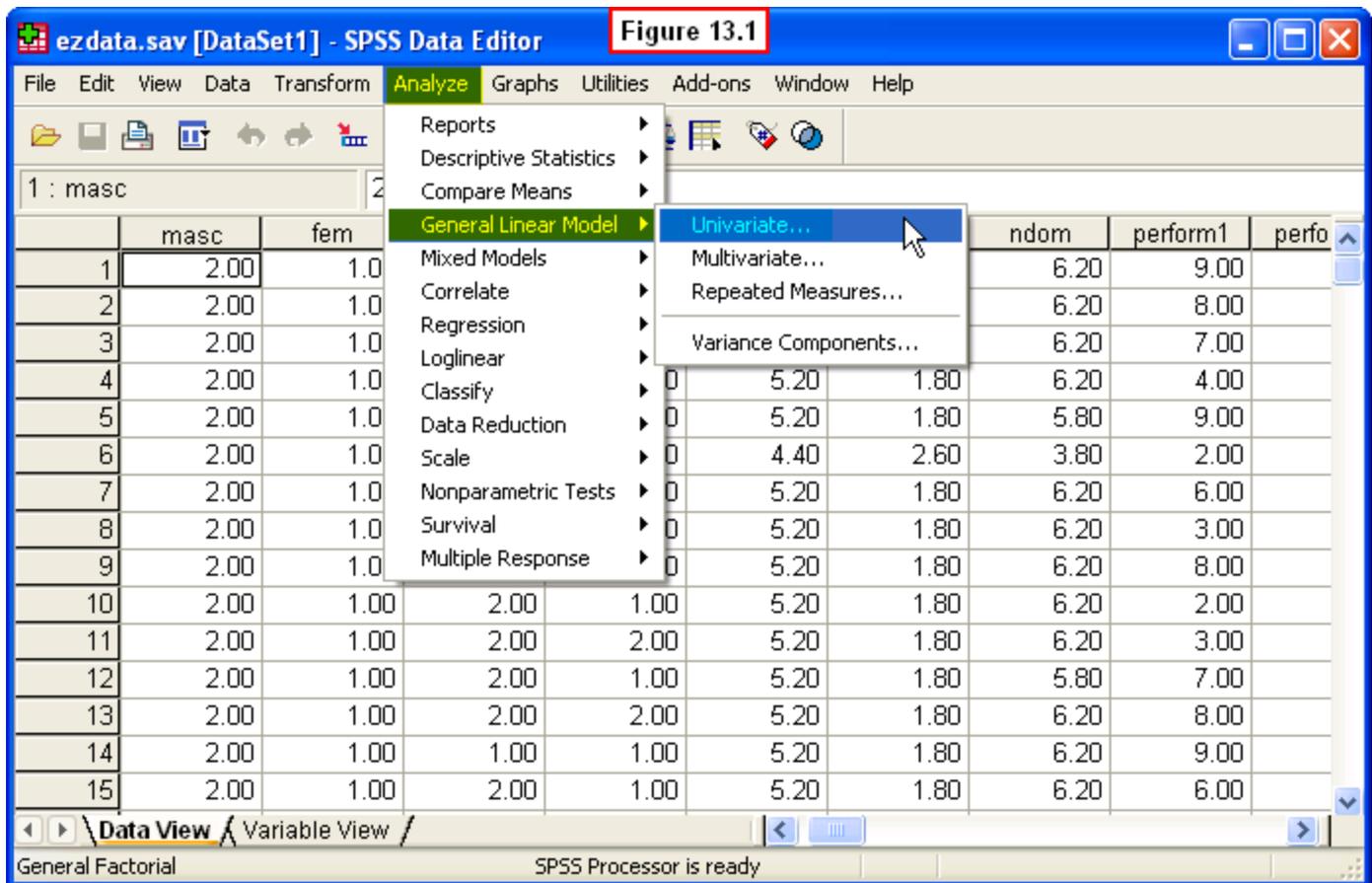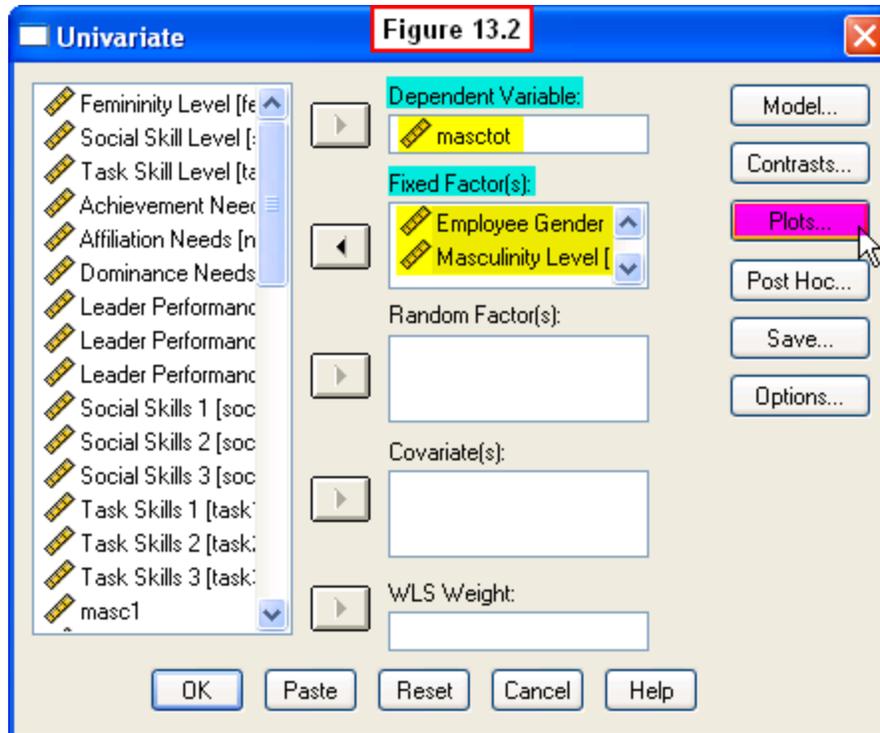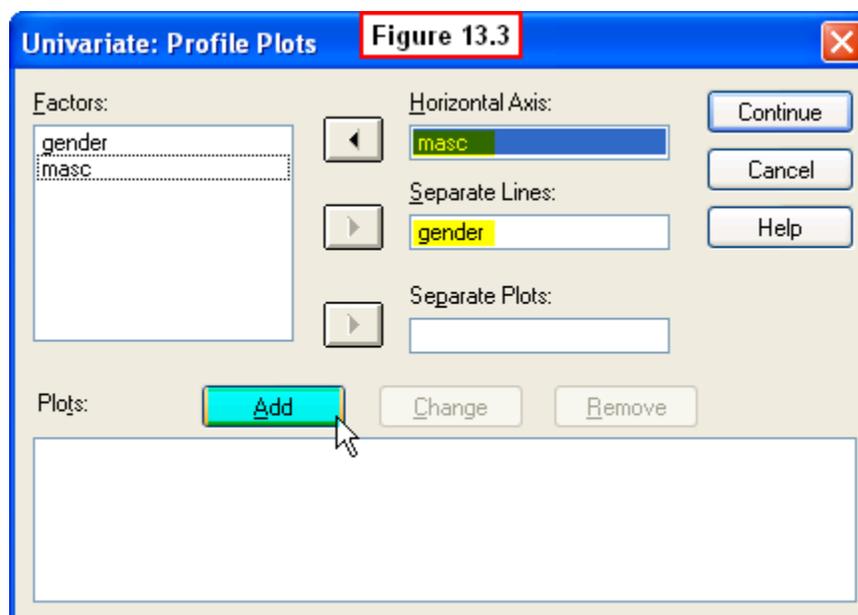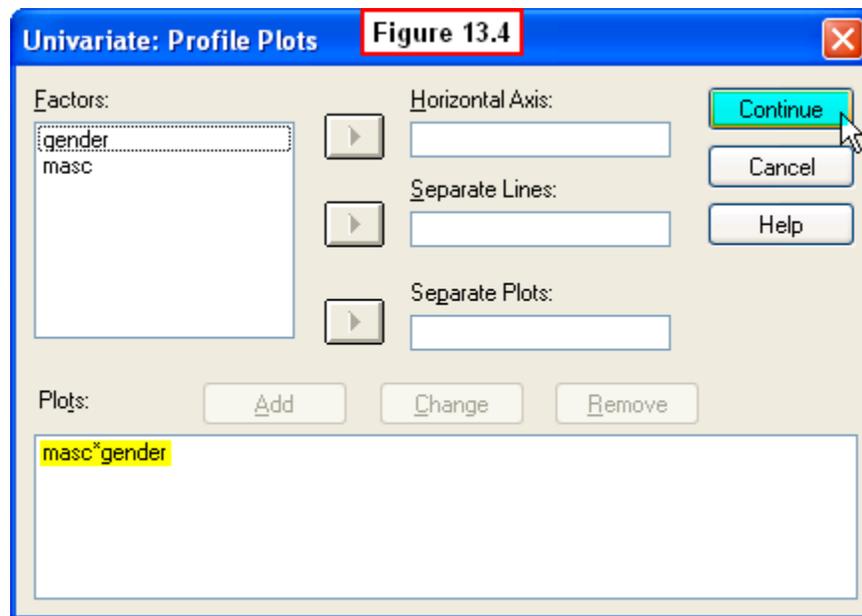
Figure 13.1

A **Univariate** dialog window will appear (Figure 13.2). In the left pane, scroll all the way down and select **masctot** (overall masculinity score), then move it to the **Dependent Variable:** box on the right. Next, in the pane on the left, select **Employee Gender** and **Masulinity Level (masc)** and move them to the **Fixed Factor(s):** pane on the right (recall that SPSS uses the term, **factor**, to refer to the independent variable, so these terms are synonymous). This will generate the ANOVA table when we run the analysis. In the event that the interaction effect is significant, we will want to generate a figure depicting this to aid our interpretation. To do this, click the **Plots...** button on the right.

Figure 13.2

A **Univariate: Profile Plots** window will appear (Figure 13.3). This dialog window wants to know which factor to represent on the horizontal axis and which one to represent by drawing two lines. While technically it doesn't matter which factor is placed where, using separate lines for gender will facilitate comparisons between the sexes. Thus, move **gender** to the **Separate Lines:** box and move **masc** to the **Horizontal Axis:** button.



Figure 13.3

Next, click the **Add** button at the bottom of this window. You will now see **masc\*gender** has been added to the **Plots:** pane at the bottom of this window (Figure 13.4).



Last, click the **Continue** button in the upper right corner. This window will close, returning you to the main **Univariate** dialog window (Figure 13.2). The last thing we need to do is to specify our options. Click the **Options** button on the right of the **Univariate** dialog window. A **Univariate: Options** window will appear (Figure 13.5).
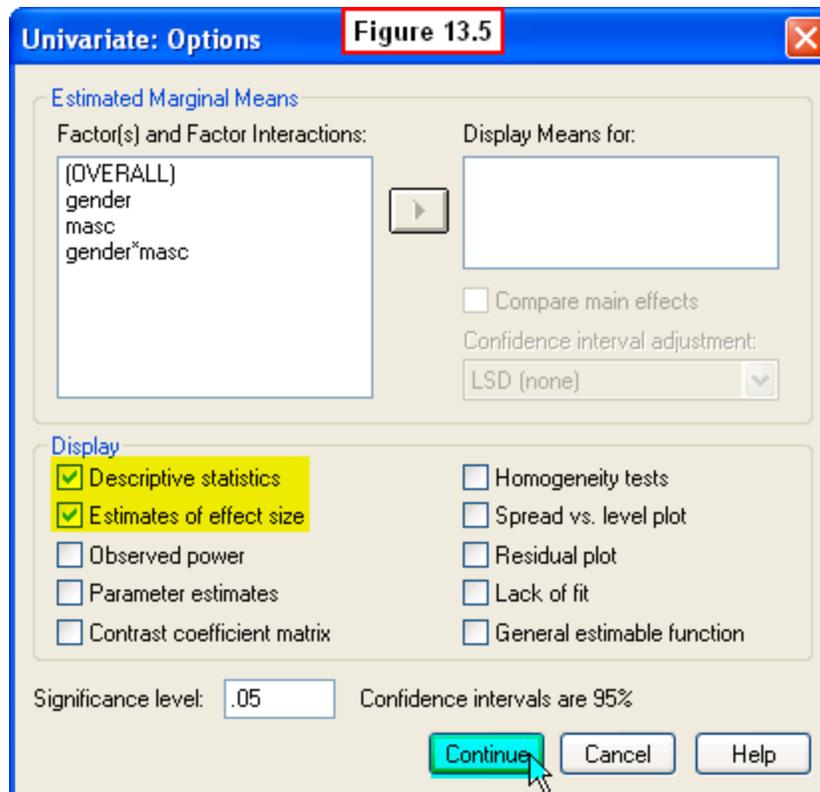
Figure 13.5

First we need to instruct SPSS to calculate the marginal means for the main effects and cell means for the interaction effect. Note that an estimate of these means could be generated by moving the factors and the interaction from the upper left pane to the **Display Means for:** pane on the right. However, these means will only be accurate if there is an equal sample size in all four cells. This is not the case for our data file.

Further, researchers are also frequently interested in other descriptive statistics (e.g., the standard deviations), and this option only generates means. Thus, instead of using the Display Means option, we will use another way to generate the means (and other descriptive statistics) - the **Descriptive statistics** option. To do this, check the **Descriptive statistics** box in the lower left of this dialog window.

In addition to determining whether or not the independent variables yield significant differences in the dependent variable (which will be assessed by the F-values), researchers typically also want to know how **big** the differences are. To do this, check the **Estimates of effect size** box also. Last, click the **Continue** button at the bottom of this window. This will return you to the main **Univariate** dialog window (Figure 13.2). We are ready to run our ANOVA now, so click the **OK** button in the lower left of the this window.

## 3.4 Interpreting the Output

The first table in the output (**Between-Subjects Factors**) displays the sample size for each of the two levels of our two factors. As mentioned, the sample sizes are not equal, which is why we chose to generate the relevant means via the **Descriptive Statistics** option. This information is shown in the second table (Figure 13.6). This table presents the marginal means of overall masculinity scores for the two main effects, the cell means for the interaction, as well as the standard deviations and sample sizes. The most relevant information for our purposes are the means.

**Figure 13.6**
**Descriptive Statistics**

Dependent Variable: masctot

| Employee Gender | Masculinity Level | Mean | Std. Deviation | N |
|---|---|---|---|---|
| Male | Low-Masculinity | 11.8276 | 2.30014 | 29 |
| | High-Masculinity | 30.0617 | .99179 | 81 |
| | Total | 25.2545 | 8.19872 | 110 |
| Female | Low-Masculinity | 10.1525 | 2.08288 | 59 |
| | High-Masculinity | 26.0339 | .31981 | 59 |
| | Total | 18.0932 | 8.11139 | 118 |
| Total | Low-Masculinity | 10.7045 | 2.28518 | 88 |
| | High-Masculinity | 28.3643 | 2.14312 | 140 |
| | Total | 21.5482 | 8.89104 | 228 |

The overall marginal means for male vs. female employees are highlighted in blue, and the marginal means for the two levels of masculinity category are highlighted in green. Last, the four cell means relevant to the interaction are highlighted in yellow. Before we can interpret these means, however, we must first examine the results of the ANOVA displayed in the **Tests of Between-Subjects Effects** table (Figure 13.7).

**Figure 13.7**
**Tests of Between-Subjects Effects**

Dependent Variable: masctot

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | 17460.081a | 3 | 5820.027 | 2691.405 | .000 | .973 |
| Intercept | 75511.824 | 1 | 75511.824 | 34919.583 | .000 | .994 |
| gender | 402.875 | 1 | 402.875 | 186.305 | .000 | .454 |
| masc | 14417.369 | 1 | 14417.369 | 6667.148 | .000 | .967 |
| gender * masc | 68.572 | 1 | 68.572 | 31.710 | .000 | .124 |
| Error | 484.389 | 224 | 2.162 | | | |
| Total | 123811.000 | 228 | | | | |
| Corrected Total | 17944.469 | 227 | | | | |

You can ignore the first two rows of this table and focus on the three rows highlighted in yellow. As we saw in the previous chapter, the most important columns in this table are

the last three, which present the F-values (**F**), the significance levels (**Sig.**), and the effect size indices (**Partial Eta Squared**). These present the tests of the two main effects and the interaction effect. Since **Sig.** = .000 for all of these, we can conclude that both main effects and the interaction effect are statistically significant ($p < .05$).

The significant main effects tell us that there are significant differences in overall masculinity score between men and women and between employees in the low-masculine vs. the high-masculine category. The significant interaction effect indicates that the differences between men and women depend on the masculinity category. As discussed in Chapter 12, the partial eta squared values indicate the size of the differences. We can conclude that the size of the difference in overall masculinity between men and women is moderate, and that the size of the difference in overall masculinity between employees in the low-masculine vs. high-masculine categories is huge! This shouldn't be surprising, since the overall masculinity score was used to create the categorical variable, low-high masculinity (see Chapter 5). In fact, the main reason we used the masculinity category variable was to investigate its possible interaction with participant sex.

Recall that to provide a complete interpretation of the results, the researcher needs to explicitly state what the differences between the means are (see Figure 13.6). Note that since there are only two levels of our two independent variables, it is not necessary to conduct comparisons between the marginal means, i.e., the F-value already tells us that they are significantly different. However, multiple comparisons would be needed to interpret the differences between the four cell means relevant to the interaction effect (see Chapters 11 and 12). For now, however, below is an example of how these results concerning the main effects would be written in APA format:

A significant main effect was obtained for gender, $F(1,224) = 186.31$, $p < .001$. Male employees had significantly higher masculinity scores ($M = 25.26$) than did female employees ($M = 18.09$). This was a moderate difference (Partial Eta Squared = .45). A significant main effect was also obtained for masculinity level, $F(1,224) = 6667.15$), indicating that employees in the high-masculinity category had significantly higher masculinity scores ($M = 28.36$) than did employees in the low-masculine category ($M = 10.71$). This was an extremely large difference (Partial Eta Squared = .97).

Neither of these effects should be surprising in light of the previous analyses in this text concerning gender differences in masculinity. And, of course, the main effect for masculinity category was guaranteed to occur given that the overall masculinity scores were used to create this independent variable. However, there is more to the story given that the interaction effect was significant, $F(1,224) = 31.71$, $p < .001$. Although the effect size is relatively weak (Partial Eta Square = .12), this is actually the most interesting and informative part of this ANOVA, because it indicates that the sex difference in overall masculinity *depends* on the particular masculinity category (low or high).

Interpreting interaction effects and be difficult because it requires the researcher to think of both independent variables at the same time. Due to the complexity of interaction effects, researchers typically graph the four cell means. This visual depiction facilitates comprehension of the interaction, and that is why we generated the graph using the Plots option (see Figure 13.8).



Figure 13.8
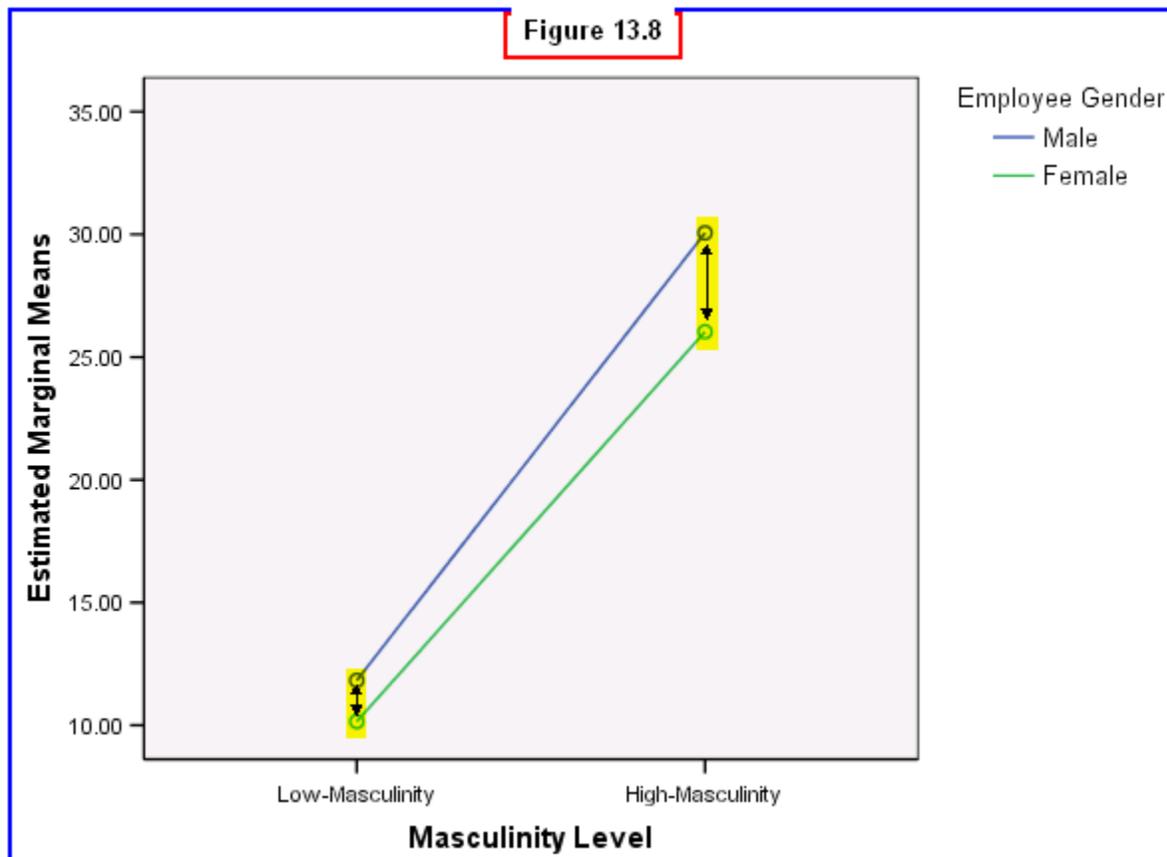
The main effect of masculinity level can be seen for both sexes along the X-axis. That is, both men and women in the high-masculinity category scored higher in masculinity than did their counterparts in the low-masculinity category. However, comparisons between the sexes (the blue line for men, the green line for women) shows that the main effect for gender only appears in the high-masculinity category.

Recall that the cell means are highlighted in yellow in Figure 13.6. Assume that we computed multiple comparisons among these four means (unfortunately, SPSS does not provide this option). Here is how the interaction would be interpreted in APA style based on these comparisons:

In the high-masculinity group, male employees had significantly higher masculinity scores ($M = 30.06$) than did female employees ($M = 26.03$), reflecting the main effect of gender. However, in the low-masculinity group, men did *not* have higher masculinity scores ($M = 11.83$) than did women ($M = 10.15$).

While the implications of this interaction are beyond the scope of this chapter, suffice it to say that two-way factorial ANOVA frequently yields much more interesting information than could be obtained in simpler research designs involving only one factor. Factorial designs allow us to determine qualifying conditions under which simple main effects either do or do not hold. Again, while interpreting interactions can be cognitively taxing, it becomes easier with practice! You will get more practice in the exercise at the end of this chapter!

# 13.5 Chapter Review Video

Review Me!

# 13.6 Try It! Exercises

## 1. Using Two-Way ANOVA: Overall Femininity Differences in Relation to Gender and Femininity Category

Use the procedures described in Section 13.4 to compute a 2 x 2 ANOVA to examine differences in overall femininity (i.e., **FEMTOT**, the dependent variable) in relation to EZ employees' sex (**GENDER**) and femininity category (**FEM**), the two independent variables.

Thus, the four cells of the factorial combination of this **2 (Gender) x 2 (Femininity)** factorial design would result in the following four groups of EZ employees:

(1) Low-feminine males
(2) High-feminine males
(3) Low-feminine females
(4) High-feminine females

After you have completed this two-way ANOVA:

- **Print** your output file to submit to your instructor.
- Answer the following questions for the affiliation variable:
  - Is the main effect of gender significant, and how large is the difference?
  - Is the main effect of femininity category significant, and how large is the difference?
  - Is the GENDER x FEM interaction significant, and what does this mean?
  - 
- **Write an interpretation of the results** along with your answers to the above questions to submit to your instructor.
  - Follow the example of APA write-up in Section 13.4 to write your interpretations of the two main effects.

o Assume you conducted multiple comparisons of the four cell means relevant to the interaction effect. Write an interpretation of the interaction graph following the example in Section 13.4.