Mapping Your Decisions 13

You Can Get There From Here

So here you are near the end of the book and you might be thinking, "How do I remember all of this when I am working on my own research?" This chapter is designed to help you make the correct decisions about matching your research design and statistical analysis. We offer decision trees (flow charts) that will serve as reference tools you can use to ensure that you have a "good" marriage between your methods and statistics. Follow the flow of the branches by answering questions about your design and method. At the end, you will know which analysis to use!

Making Basic Decisions About Your Design

In this chapter, we include three decision trees for you to use that will allow you to make sure you made the correct connection between the design of your research and the type of statistic you use. You can find the three complete decision trees at the end of the chapter. Throughout this chapter, we include parts of the decision trees to illustrate the decision-making process that connects your design and statistic. The first two decision trees (i.e., Decision Trees 1 and 2) follow the decisions made when leading to statistics used with distinct groups when you want to compare the means for those groups or conditions. For example, you have distinct groups in your research when you include grouping variables such as sex or drug condition. Note that when we say *distinct groups*, we mean discrete levels of a variable; we do not mean you have to have different people in the groups. For example, you might have two levels (i.e., groups or conditions)

©SAGE Publications

of drug condition, such as a placebo and an antidepressant. You could either randomly assign people to the two conditions or test the same people under each level on different days.

Use Decision Tree 1 for research including just one variable with distinct groups (see Decision Tree 1 at the end of the chapter), and use Decision Tree 2 (also at the end of the chapter) when you decide to include two or more variables with distinct groups. In many cases, the distinct groups are levels of an independent variable (IV), which is the manipulated variable in an experiment. However, even a study with a non-manipulated variable such as participant's sex contains distinct groups; we sometimes refer to this type of variable as a pseudo-IV because it is not an IV that we can manipulate. Throughout this EasyGuide we referred to both manipulated IVs and pseudo-IVs as just variables. Regardless of whether you include a manipulated IV or a pseudo-IV, as you see in the decision trees, the statistic you use is the same. For these two decision trees, your variables can be nominal, ordinal, interval, or ratio (see Chapter 3 for descriptions of each), as long as a few distinct levels are used (e.g., drinking 1, 2, or 3 cans of soda). Importantly, the dependent variable (DV) must be interval or ratio data to allow for the powerful calculations reviewed in this book.

In the third and final decision tree, Decision Tree 3, we show you how to analyze a study with continuous variables that are not separated into distinct groups or categories but rather contain variables with many possible values, such as ounces of soda consumed, and you want to measure the relationship between those variables. The variables in such a study must be at the interval or ratio level to conduct these analyses. Use Decision Tree 3 when your research question relies on variables that do not have a few distinct categories. In such designs, you will analyze the data using either Pearson's *r* (correlation) or linear regression for prediction, depending on your goals.

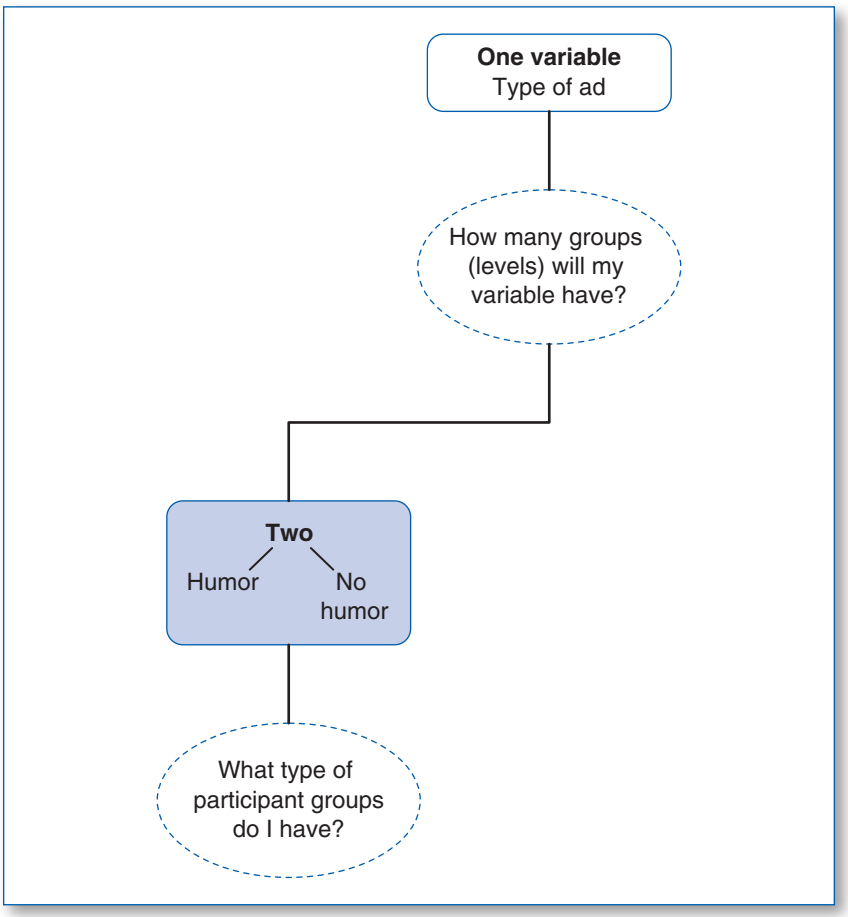
In all three of the decision trees, we map out questions you need to ask to identify the components of your design and in turn to identify the type of statistic that your data require. To help you refer back to the appropriate details, the decision trees also include the chapter numbers from this EasyGuide. That way, you can find details on a particular design, the relevant statistic, and how to run the analysis, interpret the output, and write results in APA style. Year after year, we suggest to our students that they laminate these decision trees and keep them handy for any class that requires a research proposal or project. After you have stated your hypotheses and identified the variables you want to investigate, the next steps are easy with the correct decision tree in hand. In this chapter, we also discuss in detail the different components of the decision trees. First, we go over the two decision trees for data with distinct groups (i.e., levels or conditions), and later we explain the components included in the decision tree for data with many values (such as ounces of cola people consume each month). Finally, we include examples to help you understand the differences among the decision trees.

©SAGE Publications

Data With Distinct Groups

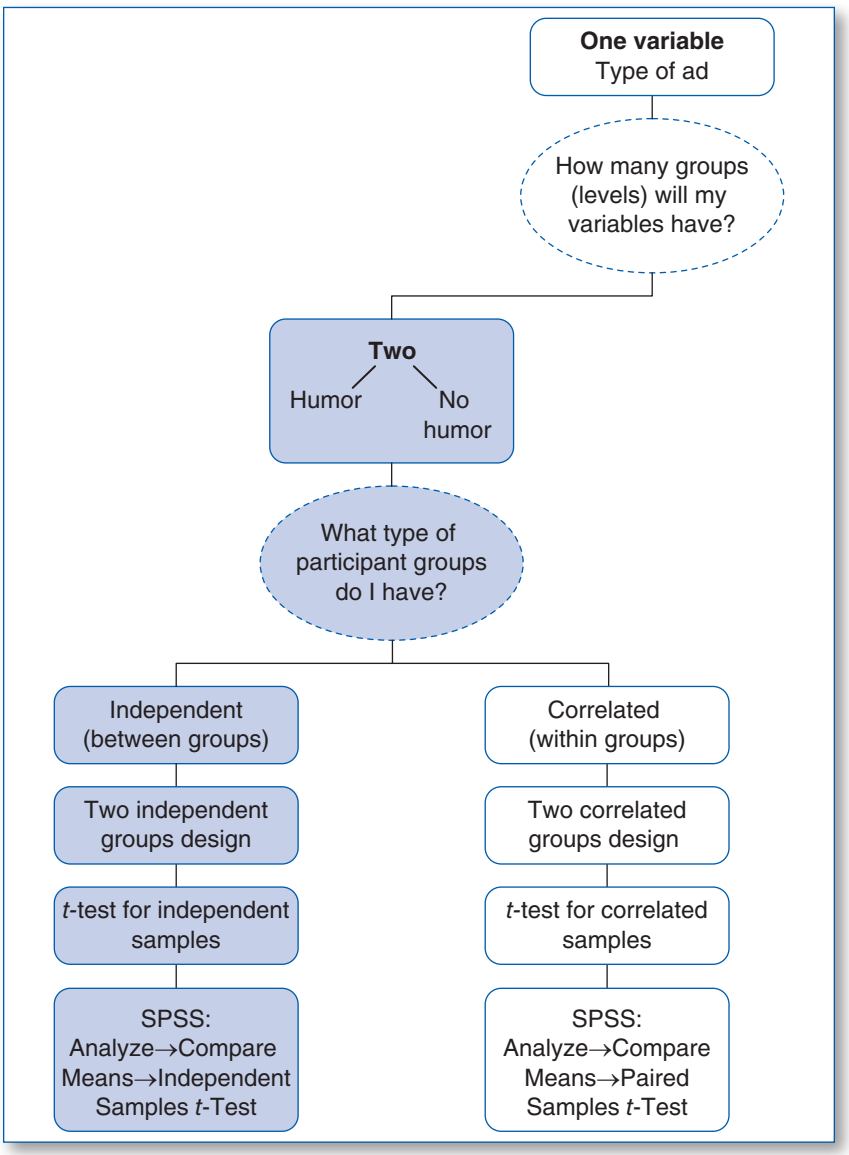
Keeping It Simple With One Variable

Let us start with Decision Tree 1 and a design with only one variable. For example, consider the ads you see either on TV or online. Often, advertising agencies use humor to help sell products. Does humor help us remember the product and in turn lead us to buy the product? Or does humor distract our attention away from the product itself, and in the end, we cannot remember what the ad was for because we focused on the humor? Based on the research literature (see Cline & Kellaris, 2007), we might predict that humor will lead to an increase in memory for the product. We could test this hypothesis by creating a design with one factor. In this example, type of ad (presence or absence of humor) would be our variable with two levels. One level would be ads with no humor, and the other level would be ads with humor. Notice on the decision tree below that you have made the choice of including two levels for one variable.



©SAGE Publications

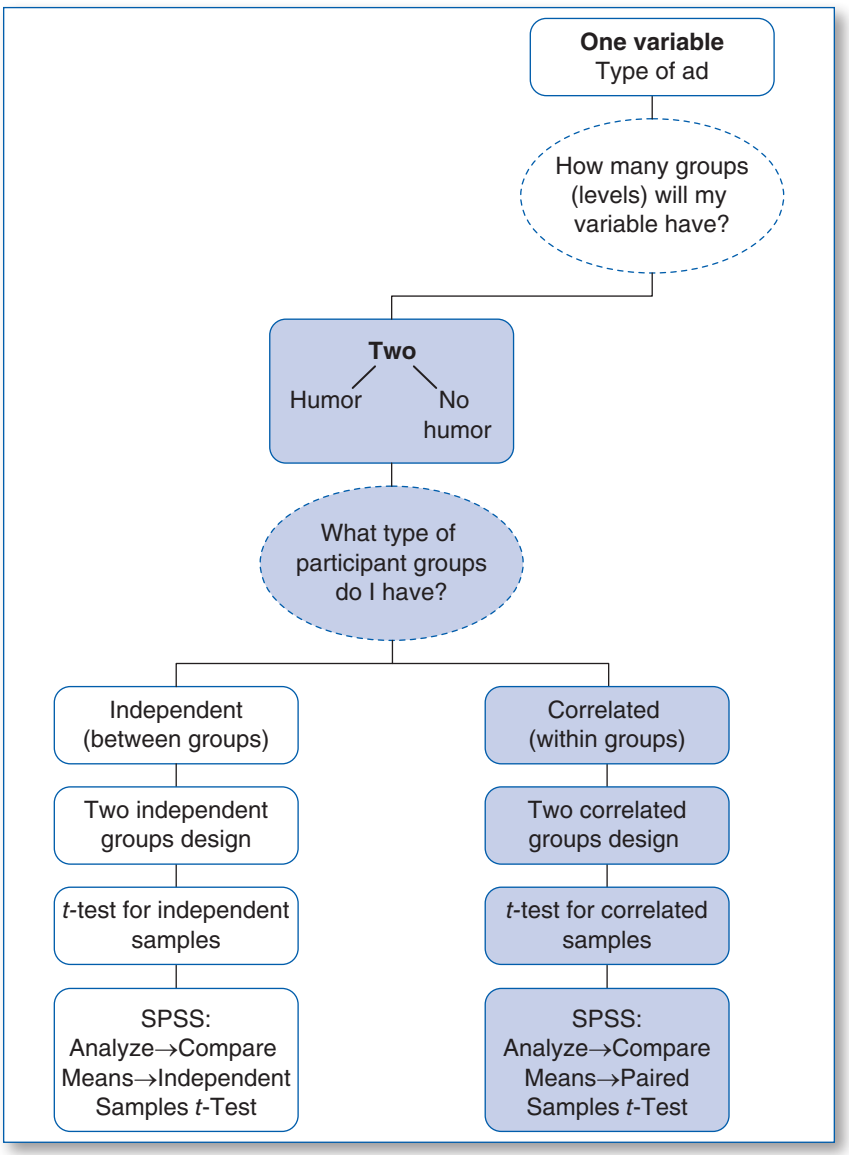
The next decision concerns *how* to present participants with these two levels. As one option, we could choose to have different people see the ad with humor versus the ad without humor. In this case, we would be using a between-groups design because separate groups of participants would experience only one of the two levels of our manipulated IV. As shown in the far left of the decision tree (following the blue highlighted path), this independent-samples design with two groups would require that we use a *t*-test for independent samples. You would use this same approach even with existing groups (a pseudo-IV). For instance, you might be interested in using only ads with humor and comparing memory differences between men and women. You would use all ads that contain humor



©SAGE Publications

and simply present those ads to both men and women. Then, compare memory for the products between men and women. Chapter 8 provides all the details needed for a between-groups design regardless of whether the variable was manipulated. In Chapter 8, you will also find guidance on how to set up a data file and run an analysis using SPSS as well as write your outcome in APA style.

Now, let us return to the two different types of ads. Imagine that you would like all of your participants to see *both* the humorous and the non-humorous ads. In this case (following the blue highlighted path below), you need to use a **within-groups design** in which the same participants see the two different types of ads. As you see, exposing all participants to both levels of your study changes the design and



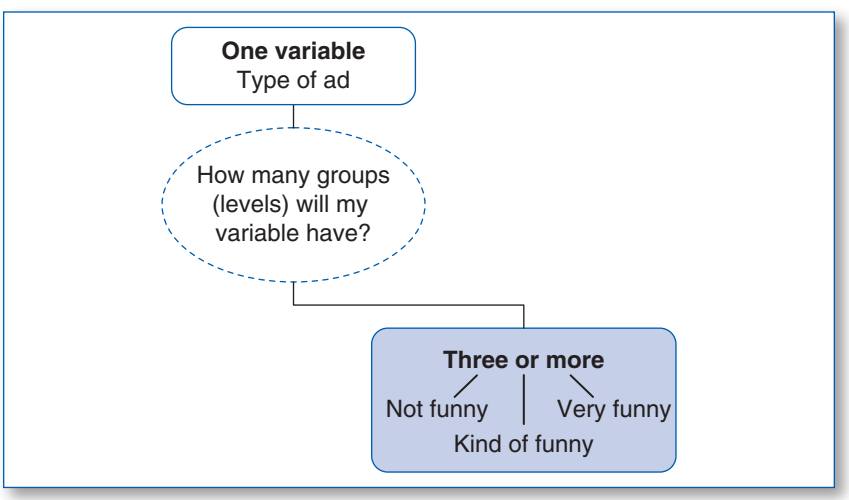
©SAGE Publications

statistic. Remember, if that were the case, you would be comparing the responses of the same participants who were asked to remember details about both types of ads.

There are advantages and disadvantages to choosing a between-groups design versus a within-groups design. When you choose a between-groups design, you do not need to be concerned about carry-over effects. In other words, participants who see the humorous ads are not influenced by previously seeing the non-humorous ads. Why? Because they are presented with only one or the other. That is the advantage of a between-groups design . . . no worries about carry-over effects. So why use a within-groups design when carry-over effects might cause a problem? The answer is all about statistical **power**. Power is the likelihood of accepting your research hypothesis when it is true. Statistically speaking, when you use the same participants in both conditions, variability produced by individual differences has less effect on the analysis because participants serve as their own comparison condition. In this case, if you have significant differences between your conditions, you will be more likely to detect those differences. When you can control for carry-over effects, this increase in power is well worth the effort.

Look at Chapter 9 to learn when to use a within-groups design with only two conditions. Some researchers refer to this design as a *two correlated-groups design*, a *related-samples design*, a *repeated-measures design*, or a *dependent-samples design*. Regardless of what you decide to call it, you analyze this design with a paired-samples *t*-test in SPSS. Chapter 9 will also take you step-by-step through how to set up the data file and what to click in SPSS for this type of design with two levels of the variable.

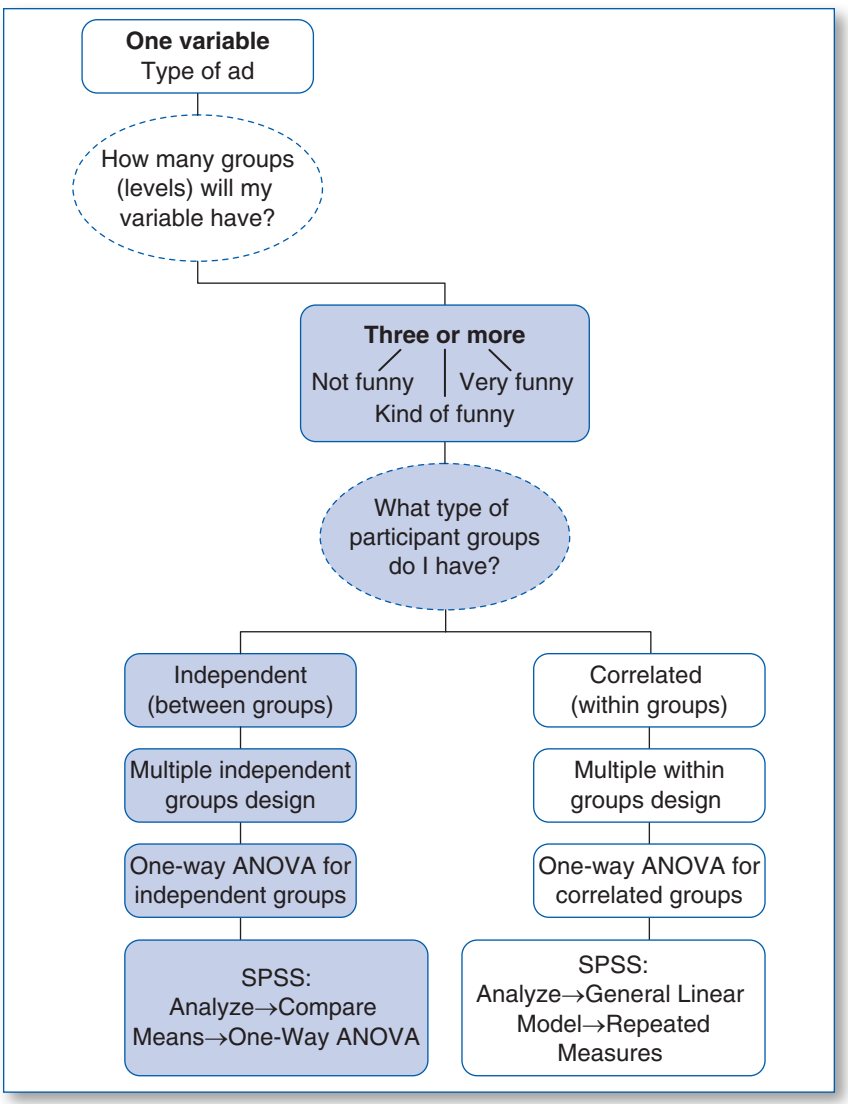
In the example above, we included only two levels of our variable. However, there are times when we want to compare more than two levels of a variable. For example, the literature might clearly indicate that humor improves attention to the product, but you realize that no one has investigated the impact of *how much* humor. Perhaps a small degree of humor causes a laugh but still allows a consumer to focus on the product, but a large amount of humor actually detracts from remembering the product. In other words, we could compare one ad without humor, one ad with slight humor, and a third ad with a lot of humor. Including



©SAGE Publications

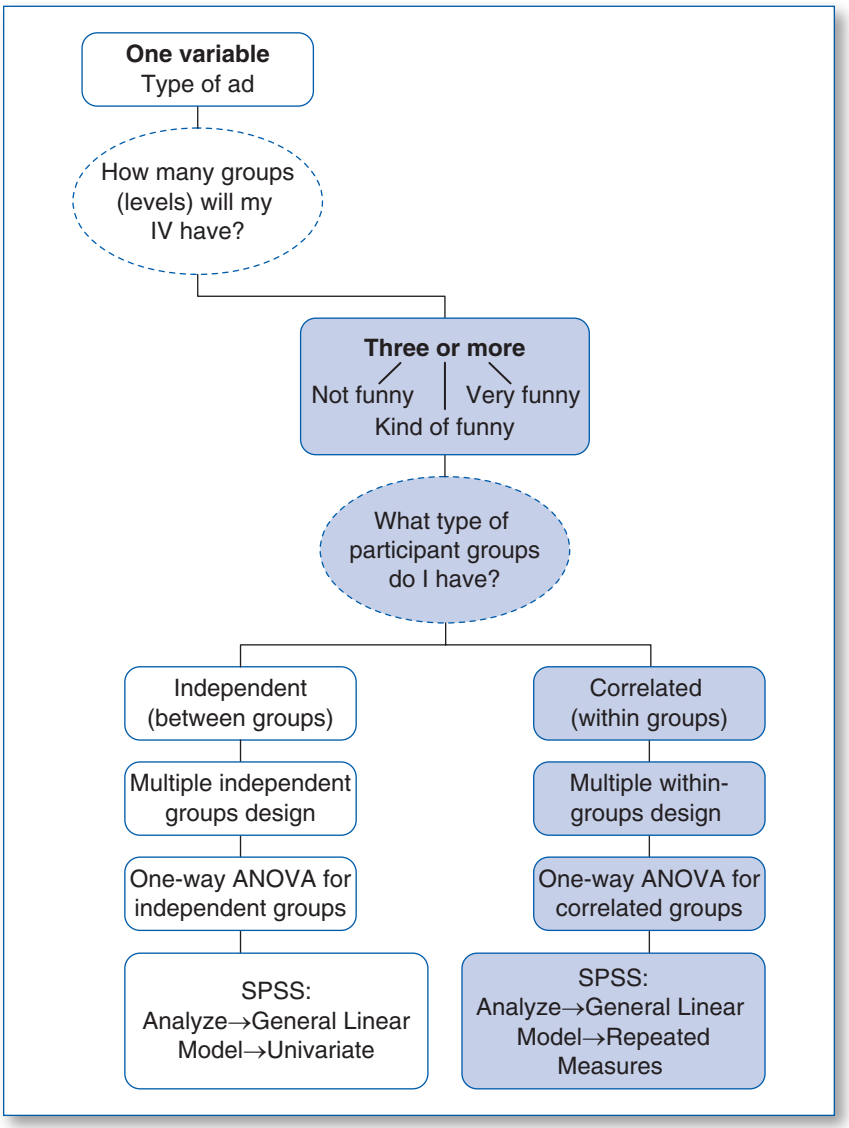
three levels of the humor variable would allow us to test our new hypothesis. When you look at decision tree below, you will see the two different ways we can include additional levels of our variable.

We are now on the right half of Decision Tree 1 (see the full tree at the end of this chapter) where we have three or more levels of a variable. We are still using one variable, degree of humor, but now our design employs three levels of that variable. Because we added another level of the variable, the design name changes slightly from a two-groups design to a multiple-groups design. Just as in the last example, you need to decide if you want to use a between-groups design or a repeated-measures approach to testing participants. When each participant experiences only one of the three levels, you have a between-groups design (Chapter 8). Follow the blue-highlighted section of the next decision tree.



©SAGE Publications

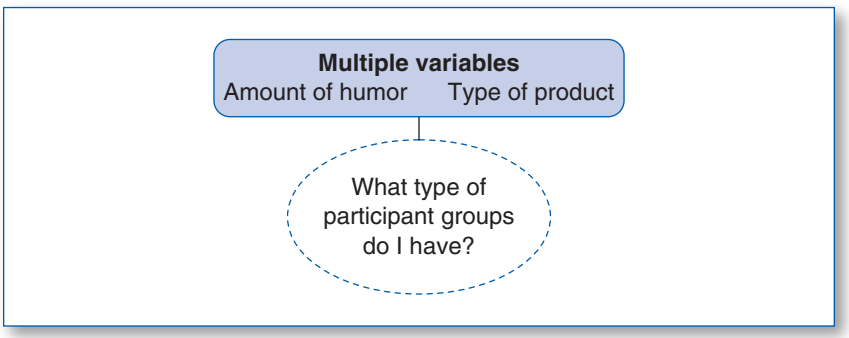
Or, the same participants can view all three ads, which would be a repeated-measures design, sometimes called a *within-groups design* (see Chapter 9). If that is your choice, your design and statistic are highlighted below.



©SAGE Publications

Designs With More Than One Variable

Published articles typically include more than one factor. After all, a researcher learns much more information when using multiple variables. In the humor examples used so far, we bet you could easily think of an additional variable that might increase our understanding of how humor affects memory for products. For example, you might wonder whether the type of product makes a difference when it comes to including humor. Perhaps an ad for a car would benefit more from humor than an ad for life insurance. You can answer this more complex question by including two variables in your research design: amount of humor and type of product. We now move to using Decision Tree 2.



Notice that the difference between the start of this decision tree and the one earlier in the chapter is that we added one more variable to the design. Once again, the decision must be made as to whether to use different participants in your variable levels or the same participants. However, this time the decision needs to be made separately for *each* of the variables included.

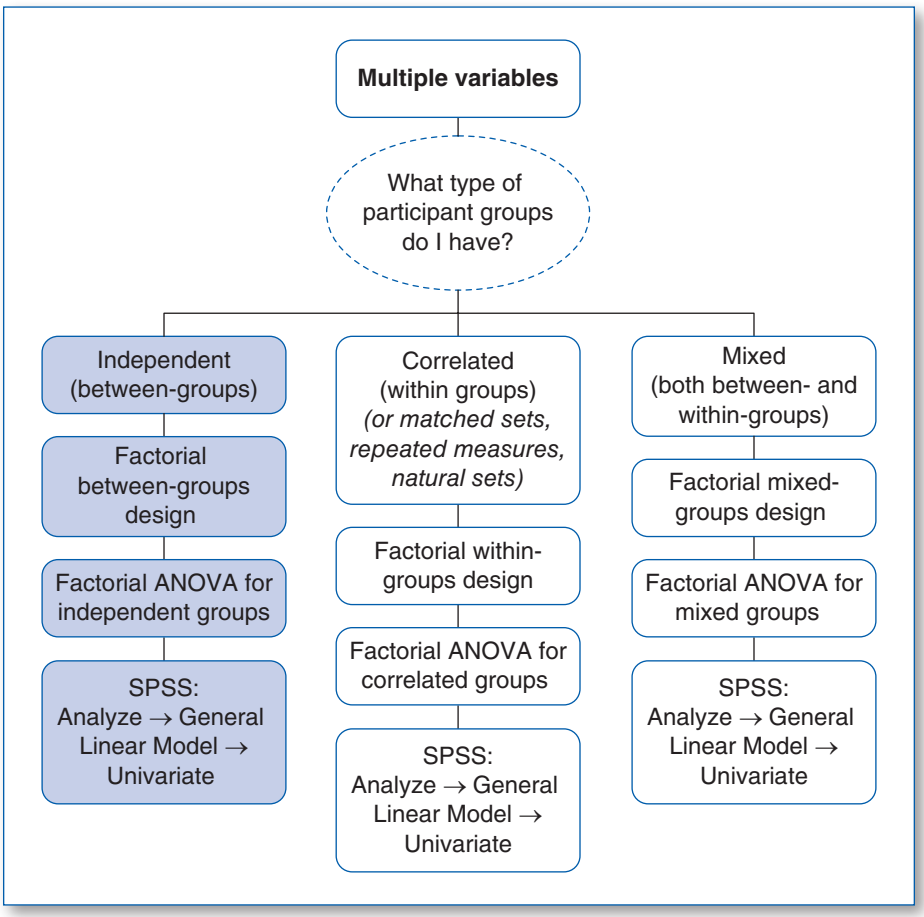
You could decide that both variables should be between-groups factors; you would be using a factorial design for between groups. Different participants would be assigned to all four of the different conditions in your research. Notice in the table below that combining the two levels of each of the two different variables leads to four separate conditions.

		Type of commercials	
		With humor	Without humor
Car ads		Car ad with humor	Car ad without humor
Insurance ads		Insurance ad with humor	Insurance ad without humor

In this factorial between-groups design, some participants would watch an ad for a car with humor, a different group of participants would watch a car commercial without humor, a third group of participants would watch a life insurance

©SAGE Publications

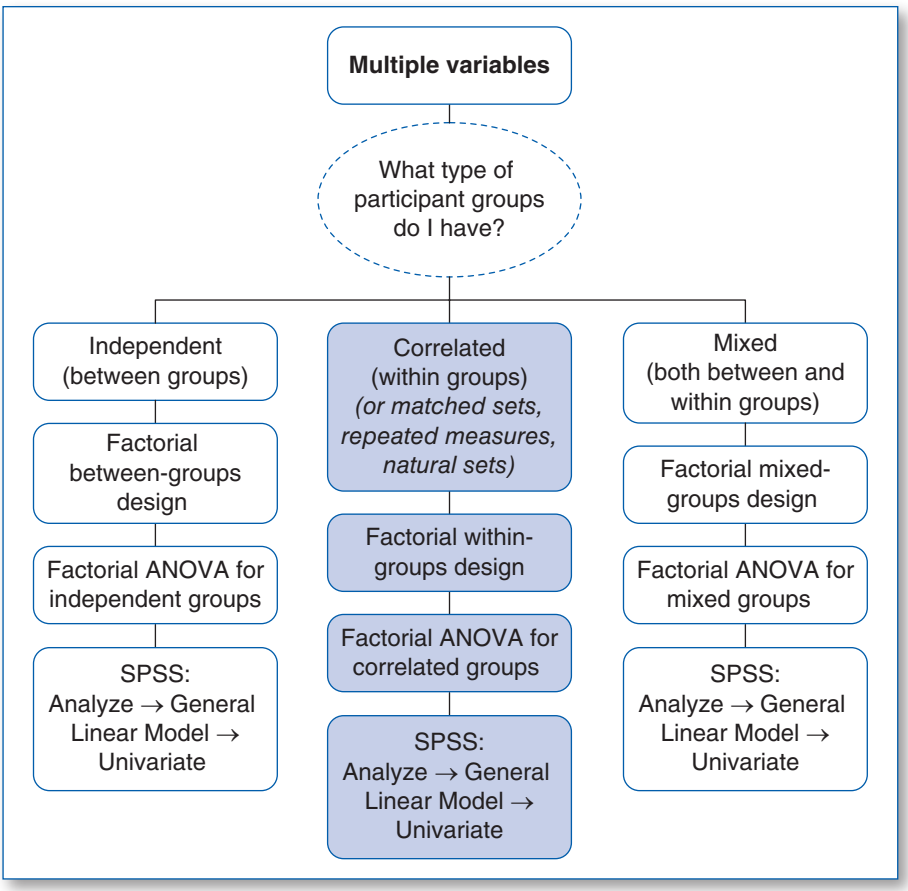
commercial with humor, and a fourth group of participants would watch a commercial for life insurance without humor. This follows the highlighted path in the decision tree on the far left as shown below.



You can find the reasons for making this choice, which statistic to choose, and how to set up the data file and run analyses in Chapter 8.

As an alternative, you could decide that you want all participants to experience the conditions created by the levels of both variables. In that case, the same participants would watch all four types of commercials in your research, and you have a factorial within-groups design. For details, follow the blue pathway in the center of the decision tree on the next page.

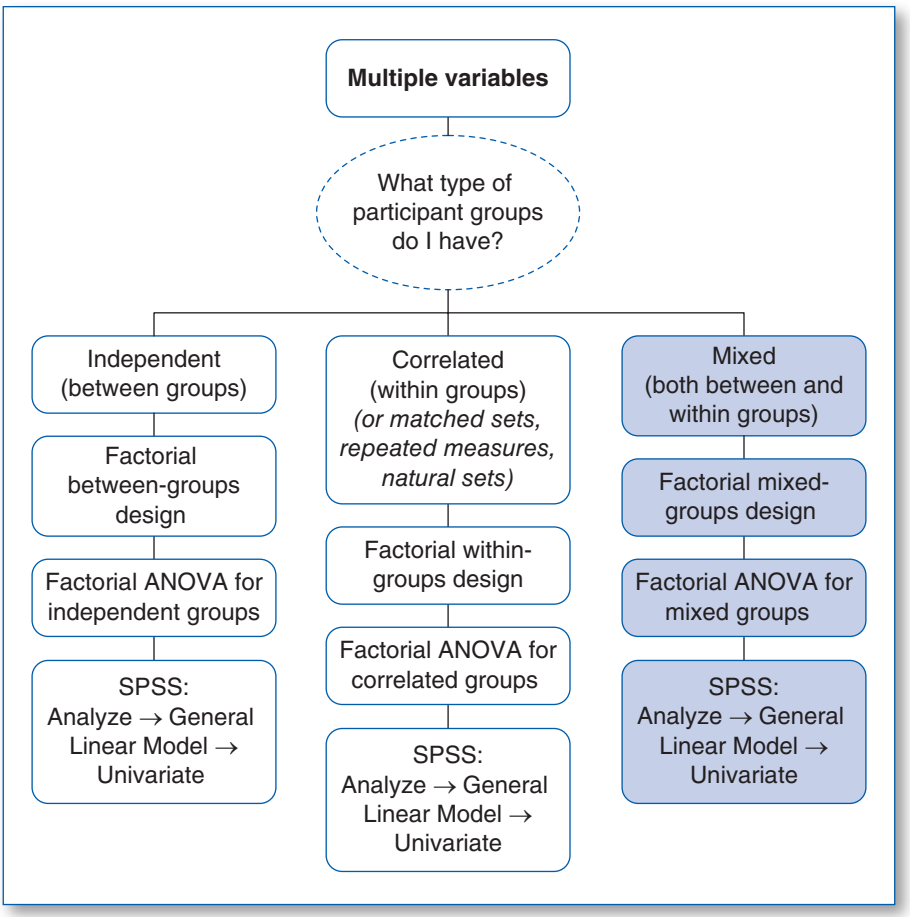
©SAGE Publications



Turn to Chapter 9 for a discussion of when you would want to use a within-groups approach and review how to run the appropriate analyses.

As a third alternative when you have more than one variable, you might decide to include one between-groups variable and one within-groups variable. In that case, you would be using a factorial mixed-groups design. We call it a *mixed* design because the study includes both types of variables: One is between groups, and one is within groups. Perhaps half of your participants watch the humorous commercial for both products, and the other half of your participants watch the non-humorous commercials for both products. In this case, can you figure out which variable is within and which is between? Remember, the within-groups variable is the one that allows all participants to see both levels of your variable (type of product). However, when it comes to the humor in the ad, participants will see either the ads with humor or the ads without humor (a between-groups variables). Your design and analysis choices are highlighted on the far right of the decision tree on the next page.

©SAGE Publications



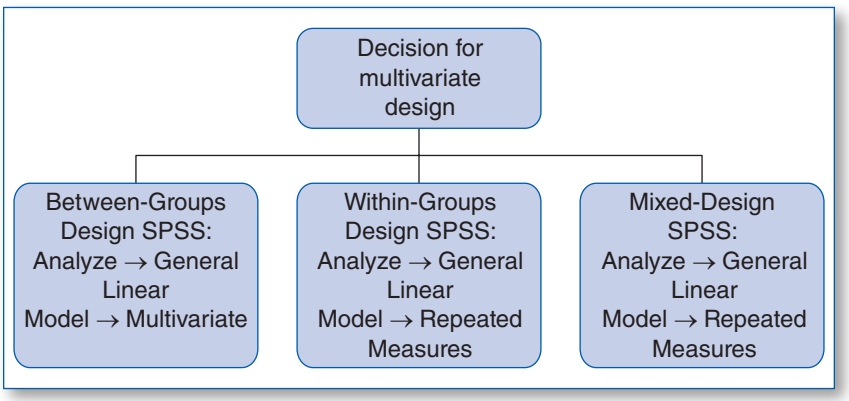
Chapter 10 provides the details you need when using a mixed-groups design and will provide information about what type of statistic to use.

As you can see from our first two decision trees, you can design studies with one or more variables, as many levels as you would like, and you can decide how to assign participants to levels. Follow these decision trees to locate the appropriate statistic. Keep in mind, as discussed earlier, that you can have either manipulated IVs (and show cause and effect) or pseudo-IVs (and know relationships between variables); both types of variables are analyzed the same way.

All of the designs you read about in the previous pages have a single DV. You might be thinking about a research project that will require several dependent measures. For example, you might be thinking about measuring driving ability and desire to measure acceleration from a stop, number of times the fog line is crossed, and stopping distance in both ordinary and emergency circumstances.

©SAGE Publications

Any design that requires more than one dependent measure is called a *multivariate design*. A multivariate design may use any of the experimental designs just described, so it could be a simple within-groups design or a factorial between-groups design. All of them rely on the same statistical procedure, the multivariate analysis of variance, or MANOVA.



In the next section, we will offer a decision tree to use when you do not have a few distinct levels for your variables.

Interval or Ratio Data With Many Levels

When variables are interval or ratio data with many levels, we often cannot rely on the analyses offered above. Instead, we turn to statistics that can handle a wide range of numbers. You will notice that the statistics include correlation (Pearson's *r*) and linear regression, but remember that correlational analysis and design mean two different things! Refer to Chapter 2 for a refresher, if needed. In the decision trees, we give you the correct path to a statistical analysis. However, keep in mind that your *design* will dictate whether you talk about cause-and-effect or relationships when you have results in your hand. If you manipulated a variable, you can discuss cause and effect; if you did not, you only can talk about relationships between variables. Either way, you can use Pearson's *r* to analyze interval or ratio data with numerous levels.

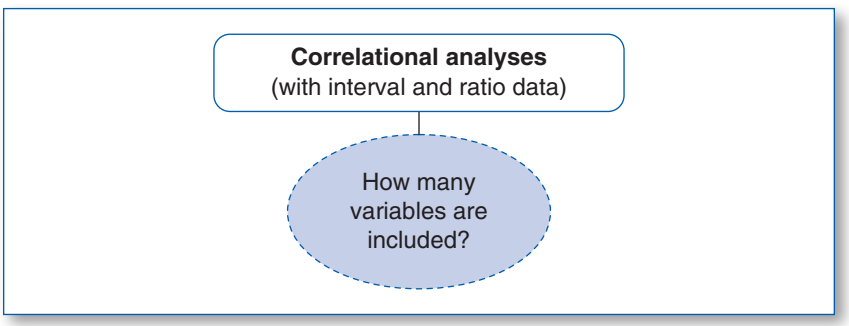
As an example, you might be interested in whether there is a relationship between the number of friends people have and self-esteem. You could ask a large group of people to fill out a survey and indicate how many friends they have (after you define what "friend" means, of course), and then you would measure their self-esteem using an established scale (let us say on a scale from 1 to 10, with higher numbers meaning higher self-esteem). Chapter 11 includes details on how to lay out your data in SPSS, run the appropriate analyses, interpret the results, and finally write up your results in APA style. Because number of friends

©SAGE Publications

and values on a self-esteem rating scale have many values in this design, you would analyze these data using Pearson's *r* (correlational statistic). In addition, if you also wanted to predict self-esteem from number of friends, you could go a step beyond Pearson's *r* to prediction with linear regression. In the following section, we offer paths on the decision tree for prediction, *but* if you only want to know whether two variables correlate, simply focus on the correlation branch of the decision tree (far left).

Keeping It Simple With Two Variables

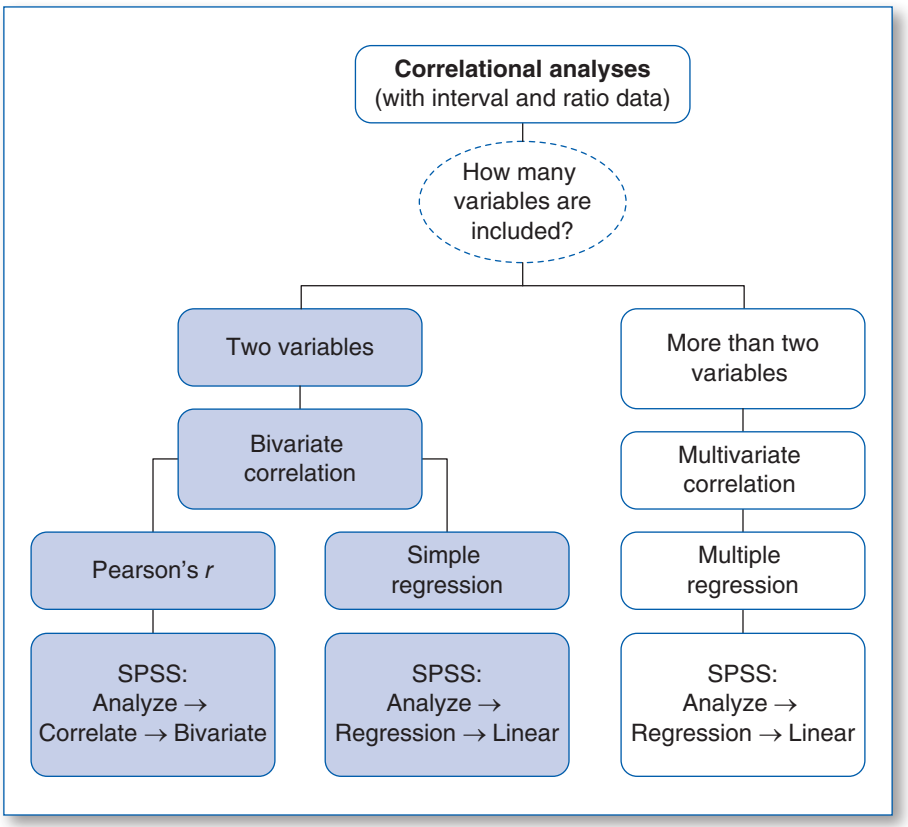
Let us start with a design having only two interval or ratio variables that were not manipulated and have many numbers possible on each variable. For correlational designs (not manipulated), instead of talking about independent variables (IVs), pseudo IVs, or dependent variables (DVs), we talk about plain old variables. If you want to know whether or not two variables are related, you can run Pearson's *r* for correlation as long as the two variables are interval or ratio data and have a wide range of numbers on each variable. Notice that at the top of this decision tree, your first question changes from "How many groups will my variable have?" or "What type of participant groups do I have?" to "How many variables are included?"



If you decide to include only two variables in your study, then you will follow the far-left path of the next decision tree; this path is for a **bivariate correlation**. A bivariate analysis means you are measuring the relationship between two variables. Sometimes, you merely want to know if two variables are related. If your research question goes a step further, and you want to predict one variable from another, simply look at the middle pathway in the decision tree on the next page, which allows you to test if one variable can predict the second variable. It should come as no surprise that the variable you choose as the predictor is called a *predictor variable*; the predicted variable is called a **criterion variable**. In the decision tree, look at the two blue

©SAGE Publications

pathways and notice that even with prediction, you first want to know if the two variables correlate. After all, if two variables are not related to each other, there is not much point in trying to predict one from the other.



Again, which blue path you choose at this point depends on the goal of your research project. Do you want to measure the potential relationship between two variables? Or are you interested in the ability to predict one variable based on knowing the value of the other variable?

As an example, you might wonder if the amount of chocolate consumption in a country is related to the number of Nobel Laureates produced in that country. This research question might seem odd, but when the relationship was actually measured, many were surprised to find out that the correlation is +.79 (and it was +.86 when Sweden was taken out ☺; Messerli, 2012).

©SAGE Publications

In this example, the findings indicate that the two variables are related to one another, and surprisingly, there is indeed a strong positive relationship between the two. To examine this relationship, you can run a bivariate correlation by using Pearson's *r*. Chapter 11 provides all the details on this type of correlational research, including how to create a scatterplot to visually present your results.

However, sometimes we are interested in more than just estimating the strength of the relationship between two variables. Researchers are also interested in predicting one variable when knowing the value of the other. Using SPSS, a simple regression will provide a regression equation and allow you to understand how well you can predict the outcome of one variable based on knowing the value of your predictor variable. In the chocolate–Nobel example, maybe you have a list of countries and the chocolate consumption of each and want to know if you can predict the number of Nobel Laureates based on chocolate consumption. Simple linear-regression analysis will provide the estimate. In Chapter 11, you can review details on how to set up the data file for both types of bivariate correlations, how to run the analyses using SPSS, how to interpret the SPSS output, and how to write the results up in APA style.

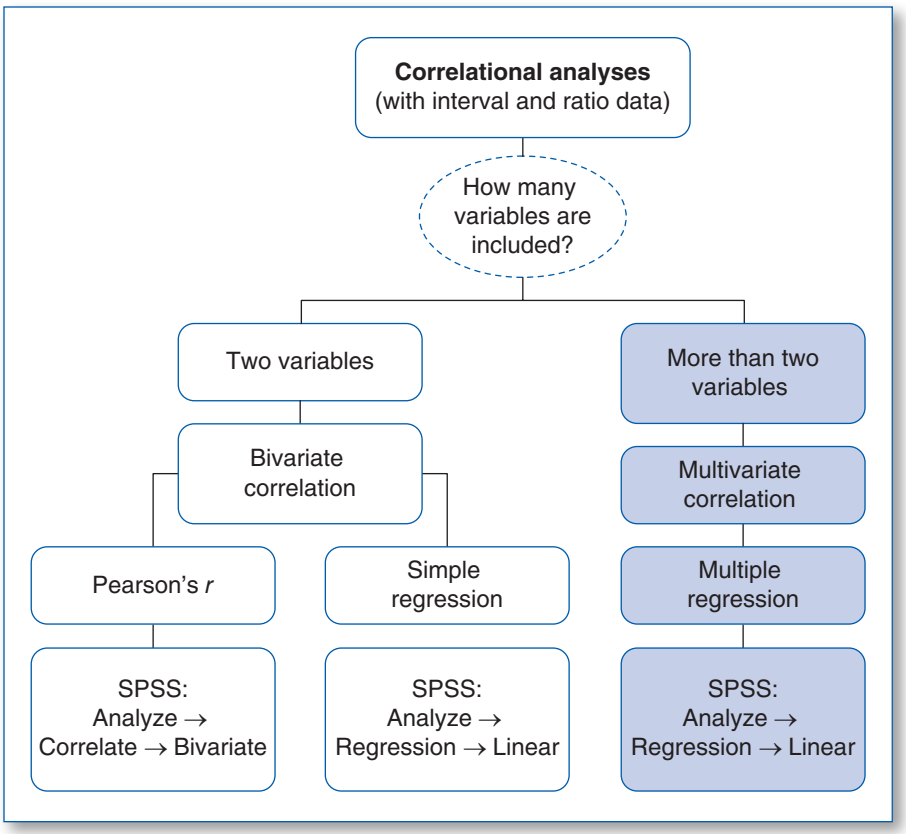
Including More Than Two Variables

After reading the example above and the surprisingly strong relationship between chocolate consumption and number of Nobel Laureates, we hope you realized that correlation does not mean causation. In other words, although chocolate consumption is related to the number of Nobel Laureates, there must be additional factors related to winning a peace prize that will help us better understand why this relationship exists. Posing that type of question brings up the final path on the decision tree. Perhaps you want to include more than one predictor. Maybe countries with more chocolate consumption also are more affluent, or have stronger education systems. Are those variables better predictors of Nobel Laureates than chocolate consumption? You now have multiple predictor variables to include in your design and analysis, so you would need to run a multiple linear regression and follow the path on the far right as highlighted on the next page.

The right side of the next decision tree illustrates the type of design you need to use and the type of statistics to run using SPSS. In the second highlighted box on the far right, *multivariate correlation* is a fancy way of saying you looked at bivariate correlations among all variables. That is, you can only look at two variables at a time when running correlations. Multivariate correlation means you ran Pearson's *r* for every possible pairing of variables in your design. When moving toward predicting, however, multiple predictors are used to predict one criterion (outcome) variable. And all predictors can be analyzed in the same equation, usually allowing better prediction of the outcome variable than using only one predictor. As you can see on the next page, this prediction analysis is called *multiple regression*.

©SAGE Publications

Take a look at the end of Chapter 11 to read about how to run this type of analysis, see what the output will look like, understand how to interpret the output, and learn how to write the results in APA style.

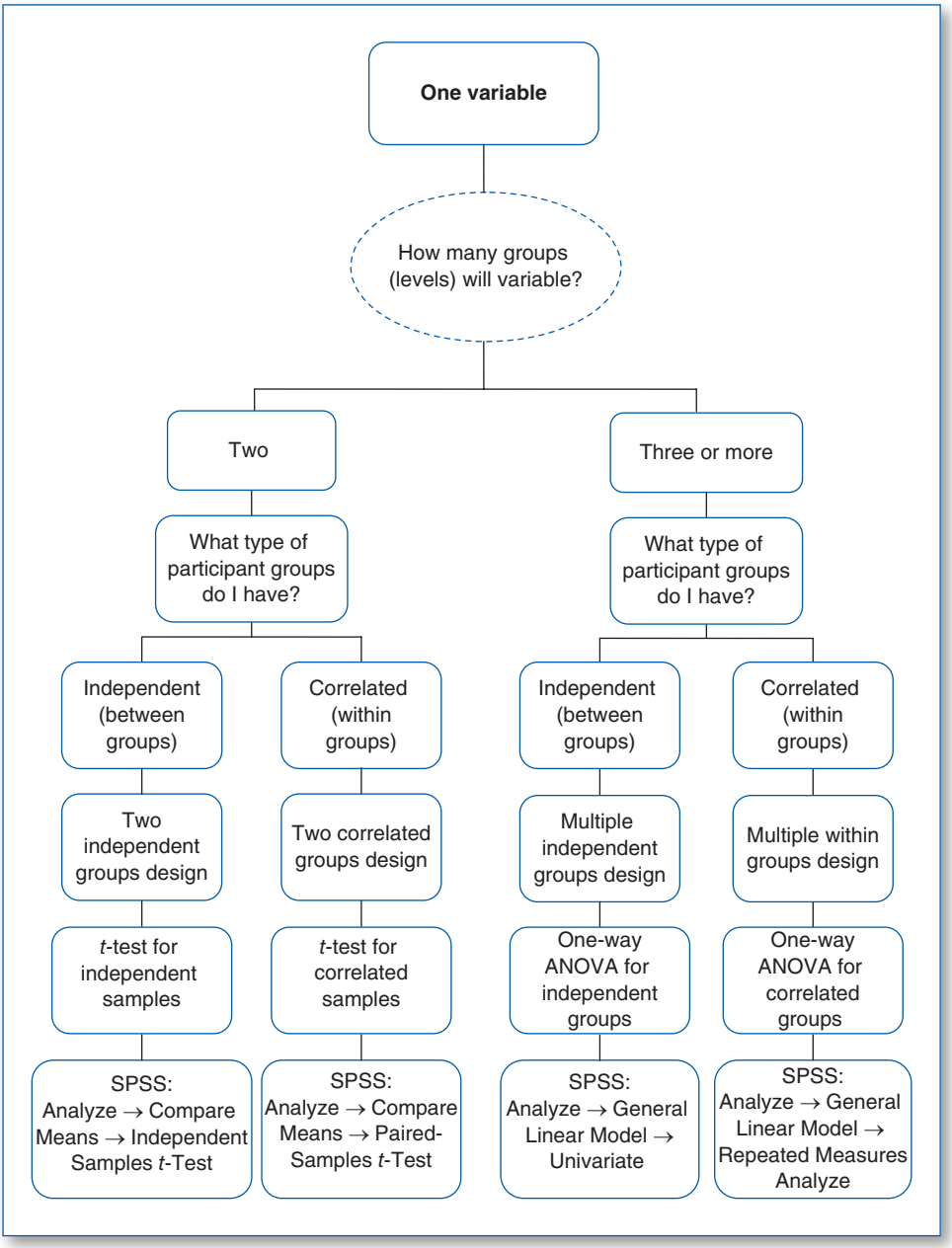


Summary

Throughout this chapter, we have offered examples to demonstrate the different paths to take on the three decision trees. We urge you to copy the complete decision trees included at the end of this chapter and put them up on your wall as a poster. Or simply keep this book handy! Either way, the following pages offer the decision trees in full for your review. As you develop your research ideas, a glance at the decision trees will remind you how each decision you make concerning the design of your research leads to a specific type of design and, in turn, the use of a specific type of analysis and name of statistic to use in SPSS. Remember, the chapters earlier in this EasyGuide cover the design choices and appropriate statistics for each branch of the decision trees. In those chapters, we promise you can find help for each step of the way!

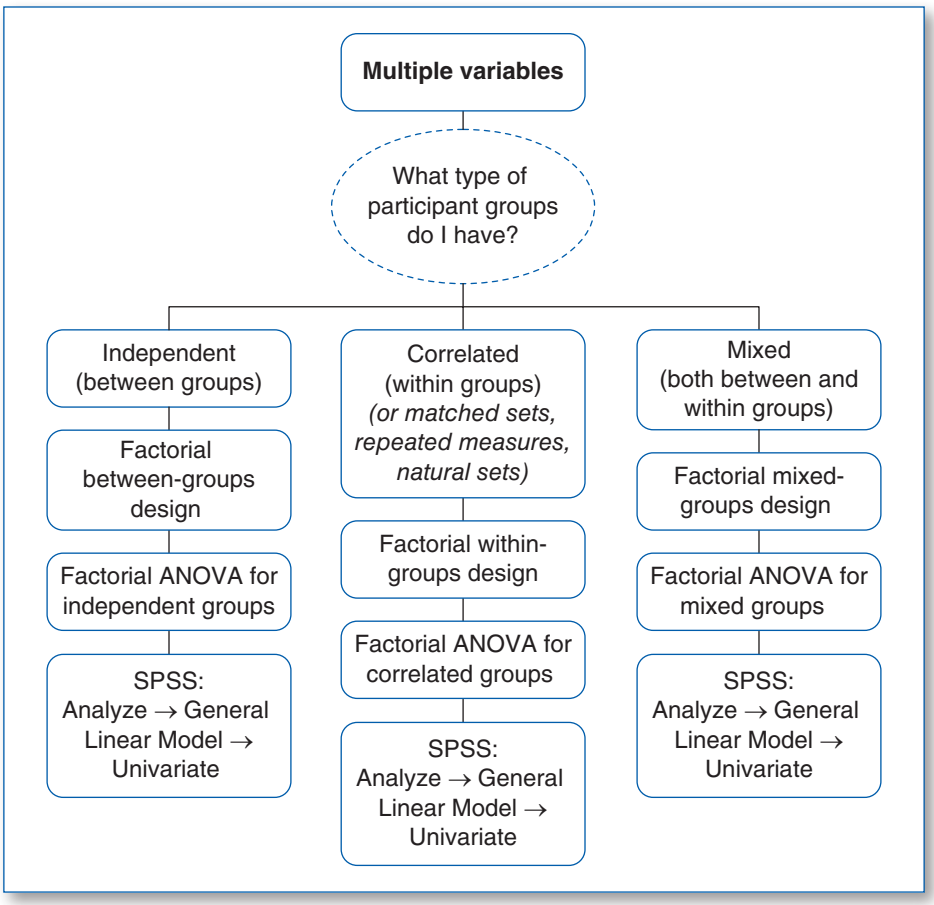
©SAGE Publications

DECISION TREE 1



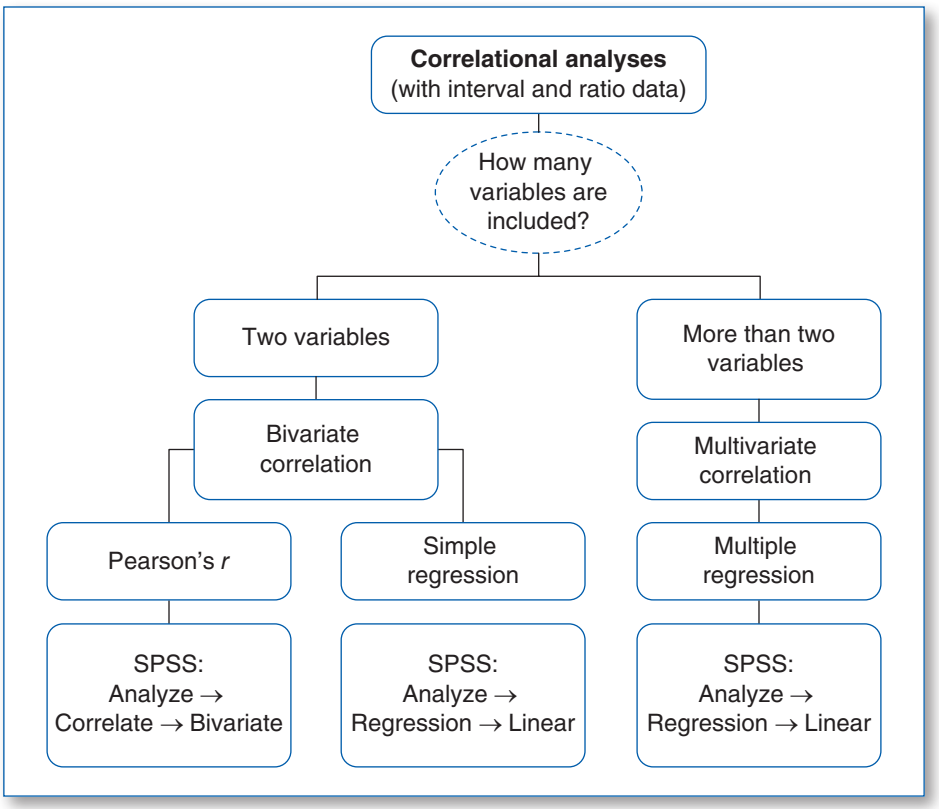
©SAGE Publications

DECISION TREE 2



©SAGE Publications

DECISION TREE 3



©SAGE Publications