

Appendix E

Advanced Statistical Procedures

CHAPTER OBJECTIVES

1. Explain the Law of Parsimony.
2. Summarize a series of different statistical analysis of variance (ANOVA) tests: factorial ANOVA, analysis of covariance (ANCOVA), multivariate ANOVA (MANOVA), and repeated-measures ANOVA.
3. Evaluate a range of advanced statistical tests that analyze relationships: path analysis, structural equation models, factor analysis, and canonical correlations.

In Chapters 15 through 19, we explained in great detail five common statistical procedures that communication researchers use when doing social scientific empirical research. However, many research articles use a wide range of additional statistical tests that you should at least be aware of at this point, so this appendix introduces you to four difference tests and four relationship tests that are currently found in communication journals.

Most of these tests are useful because they allow researchers to be parsimonious. The **Law of Parsimony** states that scientists should look *for the simplest assumption in the formulation of a theory and the simplest test to interpret data*. Although we are not as concerned with the first part of this definition (theory), we are interested in the second part because of its depiction of how researchers should carry out data analysis. As mentioned, the law of parsimony states that researchers should find the simplest way to analyze their data. And we have already conducted one statistical test in this book because of the law of parsimony, but we did not tell you so at the time. When someone conducts a one-way analysis of variance (ANOVA) that has an independent variable with three categories (A, B, and C), he or she could easily run a single one-way ANOVA or three independent *t* tests (A and B, B and C, and A and C). However, when we run three independent *t* tests, our error is compounded for each test, so the likelihood of type I error rises. Conducting a single one-way ANOVA is more frugal than running three *t* tests, so a one-way ANOVA is more parsimonious in this case. In fact, all of the advanced statistics could

be determined using the tests previously described in this text, but the tests described in this appendix are more parsimonious. Additionally, these advanced tests generally perform computations that the less parsimonious tests would not be able to compute, so we are able to ask more advanced questions about our data using a single test.

This appendix is not going to provide you with computer printouts or make you compute things by hand. To learn more about how to run these tests in Statistical Package for the Social Sciences (SPSS), we encourage you to read Mertler and Vannatta (2005), and to see how many of these tests are calculated by hand, please read Bruning and Kintz (1997).

DIFFERENCE TESTS

In this section, we examine four difference tests: factorial ANOVA, analysis of covariance (ANCOVA), multivariate ANOVA (MANOVA), and repeated-measures ANOVA.

Factorial Anova

EXAMPLE

Suppose that you wanted to find out whether there were differences between males and females and differences between political affiliations (Democrat, Republican, other, and not registered to vote) on college students' attitudes toward college. We could run two one-way ANOVAs to answer this question using both sex and political affiliations as separate independent variables (remember that in a one-way ANOVA the independent variable is always a nominal variable) looking for differences in the dependent variable (college students' attitudes toward college). (This example is based on the dataset found on the textbook's website at <https://oup-arc.com/wrench>.)

Unfortunately, one thing we know about statistics is that the more statistical tests a researcher runs to answer a question, the greater the chance that he or she will run into type I error. Running the two one-way ANOVAs is not considered parsimonious for this specific example. In statistics, we want the most frugal and simple (parsimonious) way to answer a single research question. For example, why run three linear regressions when one multiple linear regression will do the same thing? The same thing is true in this example as well. There is no need to run two separate one-way ANOVAs ("sex with attitude toward college" and "political affiliation with attitude toward college"). Instead, we run what is called a two-way or factorial ANOVA.

EXPLANATION

When we originally talked about one-way ANOVAs, we mentioned that the independent variable is called a factor, so in a **factorial ANOVA**, you are simply dealing with more than one factor. In the example here, we have two levels in the first factor

(male and female) and four levels in the second factor (Democrat, Republican, other, and not registered to vote). This design then would be considered a 2×4 factorial ANOVA.

In a factorial ANOVA, there are two types of difference tests. The first type of difference test is called **main effects** because they look for differences between each of the nominal independent variable's categories separately. In our example, the first difference computed in the factorial ANOVA looks at the first independent variable (female vs. male) and the dependent variable (attitude toward college). The second difference computed in the factorial ANOVA looks for differences among the categories in the second independent variable (Democrat, Republican, other, and not registered to vote) and the dependent variable (attitude toward college).

The third difference test examined in a two-way ANOVA is called an **interaction effect** because it is looking for differences in the combination of the two factors. In the example here, the interaction test looks for differences between all the following eight groups ($2 \times 4 = 8$): male Democrat, male Republican, male other, male not registered to vote, female Democrat, female Republican, female other, and female not registered to vote. So not only is a factorial ANOVA more parsimonious, but also it gives you a second type of difference test that cannot be done by simple one-way ANOVAs. You could even throw another factor into this study if you so desired, such as geographical location (north, south, east, or west), and get a $2 \times 4 \times 4$ design. In this case, you would end up with 32 comparisons being examined by the single interaction test and three main effect tests being reported by the factorial ANOVA test.

APA WRITE-UP

Now that we have explained what a factorial ANOVA is, let's see how the American Psychological Association (APA) write-up of a factorial ANOVA would appear in a journal:

A 2×4 ANOVA was conducted to evaluate the effects of biological sex (male and female) and political affiliation (Democrat, Republican, other, and not registered to vote) on college students' attitudes toward college. The means and standard deviations can be seen in Figure E.1. The ANOVA indicated a significant main effect for sex with attitude toward college: $F(1, 304) = 4.45, p = 0.04$; did not indicate a significant main effect for political affiliation with attitude toward college: $F(3, 304) = 0.201, p = 0.90$; and did not indicate an interaction effect for sex by political affiliation with attitude toward college: $F(3, 304) = 0.144, p = 0.94$. The biological sex main effect indicated that females have more positive attitudes toward college than males.

DISCUSSION

In this APA write-up, we see three separate F tests reported in the one factorial ANOVA conducted to analyze this research question: two main effects tests and one interaction test. One of the main effect tests was significant, indicating that females

FIGURE E.1
Factorial ANOVA
Means and SDs

Biological Sex	Political Affiliation	Mean	SD
Male	Democrat	35.86	6.19
	Republican	35.17	6.51
	Other	36.37	6.10
	Not Registered to Vote	34.46	6.02
	Total	35.52	6.29
Female	Democrat	37.78	5.19
	Republican	37.57	5.89
	Other	37.45	5.07
	Not Registered to Vote	37.50	7.48
	Total	37.65	5.54

reported having more positive attitudes toward college than males do. The other main effect test indicated that people in the four political affiliations (Democrat, Republican, other, and not registered to vote) did not differ in their attitudes toward college. Finally, we found that there were no statistical differences between the eight combined groups (male Democrat, male Republican, male other, male not registered to vote, female Democrat, female Republican, female other, and female not registered to vote) and their attitudes toward college. Female Republicans did not differ from males not registered to vote and so on. Overall, the factorial ANOVA allows researchers to answer more complex questions than could be accomplished using a simple one-way ANOVA.

Analysis of Covariance

EXAMPLE

Suppose you want to determine whether males and females differ in their level of communication apprehension (CA). You have a group of college students fill out the Personal Report of Communication Apprehension–24 (PRCA-24). However, you understand that there is a strong negative relationship between an individual's CA and her or his willingness to communicate (WTC). You want to see whether WTC is a confounding variable when determining whether females and males have different levels of CA. (This example is based on the dataset found on the textbook's website.)

EXPLANATION

The **analysis of covariance (ANCOVA)** extends the one-way ANOVA discussed in Chapter 17. Simply put, a **covariate** is a variable related to the dependent variable that can cause the participants' scores on the dependent variable to be skewed or altered, so the ANCOVA readjusts the dependent variable scores to prevent this skewing from occurring. More specifically, a researcher may opt to use an ANCOVA for two reasons.

First, the researcher may want to exclude the effect of a given independent variable. For example, maybe you are conducting a study looking at sex differences and perceptions of political speeches. However, when you collect your data, you find a significant difference between the ages of the females and males in your sample. To prevent age from becoming a factor while you are looking at the sex differences, you decide to use age as a covariate to exclude the effect that it may have on your participants' perceptions of the political speeches.

The second reason that a researcher may opt to use an ANCOVA is when two variables are strongly related to each other. In this situation, if someone is looking for a difference in a dependent variable, it is possible that the researcher will end up finding a difference in the variance accounted for by the dependent variable and another variable (the covariate). For this reason, the researcher may want to partial out the variance of the covariate so that he or she can only look for a difference in the variance not accounted for by the covariate.

In our sample ANCOVA, we will test the second use of an ANCOVA. The purpose of our example ANCOVA is to allow a researcher to determine whether a difference lies between groups (female and male) on a dependent variable (CA) after the dependent variable has been mathematically adjusted for differences associated with one or more covariates (WTC). The basic test analyzed in an ANCOVA is similar to the one-way ANOVA in that both look for differences between groups. The ANCOVA, however, increases the power of the F test for a main effect or interaction by removing the predictable variance associated with the covariate (WTC) from the error term for the F test.

APA WRITE-UP

Let's see how the APA write-up of an ANCOVA would appear in a journal:

The purpose of this research question was to examine the possibility of a significant difference in CA based on biological sex (male and female) while controlling for an individual's WTC. A one-way ANCOVA was conducted using biological sex (male and female) as the independent variable, CA as the dependent variable, and WTC as the covariate. A significant relationship was found between the dependent variable (CA) and the covariate (WTC): $F(1, 295) = 86.41, p < 0.001, \eta^2 = 0.23$. Furthermore, a significant difference was found between males ($M = 62.68, SD = 15.43$) and females ($M = 66.41, SD = 19.25$) on CA: $F(1, 295) = 8.78, p = 0.003, \eta^2 = 0.03$.

DISCUSSION

These results indicate a significant difference between males and females and their levels of CA. In Chapter 16, we ran the same test using an independent samples t test procedure and found no differences between males and females and CA. Because we

used the same data off the textbook's website sample to ask this question, what caused the difference to appear now?

The correction of CA that occurred by the covariate WTC is what ultimately caused this difference to occur. Note that two separate F tests are reported in the ANCOVA. The first F test indicated that there was a significant relationship between the dependent variable (CA) and the covariate (WTC). The eta-square ($\eta^2 = 0.23$) indicates that approximately 23% of the variance in CA can be accounted for by WTC. In an ANCOVA, you can think of eta-square as being similar to R^2 in a regression.

The second F test examined the differences between males and females on CA. This test indicated that females do have slightly higher levels of CA than males in the sample. Note again that eta-square is reported for this F test. Although males and females differ in their CA in this sample, biological sex only accounts for 2.9% of the variance, which is not much when you think about it.

MULTIVARIATE ANOVA

Example

A **multivariate ANOVA (MANOVA)** allows a researcher to examine differences using one or more nominal independent variables with one or more dependent variables. Perhaps a researcher wants to see whether males and females differ in their levels of ethnocentrism, but also wants to see whether males and females differ in their levels of WTC with strangers. This could be answered using two one-way ANOVAs (sex with ethnocentrism and sex with WTC with strangers), but again, the more tests you use, the greater the chance you will end up with type I error. So, to be parsimonious (simple and frugal), you would need to conduct a one-way MANOVA.

EXPLANATION

A MANOVA, as discussed earlier, is considered a **multivariate test** because you have multiple dependent variables (ethnocentrism and WTC with strangers), while a one-way ANOVA is considered a **univariate test** because you have one dependent variable. Often, the dependent variables analyzed in a MANOVA are different measures of the same phenomenon. In our example, we would think that people who have higher levels of ethnocentrism would be less willing to communicate with strangers, so these two variables could theoretically be related. The two do not need to be related, but they should share a common conceptual meaning and some degree of linearity (remember that all ANOVA tests are general linear model tests). In essence, the dependent variables in a MANOVA should go together in a way that makes sense, so you would not metaphorically put both apples and oranges as dependent variables in the same MANOVA. (Again, the purpose of this chapter is not to explain all of the mathematical aspects of the MANOVA procedure, but to introduce you to the basic concept. For this

reason, we strongly urge you to read more information about the MANOVA elsewhere since it is the basis of a number of advanced statistical procedures.)

APA WRITE-UP

Let's see how the APA write-up of a MANOVA would appear in a journal:

The goal of this research question was to determine whether there was a difference between males and females on ethnocentrism and WTC with strangers. To analyze this question, a one-way MANOVA was calculated using biological sex (female and male) as the independent variable and the participant's scores for ethnocentrism and WTC with strangers as the dependent variables. Box's test (Box's $M = 3.58$) reveals that equal variances can be assumed: $F(3, 31993081) = 1.19, p > 0.05$; so Wilks's lambda (Λ) will be used as the test statistic. The Wilks's lambda criteria indicate significant group differences in biological sex for the overall model: Wilks's $\Lambda = 0.975, F(2, 302) = 3.94, p = 0.02$, multivariate $\eta^2 = 0.03$. Univariate ANOVA results were interpreted using alpha at 0.05. Results reveal that males ($M = 38.43, SD = 9.08$) and females ($M = 35.63, SD = 8.08$) significantly differed on ethnocentrism: $F(1, 303) = 7.91, p = 0.005$, partial $\eta^2 = 0.03$. Results also revealed that males ($M = 47.86, SD = 24.86$) and females ($M = 48.86, SD = 26.67$) did not significantly differ on WTC with strangers: $F(1, 303) = 0.11, p = 0.74$.

DISCUSSION

Let's start an analysis of these results by remembering the purpose of this research question. The goal was to use one independent variable (sex) to examine two dependent variables (ethnocentrism and WTC with strangers) using one test. The first result reported in the MANOVA is the test that measures for the equality of variances assumption. Like in other tests we have examined, the MANOVA has a basic assumption that the variances from the groups being examined are equal. Box's M test examines the equality of variances assumption and determines whether we can use Wilks's lambda (if Box's M is not significant—we accept the equality of variances assumption) or Pillai's trace (if Box's M is significant—we reject the equality of variances assumption). Wilks's lambda or Pillai's trace are two multivariate tests that examine statistical significance of the whole model (both independent variables and dependent variables) In this example, Box's M was not significant, so we were able to utilize Wilks's lambda.

The multivariate test is then followed by a series of univariate tests (F tests) for the independent variable (sex) with every dependent variable (ethnocentrism and WTC with strangers). The multivariate test essentially lets us know that a difference exists between the independent variable and dependent variables, but not where the

difference actually is. In our example, the overall multivariate test was significant, so we needed to examine the univariate statistics (sex with ethnocentrism and sex with WTC with strangers). Here, males had higher levels of ethnocentrism than females, but there was no difference between females and males in their WTC with strangers. However, once again, biological sex only accounted for a small amount of the variance in ethnocentrism (3%).

Repeated-Measures Anova

EXAMPLE

Suppose you want to determine whether taking a public speaking course actually decreases a person's level of CA. One possible way to determine whether a person's level of CA decreases over the course of a public speaking class would be to test his or her CA level at the beginning of the course, test it again halfway through the course, and test it a third time at the end of the course. You could then calculate three paired *t* tests to determine this research question (Time 1 to Time 2, Time 2 to Time 3, and Time 1 to Time 3), but again, the more tests you run, the more error your findings will have. So, to avoid increasing your type I error, you can run a procedure called a repeated-measures ANOVA. (This example is hypothetical because the data on the textbook's website are not set up to answer a repeated-measures ANOVA question.)

EXPLANATION

A **repeated-measures ANOVA** allows a researcher to determine whether differences occur in a variable over time. In this example, we would measure these differences occurring over time by having students fill out the PRCA-24 at the beginning of the semester/quarter, in the middle of the semester/quarter, and at the end of the semester/quarter. By having the students fill out the survey all three times, we have a way of mapping what happens to CA levels throughout the course of a public speaking class. What we are testing is the null hypothesis that Time 1 = Time 2 = Time 3.

APA WRITE-UP

Let's see how the APA write-up of a repeated-measures ANOVA would appear in a journal:

The goal of this research question was to determine whether a person's level of CA changes over the course of a public speaking class. A one-way, within-subjects, repeated-measures ANOVA was conducted using three scores for CA taken at the beginning of the course ($M = 65.11$, $SD = 15.49$), the middle of the course ($M = 64.61$, $SD = 14.36$), and the end of the course ($M = 64.09$, $SD = 15.93$). This study found no significant differences between the measurements: Wilks's $\Lambda = 0.985$, $F(2, 125) = 0.96$, $p = 0.39$.

DISCUSSION

In this APA write-up, we learned that there was not a significant difference among the beginning, middle, and end of the public speaking course and individual levels of CA. Although the repeated-measures ANOVA is considered a univariate test, it still relies on a multivariate test (Wilks's Λ) to determine the overall significance of the model. If the test had been significant, then we could have used paired t tests to determine where the actual difference existed (Time 1 to Time 2, Time 1 to Time 3, or Time 2 to Time 3).

It is also easy to make the repeated-measures ANOVA even more complicated by adding what is called a between-groups aspect to the test. For example, perhaps you want to see whether male and female levels of CA changed over the course of a public speaking course, which would give you one group within (everyone and CA) and one group between (males and females). More and more layers can be added to the repeated-measures ANOVA by adding multiple dependent variables, which would create a repeated-measures MANOVA.

RELATIONSHIP TESTS

In this section, we will examine four relationship tests: path analysis, structural equation modeling, factor analysis, and canonical correlations.

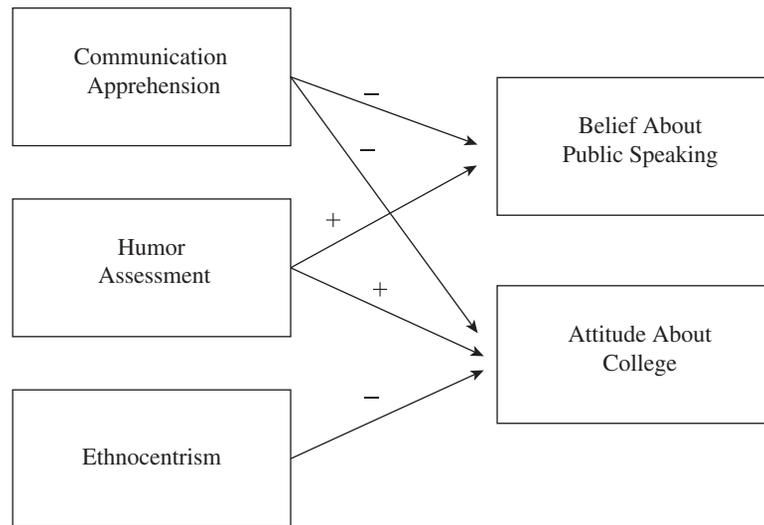
Path Analysis

EXAMPLE

Causal relationships are one area that social scientific researchers are hesitant to discuss. However, one technique that has been developed to examine causal relationships is the path analysis.

Perhaps you wanted to determine the causal relationships among CA, ethnocentrism, humor assessment (HA), attitudes toward college, and people's belief that everyone should be required to take public speaking in college. You start doodling on a piece of paper and come up with a theoretical explanation for why some people like college and others do not and why some people think everyone should take public speaking and others do not (Figure E.2). You believe that people with higher levels of CA are less likely to like college and are less likely to believe that everyone in college should take public speaking (represented by the minus sign next to the lines in the figure). Furthermore, you believe that people who are more humorous will enjoy college more and will believe that everyone in college should take public speaking (represented by the plus sign next to the lines). Finally, you believe that people who are ethnocentric are going to have more negative attitudes about college, but you do not think that there will be a relationship between an individual's level of ethnocentrism and her or his belief that everyone should take a public speaking class in college.

FIGURE E.2
Path Analysis Example



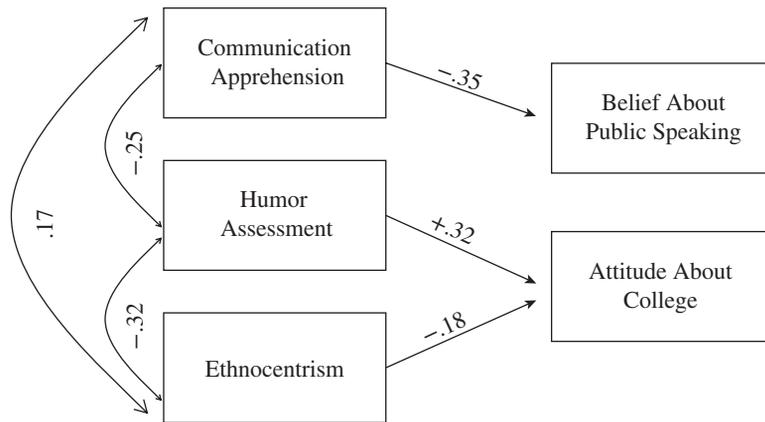
EXPLANATION

In a **path analysis**, two different types of variables must be examined: endogenous and exogenous. **Endogenous variables** are explained by one or more of the other variables in the model (e.g., belief about public speaking and attitude about college). **Exogenous variables** are taken as a given, so the model does not try to explain them (e.g., CA, ethnocentrism, and HA). To conduct a path analysis, we calculate a series of multiple linear regressions using the exogenous variables as the independent variables (CA, ethnocentrism, and HA) and the endogenous variables as the dependent variables (belief about public speaking and attitude about college). From these regressions, you simply report the beta weights for each regression on the picture.

APA WRITE-UP

Let's see how the APA write-up of a path analysis would appear in a journal:

A path analysis was conducted to determine the causal effects among the variables CA, ethnocentrism, HA, belief that everyone should take public speaking in college, and attitude about college. Before the analysis, an initial model was created (Figure E.2). This model was not consistent with the empirical data. Two of the correlations exceeded a difference of 0.05, so the nonsignificant paths were removed from the model. Thus, a revised model was generated (Figure E.3). Approximately 12% of the variance in a person's belief that taking public speaking in college should be required can be accounted for by the model, and approximately 17% of the variance in a person's attitude about college can be accounted for by the model.



Betas were significant at $p < .005$.

FIGURE E.3
Path Analysis
Calculated

DISCUSSION

Overall, you can think of the path analysis as a pictorial way of presenting information about multiple regressions. The numbers shown to the left of the exogenous variables in Figure E.3 are Pearson product-moment correlations that were calculated among the exogenous variables. The numbers above the lines originating from an exogenous variable pointing toward an endogenous variable are beta weights calculated during multiple regressions and indicate whether there is a positive or negative relationship. What we can tell from these findings is that an individual's level of CA negatively relates to her or his belief that all college students should take public speaking. We also learned that there is a positive relationship between an individual's HA and positive attitudes about college and that there is a negative relationship between an individual's ethnocentrism and positive attitudes about college. People who are more humorous feel more positively about college, whereas people who are more ethnocentric feel less positively about college.

Structural Equation Modeling

EXAMPLE

As we have noted throughout this appendix and in previous chapters, the dataset collected for this textbook has indicated that there is a negative relationship between an individual's level of CA and her or his belief that all college students should be required to take public speaking. Suppose that you wanted to examine this relationship in light of the four subscales that allegedly make up CA. We could run a path analysis as in the previous section, but a path analysis relies on a number of linear and multivariate regressions, which simply is not parsimonious. However, a newer statistical technique has been created called structural equation modeling.

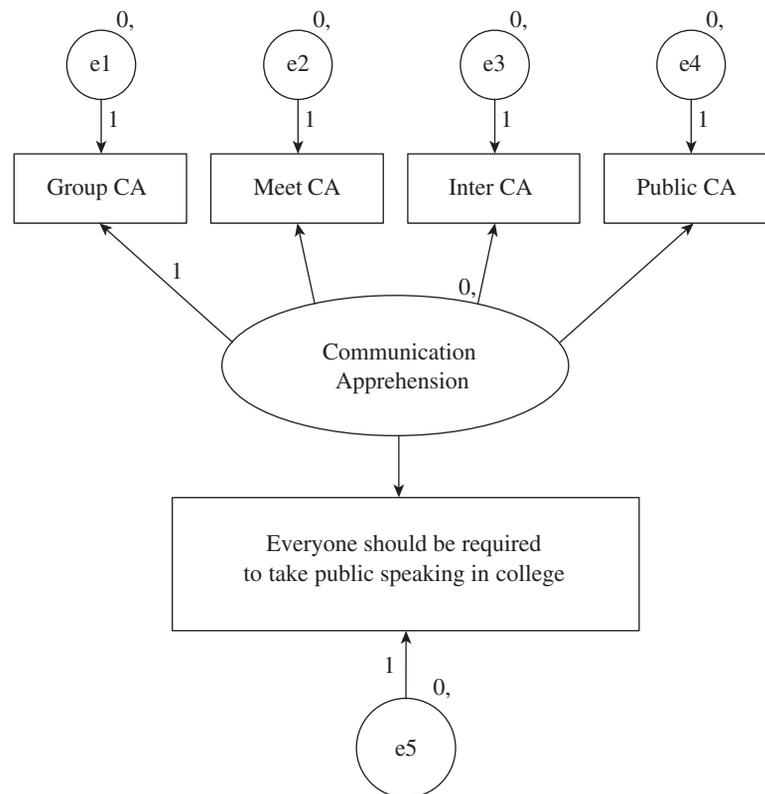
In this example, we want to see how well the four subscales of CA create this variable known as “communication apprehension” and how this variable relates to a participant’s belief that all students in college should be required to take public speaking. And to do this, we will use structural equation modeling. Figure E.4 shows how this question would look pictorially.

EXPLANATION

To further understand this research question, we will examine what a structural equation actually is. **Structural equation modeling** is similar in purpose to path analysis; however, the calculations are considerably more difficult but mathematically more meaningful.

Ultimately, structural equation modeling is concerned with observed and latent variables. An observed variable can be an observation that a researcher directly collects (self-reports on a survey, scores on an achievement test, coded responses to interview questions, etc.). Latent variables are those that are not directly measured, but that we believe our measurements help us understand. For example, we believe that CA is a

FIGURE E.4
Structural Equation
Model Hypothesis



product of four subscales (group CA, meeting CA, interpersonal CA, and public CA). Therefore, we do not measure a variable called “communication apprehension,” but rather measure the four subscales together and then combine their results to create an overall score of communication apprehension. In essence, “communication apprehension” is the latent variable being measured by the four subscales (group CA, meeting CA, interpersonal CA, and public CA). As shown in Figure E.5, we measured each of the four subscales in a rectangular box, indicating that these variables are directly measured by the researcher. However, the variable listed as “Communication Apprehension” is an oval, indicating that it is a latent variable.

In structural equation modeling, as in path analysis, there are two types of variables discussed: exogenous and endogenous. Exogenous latent variables are similar to independent variables because they can account for some of the variance in other variables in the model. Endogenous latent variables are similar to dependent variables because they are influenced by the exogenous variables. Here, CA is an example of an exogenous variable because it attempts to account for some variance in an individual’s belief that all college students should be required to take public speaking.

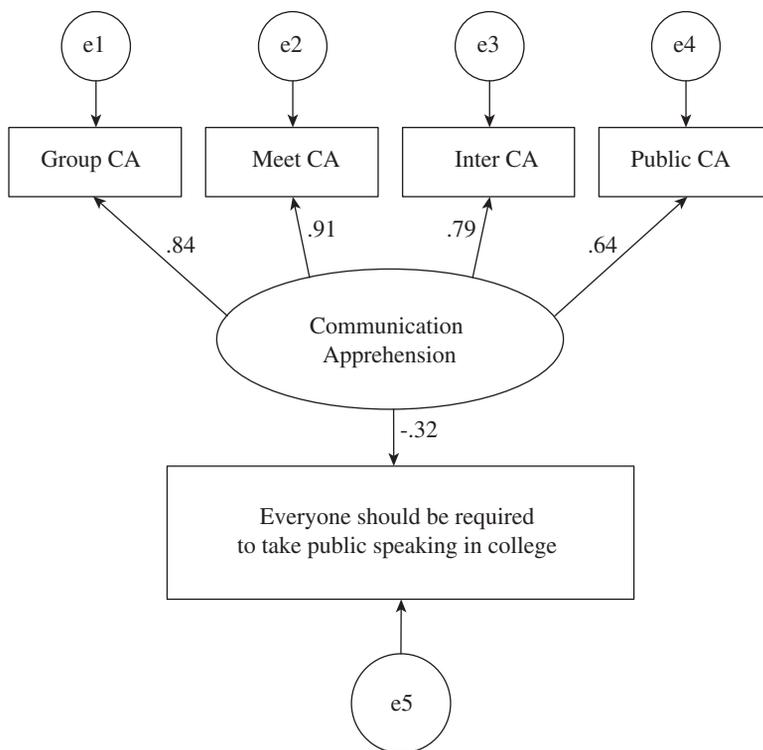


FIGURE E.5
Calculated Structural
Equation Model

APA WRITE-UP

Let's see how the APA write-up of our structural equation modeling would appear in a journal:

Using structural equation modeling, the relationships between CA, a latent variable with four indicators (group CA, meeting CA, interpersonal CA, and public CA), and an individual's belief that public speaking be a required course in college were examined. The hypothesized model is presented in Figure E.4. Circles represent latent variables, and rectangles represent measured variables. Absence of a line connecting variables implies lack of a hypothesized relationship. Results indicated that the proposed structural model was problematic: $\chi^2(5, N = 325) = 24.20, p < 0.001$. However, because this model had more than 200 participants, other goodness-of-fit indices are necessary. All of the goodness-of-fit indices far exceeded the recommended levels: normed fit index (NFI) = 0.99, comparative fit index (CFI) = 0.99, relative fit index (RFI) = 0.98, incremental index of fit (IFI) = 0.99, and the Tucker–Lewis index (TLI) = 0.99. All of the indices of fit were over the 0.95 mark, which indicates that the model proposed is a superior fit. The final structural equation model can be seen in Figure E.5.

DISCUSSION

The goal of this research question was to see whether the four subscales of the PRCA-24 actually created a latent exogenous variable called “communication apprehension” and then to determine whether this latent exogenous variable was related to an individual's belief that everyone in college should be required to take public speaking. To examine these findings, we first talk about the statistics involved and then examine the structural equation model in Figure E.5.

The first statistic that is reported is a chi-square test to determine whether the model is a good fit. When examining the chi-square test that is conducted for a structural equation model, two things should really be examined: the chi-square statistic and the degrees of freedom (df). In an ideal test, the df is less than 5, and the closer the df is to the chi-square statistic, the stronger your model is said to be. However, the chi-square test of goodness-of-fit in the structural equation model is not always the best way to determine whether you have a strong model, especially if a study has more than 200 participants (Bollen & Long, 1993). If your study has more than 200 participants, your chi-square is almost always going to be significant, which is an indication of poor fit. For this reason, a number of other indices have been developed such as the ones reported in our APA write-up for this example: normed fit index (NFI), comparative fit index (CFI), relative fit index (RFI), incremental index of fit (IFI), and the Tucker–Lewis index (TLI). Without going into the mathematical reasoning for each of these goodness-of-fit tests, each explains to us mathematically whether our model makes

sense the way we designed it. All of these goodness-of-fit indices can range from 0 to 1, with scores above 0.95 generally being seen as acceptable (Byrne, 2010). Overall, all the goodness-of-fit indices indicate that the model we proposed in Figure E.5 was a good model to explain the data.

Next to the goodness-of-fit indices, the most important part of the structural equation model is the standardized estimates seen in Figure E.5. The standardized estimates can be seen as beta weights when determining the linearity of the relationship. In essence, all four subscales contribute to create the latent variable “communication apprehension” above 0.64, which indicates that the four subscales do actually measure the latent variable effectively. There is also a negative relationship between the latent exogenous variable “communication apprehension” and an individual’s belief that all students should be required to take public speaking in college. Although this resembles the path analysis results from the previous section, the biggest difference here is that the structural equation model is more parsimonious and accounts for possible error, which a series of multiple linear regressions cannot. Attached to every observed variable (the rectangles in the model), you will note that there is an error term (a small circle with the letter “e” followed by a number from 1 to 5). Again, to avoid going into the computational mathematics, understand that the error associated with observed variables (four CA subscales and belief about public speaking) accounts for measurement error, which can come from one of two places: random error and error uniqueness. Random error is simply random measurement error that occurs as a result of measuring things. Error uniqueness indicates that there was a form of error unique to a particular variable that is considered to be a nonrandom measurement error.

Factor Analysis

EXAMPLE

Factor analysis is an extremely important technique to master if you want to understand how to create survey research measures like the PRCA-24, WTC, Sociocommunicative Orientation Scale, or any of the other scales we have used in this textbook. One of the problems when creating a new scale is that you never know whether or not your new scale is actually measuring what you say it should be measuring. Although the primary way we examine problems like this is through analyzing face or criterion validity (see Chapter 8), another way we can attempt to understand whether a scale is measuring what we say it is measuring is through a factor analysis (construct or factorial validity).

In 2001, Richmond, Wrench, and Gorham created a new research scale to measure an individual’s use of humor during interpersonal interactions. The scale itself consists of 16 Likert-type items using a five-point scoring system from 1 (*strongly disagree*) to 5 (*strongly agree*). How do we know that these 16 items actually measure anything? As discussed in Chapter 8, the HA instrument has an alpha reliability of 0.91 ($M = 63.24$,

$SD = 9.53$). So we know the scale is reliable, but do the 16 items in the HA actually measure just one thing? To determine whether a set of scale items (like the 16 items on the HA) is measuring one concept or multiple concepts, we conduct a factor analysis.

EXPLANATION

A **factor analysis** is a technique that enables researchers to determine variation and covariation among research measures. For example, suppose we had two items being measured on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The first item in the scale reads “People are innately good,” and the second item reads “Cats are the best animals.” You give these two questions to a large sample and get the results back. Assume that you get an alpha reliability of 0.80, which is considered good. So your scale is reliable, but what is the scale really measuring?

Chances are your scale is not measuring one coherent concept, but rather two concepts that happen to be related to each other. The purpose of a factor analysis is to determine how many different concepts are being measured by a set of questions on a research scale. If you recall the discussion from Chapter 9 on creating surveys, we mentioned that a single research survey can only measure one thing. In the instance of the HA, the purpose of the scale is to examine only an individual’s use of humor during interpersonal interactions. The scale does not measure an individual’s ability to use humor, an individual’s sense of humor, or anything else. The scale has a single purpose, and all 16 scale items were written to reflect that conceptualization.

Two basic types of factor analyses can be calculated. The first is called an exploratory factor analysis. An exploratory factor analysis is when a researcher has a set of scale items and wants to determine how many concepts the set of scale items is measuring. If you have 20 scale items, it is theoretically possible that each scale item is measuring a completely different concept and that there is no unity between the items. However, it is also possible that all 20 items are only measuring a single variable. Most communication scales tend to measure between one and five distinct concepts. For example, of the scales used in this textbook, only one scale measures two distinct variables. Know which scale it is? If you guessed the Sociocommunicative Orientation Scale, then you were correct. The Sociocommunicative Orientation Scale measures the degree to which an individual is assertive and the degree to which an individual is responsive.

Each concept measured by a scale is called a factor. A factor analysis helps researchers group the individual scale items into coherent sets of concepts called factors. Sometimes you may think you have written 30 strong scale items to measure one single concept—say, cooperative communication—only to find out that your 30 items are actually measuring three different concepts—cooperative behavior, competitive behavior, and trust. Other times, the 30 items will hold strong and clearly measure one variable—cooperative communication.

The second type of factor analysis is called confirmatory factor analysis (CFA). This analysis is when a researcher uses a factor analysis to make sure that a previously determined factor structure is consistent with present results. For example, in the last example looking at structural equation modeling, we used structural equation modeling to determine whether the four CA subscales actually measure a variable called “communication apprehension.” Although not an exact example of CFA, this was similar to what a CFA actually does, except that in a true CFA, we would have also tested whether each of the individual scale items clearly helps in creating the individual subscales (meeting CA, group CA, interpersonal CA, and public CA).

Another area of factor analysis that we must explain is what we refer to as extraction methods and rotation. Extraction refers to the specific type of factor analysis that an individual is conducting. The most basic factor analysis extraction method is called a principal component factor analysis. Although we will not discuss the mathematical details of a principal component analysis, you should know that there are other forms of factor analysis, each of which is best used in different circumstances (unweighted least squares, generalized least squares, maximum likelihood, principal axis factor, etc.), and that each factoring method contains differing mathematical reasoning and computations. For a good explanation of the different types of extraction methods and when to use one rather than another, we strongly encourage you to read Tabachnick and Fidell (2001) or Grimm and Yarnold (2000a, 2000b).

Still another important concept in the world of factor analysis is factor rotation. When researchers calculate factor analyses, it is often difficult to ascertain a clear factor structure for the scale items involved. Trying to find factor structures is akin to looking at Impressionist art. Sometimes you have to step back from the piece to see what is actually on the canvas, or maybe you have to turn the canvas sideways to see what the artist intended. Examining factor structures often involves a certain amount of scrutiny and manipulation. Imagine that we have all of the data points on a graph. If we look at the graph straight on, we may not see much, but if we turn the page slightly, a clear linear structure may appear on the paper. When we rotate the data points along either the x or the y axis, we have rotated the factor structure. This is a simplistic view of factor analysis rotation, but it gives you a basic idea of what happens when a researcher must rotate a factor to determine the actual structure of the factor analysis.

APA WRITE-UP

Let’s see how the APA write-up of a factor analysis would appear in a journal:

The HA instrument was developed to measure an individual’s predisposition to use humor as a communicative tool during interpersonal situations. The HA is a 16-item, self-report measure that uses a 5-point Likert format ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

The dimensionality of the 16 items for the HA was analyzed using an unrotated principal component factor analysis. Four criteria were used to determine the number of factors to rotate: sampling adequacy, the a priori hypothesis that the measure was unidimensional, the scree plot, and the interpretability of the factor solution. To examine sampling adequacy, Kaiser's measure of sampling adequacy (MSA) was used. The MSA obtained was 0.92, which is considered "marvelous" for conducting a factor analysis (Kaiser, 1974). The scree plot indicated that our initial hypothesis of unidimensionality was correct. The principal component analysis revealed a strong primary factor. The factor loadings can be seen in Figure E.6.

DISCUSSION

First, let's remember the basic research question involved in calculating the factor analysis. Our goal was to determine whether the 16 scale items created by Richmond et al. (2001) measure an individual's use of humor in interpersonal interactions using one factor. To analyze the results in this example, we start by discussing some of the main features mentioned in the APA write-up.

FIGURE E.6
Factor Analysis Humor
Assessment

1. I regularly communicate with others using humor.	.62
2. People usually laugh when I make a humorous remark.	.69
3. I am not funny or humorous.	-.66
4. I can be amusing or humorous without having to tell a joke.	.63
5. Being humorous is a natural communication orientation for me.	.70
6. I cannot relate an amusing idea well.	-.68
7. My friends would say that I am a humorous or funny person.	.68
8. People don't seem to pay close attention when I am being funny.	-.62
9. Even funny ideas and stories seem dull when I tell them.	-.67
10. I can easily relate funny or humorous ideas to the class.	.64
11. My friends would say that I am not a humorous person.	-.71
12. I cannot be funny, even when asked to do so.	-.68
13. I relate amusing stories, jokes, and funny things very well to others.	.64
14. Of all the people I know, I am one of the "least" amusing or funny persons.	-.65
15. I use humor to communicate in a variety of situations.	.68
16. On a regular basis, I do not communicate with others by being humorous or entertaining.	-.61

Kaiser's Measure of Sampling Adequacy (MSA)

The way to interpret a Kaiser's MSA is to use the system Kaiser (1974) created for determining whether the dataset is appropriate for the factor analysis:

- 0.9 and above is marvelous
- 0.8–0.9 is meritory
- 0.7–0.8 is middling

- 0.6–0.7 is mediocre
- 0.5–0.6 is miserable
- 0.5 and lower is unacceptable

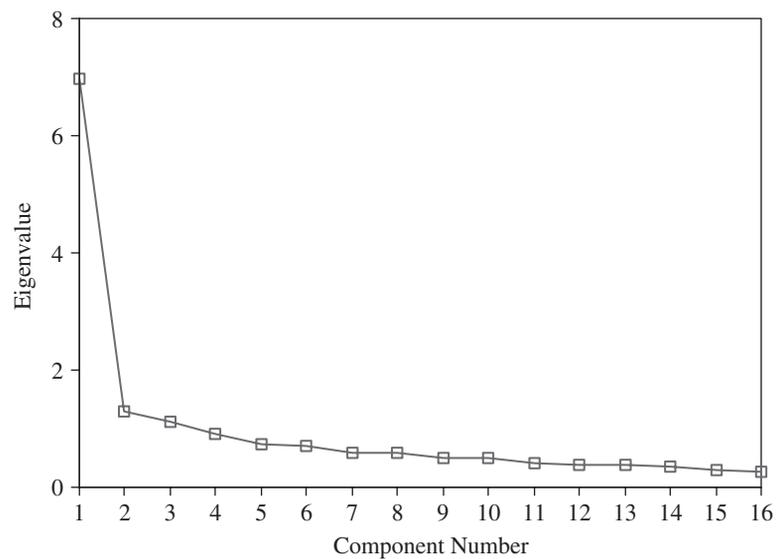
As a general rule, Kaiser's MSA must be at least 0.6 or above, but most journals expect a Kaiser's MSA of 0.8 or above, with preference given to scores above 0.9.

First, the researchers utilized an unrotated principal component analysis. In this example, there was no need to rotate the principal component analysis because the factor structure was clear without a rotation.

Second, Kaiser's MSA was used to determine whether the sample was adequate for performing a factor analysis. The Kaiser's MSA is a tool to determine whether your sample is sufficient (robust) enough to perform the factor analysis on the number of items in a scale.

Third, a scree plot was used to determine whether the model was unidimensional (contained only one factor). A scree plot is a plot of eigenvalues. Without going into detail, an eigenvalue describes the variance of the set of data points in a multivariate space that has one axis for each variable (Tabachnick & Fidell, 2001). And you probably just hit your head on the table and thought that sentence makes absolutely no sense whatsoever. Do not fear, however; we are here to help. In any dataset, the maximum number of the eigenvalues is equal to the number of scale items being factor analyzed. For example, the HA has 16 scale items, so the combined eigenvalue for the scale is 16. Then the question becomes can we minimize the eigenvalue sum (16) so that only one eigenvalue is above the number 1? For example, in our survey, the first component had an eigenvalue of 6.966 (Figure E.7), which accounted for 43.54% of the variance. In other words, one factor accounted for 43.54% of the variance in the HA scale. Ideally, for each eigenvalue above 1.0 in a factor analysis, you will extract one factor. In our example here, Figure E.7 shows that there were actually three factors that had eigenvalues above 1.0, with the others being component 2 (eigenvalue of 1.30) and component 3 (eigenvalue of 1.12). However, researchers have shown that relying on the eigenvalues alone for extracting factors can be misleading. For this reason, it is also encouraged that you examine a scree plot when extracting factors.

FIGURE E.7
Scree Plot for Humor
Assessment



In Figure E.7, the scree plot is actually the plot of the eigenvalues themselves. The word “scree” is actually a geologic term referring to the debris that collects toward the bottom of a rocky slope. If you look at Figure E.7, the figure is kind of reminiscent of a rocky slope. The scree plot has a clear downward trajectory that then levels off and flattens out to the right. The first data point is placed at 6.966, and the second point is plotted at 1.30. To determine how many factors to extract using a scree plot, it is best to use the concept of looking at the elbow.

On a scree plot, there will always be at least one eigenvalue above 1.0. And in our example, the first eigenvalue 6.966 is considerably larger than the rest of the eigenvalues. In fact, if you look at the scree plot, it is hard to clearly delineate the other 15 eigenvalues from each other as far as how they are plotted. This leveling off of a scree plot is called the elbow of the plot. Any eigenvalues separated from the elbow are considered actual factors, and those eigenvalues in the elbow or the tail of the scree plot are considered residuals of that primary factor.

Figure E.8 shows the scree plot associated with the factor analysis of the Sociocommunicative Orientation Scale. Remember, this scale has two factors—assertiveness and responsiveness. You will note that in Figure E.8 there are two clear eigenvalue plots—one at 6.17 and one at 3.36—and then you see a big dip and the creation of the elbow of the scree plot. This is an example of what it looks like when a factor analysis has two clear factors that must be extracted from the data.

Once we have determined how many factors must be extracted from the data, we can examine how the values load. In this example, we only extracted one actual factor, which can be seen in Figure E.6. When examining the factor loadings (number in the

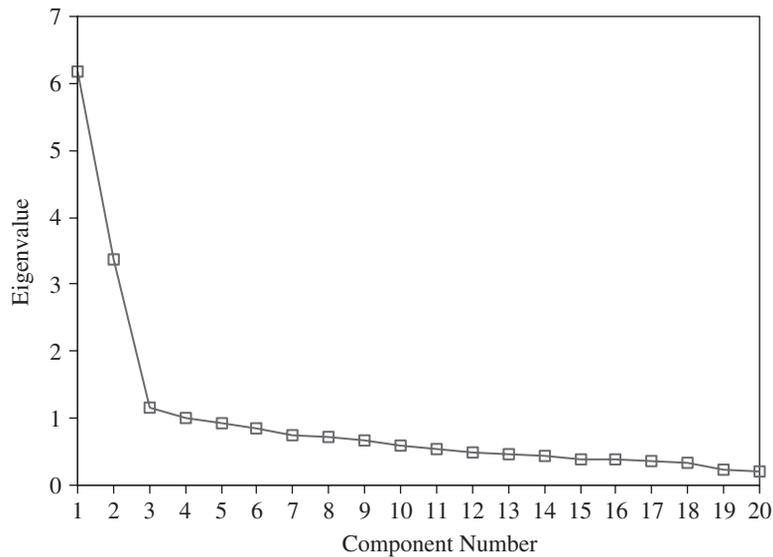


FIGURE E.8
Scree Plot for
Sociocommunicative
Orientation

right column in Figure E.6), you will note that all the loadings are 0.61 and higher. Even if a number has a minus in front of it, it is considered a high loading on the factor analysis. If you look carefully, you will see that all of those items that have negative factor loadings are items that are reverse coded on the scale. However, a factor analysis can analyze a set of variables even if items have not been reverse coded because the factor analysis is attempting to determine whether people answer a set of scale items in a coherent fashion or whether participants respond one way to a certain type of question and another way to a different type of question. As a rule of thumb, items should load on a single factor at 0.50 or higher and not load another factor at 0.30 or higher. Often, when more than one factor is extracted from a factor analysis, an item will load equally high for two different factors. In essence, this means that the scale item is actually not measuring either factor but is instead an index of both factors, and since a scale can only measure one thing, that scale item must be removed if the scale is to be statistically valid.

Canonical Correlations

EXAMPLE

In the variable CA, there are four subscales that can be examined: group CA, meeting CA, interpersonal CA, and public CA. The variable WTC also has four subscales: group WTC, meeting WTC, interpersonal WTC, and public WTC. Suppose that you wanted to determine the nature of the relationship between the four subscales of CA with the four subscales of WTC. You could run a lot of correlations, or you could conduct one canonical correlation.

EXPLANATION

A **canonical correlation** is a statistical tool that allows a researcher to investigate the relationships among two or more variable sets. In our example, we have two different variable sets: CA (group, meeting, interpersonal, and public) and WTC (group, meeting, interpersonal, and public). Variables in a canonical correlation must be either interval or ratio level variables. All of the variables used in our example here are interval variables, so the canonical correlation is a good statistical tool to determine the interrelationships among the variables. Ultimately, a canonical analysis is the best test to use when examining statistical relationships between multiple interval/ratio independent variables and multiple interval/ratio dependent variables.

APA WRITE-UP

Let's see how the APA write-up of a canonical correlation would appear in a journal:

The goal of this research question was to examine the relationships among the four CA subscales (group, meeting, interpersonal, and public) and the four WTC subscales (group, meeting, interpersonal, and public). A canonical correlation was calculated using the four CA subscales as the predictors of the four WTC subscales. Using Wilks's Λ , the overall model was significant—Wilks's $\Lambda = 0.68$, $F(16, 889.66) = 7.54$, $p < 0.001$ —which indicates that the two variable sets are significantly associated by the canonical correlation. However, only the first two canonical correlations were found to be significant in this study: canonical correlation 1, Wilks's $\Lambda = 0.68$, $F(16, 889.66) = 7.54$, $p < 0.001$; and canonical correlation 2, Wilks's $\Lambda = 0.91$, $F(9, 710.80) = 3.10$, $p = 0.001$. Canonical correlations 3 and 4 were not significant: canonical correlation 3, Wilks's $\Lambda = 0.99$, $F(4, 586) = 0.94$, $p = 0.44$; and canonical correlation 4, Wilks's $\Lambda = 0.99$, $F(1, 294) = 0.07$, $p = 0.80$. The first variate accounted for approximately 26% of the variance in the dependent variable (canonical correlation [r_c] = 0.51), and the second variate accounted for approximately 7.84% of the variance in the dependent variable ($r_c = 0.28$). The exact canonical loadings for each variable can be seen in Figure E.9.

DISCUSSION

When examining a canonical analysis, always remember what is actually being correlated. In this example, we are examining the relationships between the four CA subscales (group, meeting, interpersonal, and public) and the four WTC subscales (group, meeting, interpersonal, and public).

The first statistic reported in the APA write-up is the significance test for the whole model. Because our independent variable (CA) had four variables (group, meeting,

Variable	Variate One	Variate Two
Group Communication Apprehension	-.81	.16
Meeting Communication Apprehension	-.94	-.13
Interpersonal Communication Apprehension	-.88	.41
Public Communication Apprehension	-.74	-.46
Group Willingness to Communicate	.81	-.41
Meeting Willingness to Communicate	.89	-.17
Interpersonal Willingness to Communicate	.61	-.67
Public Willingness to Communicate	.94	.08

FIGURE E.9
Canonical Variate
Loadings

interpersonal, and public), we ended up with four canonical correlations to be calculated. The results indicated that only the first two canonical correlations were significant. This does not mean that only the relationships between group CA and meeting CA are significant. Instead, in a canonical correlation, the model allows the possibility that each independent variable may function uniquely and thus need its own variate. To understand what the two significant canonical correlations mean, you must look at how all the variables (both independent and dependent) load on the significant canonical variates (see Figure E.9). We interpret each significant canonical variate separately in a fashion similar to interpreting a factor analysis. The stronger a variable loads on an individual variate (the closer that variable is to 1), then the more commonality that variable has with that specific variate. When looking at Figure E.9, you will see that the CA variables load negatively on the first variate, but that the WTC variables load positively on the first variate. This should be expected since CA and WTC are negatively related constructs. However, the story does not end there. When you look at the second variate, a new story is developing. The cutoff point for meaningfulness of a loaded variable on a variate is 0.30 (Tabachnick & Fidell, 2001), so in our example in Figure E.9, group CA, meeting CA, meeting WTC, and public WTC did not load on the second variate at all. Most of the variables loaded moderately on the variate (interpersonal CA, public CA, and group WTC). Interpersonal WTC is actually loaded negatively higher on the second variate than on the first variate. It should also be noted that interpersonal CA loads positively on the second variate.

So what does all of this mean about the relationship between the CA subscale variables and the WTC subscale variables? A lot. First, we learn that CA and WTC are clearly negatively related. However, the exact nature of that relationship depends on the subscales themselves because the subscales do not simply relate negatively. The second variate indicates that although interpersonal WTC, public CA, and group WTC may be low, it is still possible for someone to have a higher level of interpersonal CA. In other words, there are some people who do not exhibit high levels of public speaking CA, but who may still experience high levels of interpersonal CA.

CONCLUSION

In this appendix we have introduced you to a variety of advanced statistical techniques, shown you how their results may appear in a research article, and discussed seven research questions (that could be answered using the dataset on the textbook's website) and one hypothetical research question. We have looked at four difference tests (factorial ANOVA, ANCOVA, MANOVA, and repeated-measures ANOVA) and four relationship tests (path analysis, structural equation modeling, factor analysis, and canonical correlation). This appendix concludes our investigation of actual statistical techniques that communication researchers use to answer social scientific research questions.

GLOSSARY

Analysis of Covariance (ANCOVA): Statistical test that allows a researcher to determine whether a difference lies between groups on a dependent variable after the dependent variable has been mathematically adjusted for differences associated with one or more covariates.

Canonical Correlation: A statistical tool that allows a researcher to investigate the relationships among two or more variable sets.

Covariate: A possible predictive or explanatory variable that is observed rather than manipulated but that can have an effect on the dependent variable.

Endogenous Variables: Variables in a path analysis or structural equation model that are explained by one or more of the other variables in the model.

Exogenous Variables: Variables in a path analysis or structural equation model that are taken as given, measured variables.

Factor Analysis: A statistical device that enables researchers to determine whether the responses on a set of scale items actually measure a single construct or multiple constructs.

Factorial ANOVA: Statistical test in which a researcher has multiple nominal independent variables and one interval/ratio dependent variable.

Interaction Effect: Statistical difference test in a factorial ANOVA where differences on a dependent variable are examined by analyzing the combinations of the independent variables.

Law of Parsimony: Scientists should look for the simplest assumption in the formulation of a theory and the simplest test to interpret data.

Main Effects: Statistical difference test in a factorial ANOVA where differences on a dependent variable are examined for each independent variable separately.

Multivariate Analysis of Variance (MANOVA): Difference test that can utilize one or more nominal, independent variables and two or more related interval/ratio dependent variables.

Multivariate Test: Statistical test with two or more related dependent variables.

Path Analysis: Extension of the general linear model used to test the fit of two or more causal models.

Repeated-Measures ANOVA: Statistical test that allows a researcher to determine whether differences in the same interval/ratio variable occur over three or more measurements of the variable.

Structural Equation Modeling: Powerful multivariate analysis that enables a variety of specialized versions of other statistical tests, including regression models, causal modeling, confirmatory factor analysis, second-order factor analysis, covariance structure models, and correlation structure models.

Univariate Test: Statistical test with one dependent variable.

REFERENCES

- Bollen, K. A., & Long, S. J. (1993). *Testing structural equation models*. Newbury Park, CA: Sage.
- Bruning, J. L., & Kintz, B. L. (1997). *Computational handbook of statistics* (4th ed.). New York, NY: Longman.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications and programming* (2nd ed.). New York, NY: Routledge.
- Grimm, L. G., & Yarnold, P. R. (Eds.). (2000a). *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association.
- Grimm, L. G., & Yarnold, P. R. (Eds.). (2000b). *Reading and understanding more multivariate statistics*. Washington, DC: American Psychological Association.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 401–415.
- Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods* (3rd ed.). Glendale, CA: Pyrczak.
- Richmond, V. P., Wrench, J. S., Gorham, J. (2001). *Communication, affect, and learning in the classroom*. Acton, MA: Tapestry Press.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn & Bacon.

FURTHER READING

- Asher, H. B. (1983). *Causal modeling* (2nd ed.). Newbury Park, CA: Sage.
- Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Los Angeles, CA: Sage.