CHAPTER

7

Examining Relationships

In the previous chapter, we discussed the two-way chi square (χ^2) as a way to find if two variables are related. Chi-square analysis requires simple frequencies in separate groups. The research question asks how many people are in each level across two variables, and data analysis with the χ^2 reveals a relationship between the two variables, if a relationship exists. For example, your research question might ask, *Is money in a person's wallet related to happiness?* The two variables could be whether or not people in a sample have money in their wallets and whether or not they are happy. Data from such a design might consist of the observed simple frequencies in the four cells below.

	Not Happy	Happy
No money in wallet	11	15
Some money in wallet	9	20

Notice that 11 people in the sample were unhappy and reported having no money, 15 people had no money and were happy, and so on. The outcome data are simple frequency counts. The research design called for how many people fell in each of the four categories. Analysis of this design using a two-way χ^2 will tell you if money in a person's wallet is related to happiness. The data above are simple, yielding only frequency counts. Such simple data cannot be analyzed with a powerful statistic based on the definition of power in Chapter 4. We must rely on a nonparametric statistic. Not to worry, a different research design can be analyzed using a more powerful option.

PEARSON'S *r*: SEEKING A RELATIONSHIP

Recall from Chapter 3 the four levels of data: nominal, ordinal, interval, and ratio. Interval and ratio data are defined by math properties and allow advanced data analysis. Your knowledge of measurement levels will serve you well as you design a study based on a more advanced research question: *Is the amount of money people have in their wallets related to happiness such that more money is related to greater happiness?* We still want to know about money, but now we can ask for an exact amount rather than whether or not a person has any money at all. Amount of money in wallets is a ratio variable. We also still want to know about happiness, but this variable can be measured on a rating scale from 1 (*very unhappy*) to 4 (*very happy*). A rating scale represents interval data, a big improvement over simply asking people if they are happy or not (nominal data).

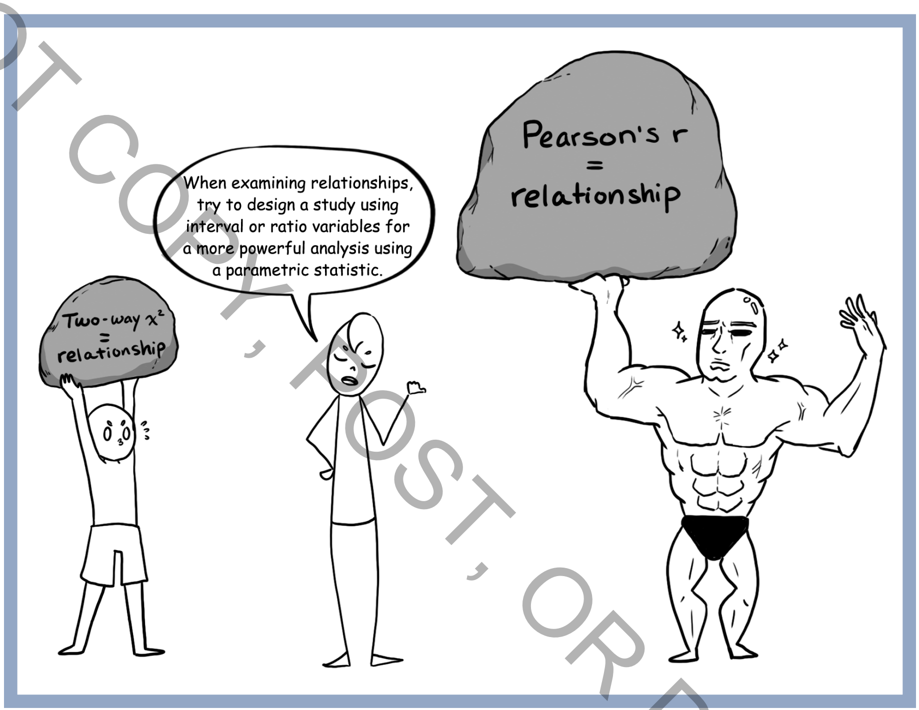
In our example, each participant provides information on both variables. A fictional data set is below.

Dollars in Wallet	Happiness
2	3
10	3
20	2
5	4
12	3
15	4
26	3
5	4
21	2
7	1

Just as with the two-way χ^2, we can find out if the two variables are related. As an added bonus, locating a relationship when one exists is more likely with interval or ratio data as opposed to simple frequency counts because quantitative variables allow a more powerful, parametric statistic. Of course, data analysis must be different to accommodate for interval or ratio data rather than simple frequency counts. To

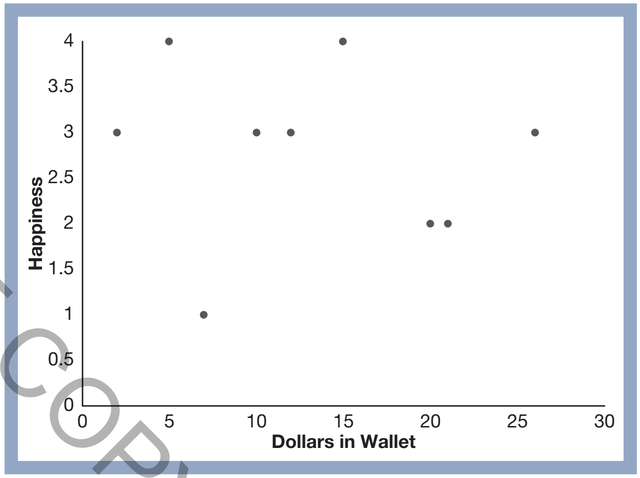
Pearson's *r* (Correlation Coefficient). Pearson's *r* is the statistic used to analyze a potential relationship between two interval or ratio variables.

analyze a potential relationship between two interval or ratio variables, we turn to a statistic called **Pearson's *r***. The statistic is equally often called the **correlation coefficient** because the analysis examines the potential relationship between two variables, a *co*-relation.



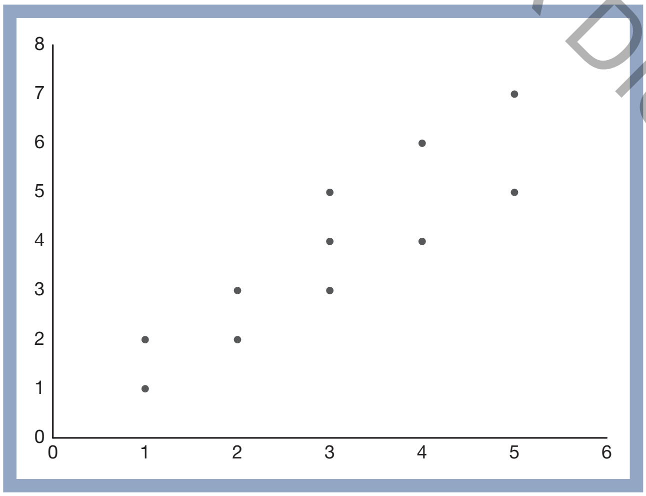
Scatterplot. A scatterplot is a graph that shows all data points in a data set, with one point for each participant.

The Pearson's *r* statistic indicates the relationship between two interval or ratio variables. Specifically, it measures how related they are in a linear (straight line) way. If the two variables are linearly related, a graph should depict the relationship. A **scatterplot** shows all data points, one for each participant. Usually researchers label the X-axis (horizontal line on the bottom of the graph) with the variable in the first column. The Y-axis (vertical line on the left side of the graph) is labeled with the variable in the second column. Choose values for each axis based on the values in each column. In our example, numbers in the first column range from 2 to 26, and numbers in the second column range from 1 to 4. In the following figure, we graphed data from the 10 participants reporting the money in their wallet and happiness. You will see only nine dots because two participants had identical data.



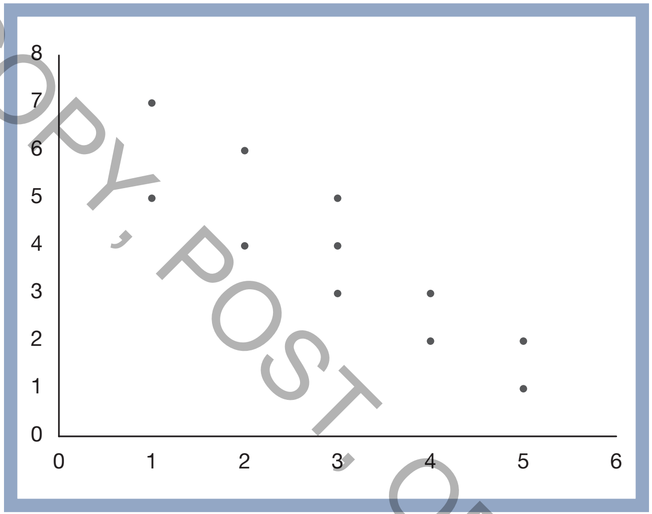
Do the points on the graph appear to fall approximately in a line? If you do not see a linear pattern, you are not alone. We do not see one either. Lack of a clear linear pattern suggests that these two variables are not related. But just to be sure, we should calculate the correlation coefficient. The Pearson's r value can range from –1.00 to +1.00, and both the absolute value and sign are important.

The correlation coefficient provides two pieces of information: *direction* of the relationship and *strength* of the relationship. The sign of the r-value indicates the direction of the relationship. If the scatterplot shows points in a line pattern from the lower left of the graph to the upper right of the graph, the two variables have a positive relationship.

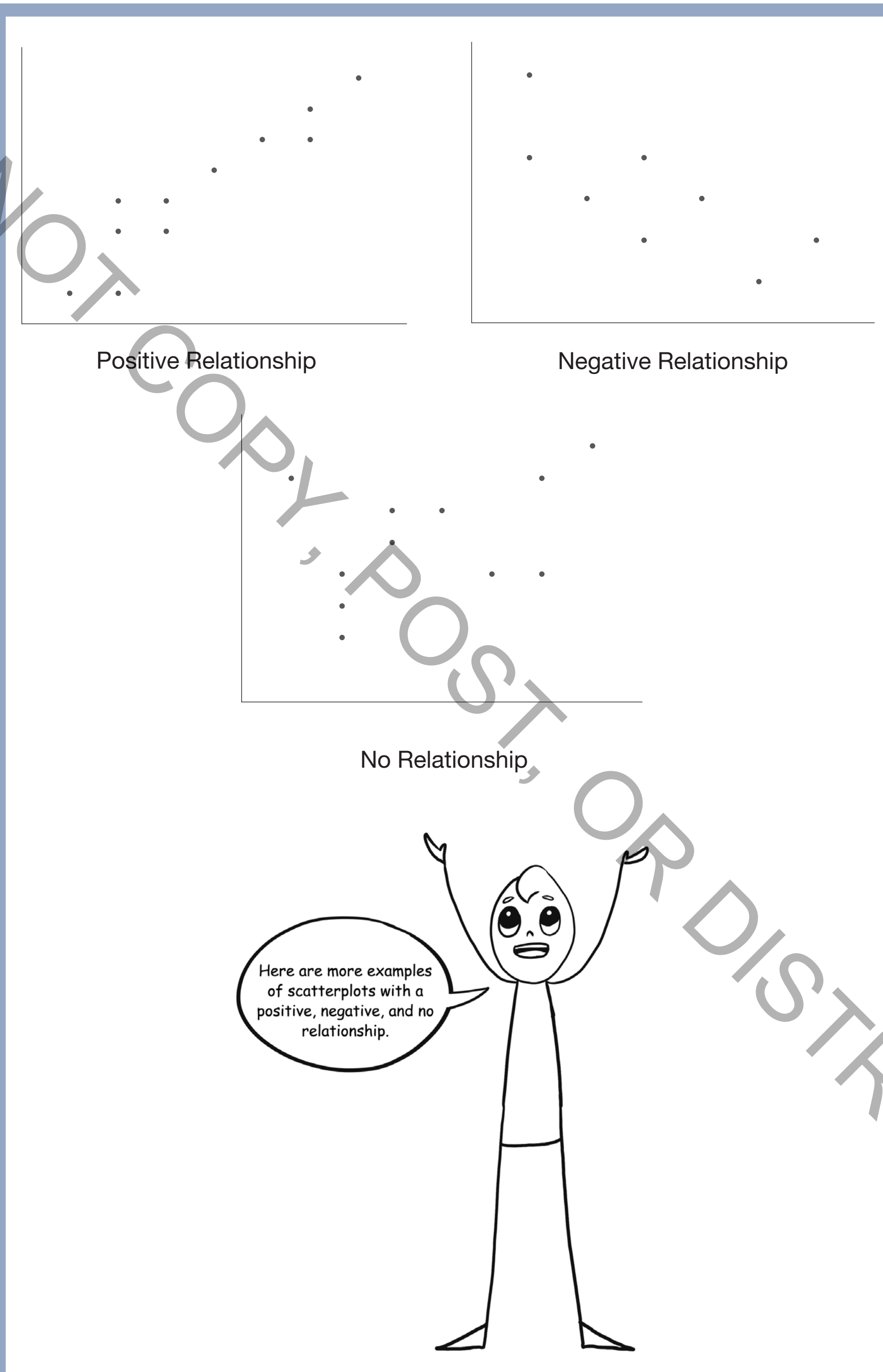


The *r*-value will be positive. As values on one variable increase, so do values on the other variable. The scatterplot below illustrates a positive relationship.

However, if the scatterplot shows an approximate line pattern from the upper left of the graph to the lower right, the two variables share a negative relationship. The *r*-value will be negative. As values on one variable increase, values on the second variable decrease. The scatterplot below represents a negative relationship.



Pearson's *r* also tells the strength of the relationship. Values of either +1.00 or −1.00 for *r* means the two variables are perfectly related. As an example of a perfect positive relationship, numbers on the first variable might increase by 1, while numbers on the second variable increase by 5. For a perfect negative relationship, numbers on the first variable might increase by 3, while numbers on the second variable decrease by 10. In real life, perfect relationships—whether positive or negative—do not exist. Likewise, a Pearson's *r* of 0.00, indicating absolutely no relationship at all between two variables, rarely exists. Life is messy, and two variables will randomly overlap a bit, yielding an *r*-value *close* to .00 but almost never exactly .00.


Positive Relationship
Negative Relationship
No Relationship
Here are more examples of scatterplots with a positive, negative, and no relationship.

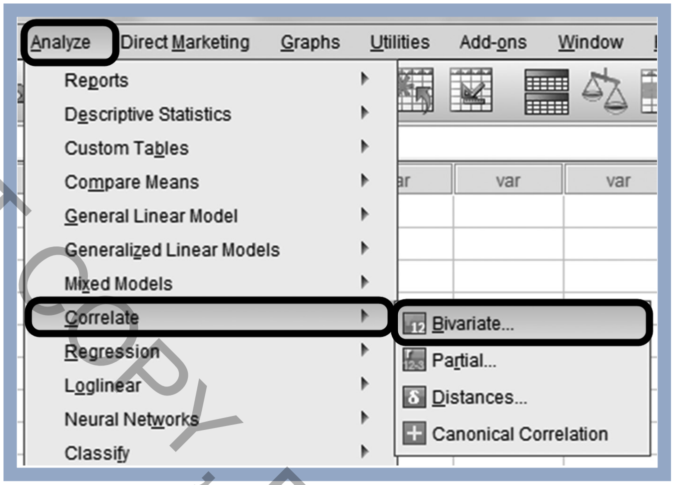
As researchers, we hope for a value beyond .00, but we do not expect to find a perfect relationship between our variables. In fact, a Pearson's *r* of +/–.10 is considered a weak relationship, +/– .30 depicts a moderate relationship, and +/–.50 or beyond shows a strong relationship (Cohen, 1992). Although Cohen has given us these guidelines for the strength of a relationship, we know the ultimate test of significance lies in analyzing an *r*-value to see how unusual it is. If the value is unlikely to occur normally (meaning it is very different from zero), we can take credit for finding an interesting relationship. To test for significance, we turn to SPSS.

SPSS: Pearson's *r* (Seeking a Relationship)

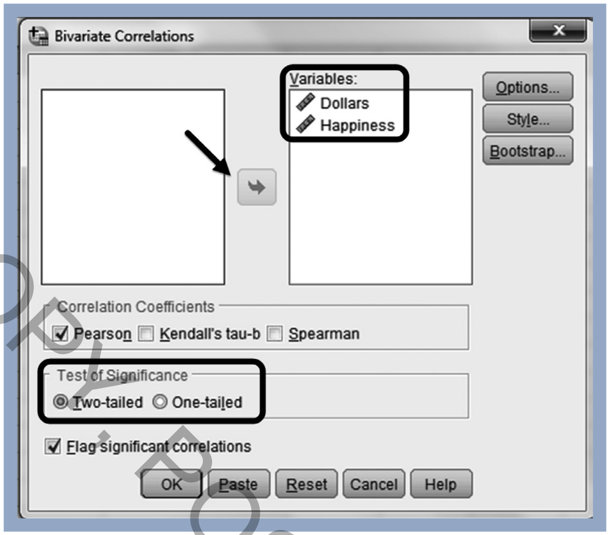
In Variable View, enter the variable names as you have in prior chapters. We do not typically assign labels to numbers under Values when analyzing interval or ratio data that are not grouped into distinct levels. Enter the two variable names and click Data View at the bottom left of the screen. In Data View, enter values for dollars and happiness, and make sure the data from each participant are on one row together. For example, the first participant reported having $2.00 and a happiness rating of 3.

Dollars	Happiness
2.00	3.00
10.00	3.00
20.00	2.00
5.00	4.00
12.00	3.00
15.00	4.00
26.00	3.00
5.00	4.00
21.00	2.00
7.00	1.00

Click Analyze, Correlate, and Bivariate to request Pearson's *r*.



In the box that opens, move both variables over to the right by highlighting each and clicking the center arrow.

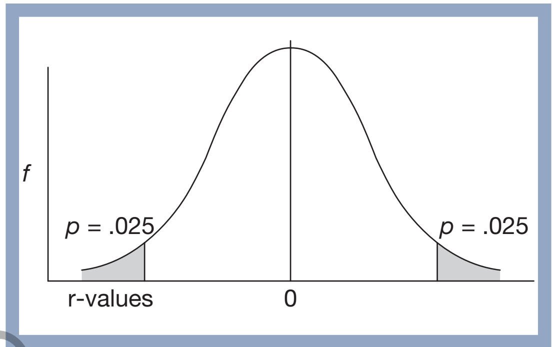


Under Test of Significance, notice your option of a Two-tailed or One-tailed test.

Two-Tailed Test

Recall that for significance, we need a *p*-value of .05 or less. In other words, we need to find an unusual, interesting relationship between money and happiness. Of course, the null hypothesis illustrates no unusual relationship: *Amount of money in wallets and happiness are not related*. Because no relationship would be expected in the normal population, a value of $p \leq .05$ means we found an unusual relationship very different from 0. If we design a study and have no idea whether the *r*-value will be positive or negative, the .05 value must be split across a positive *r*-value (.025) and a negative *r*-value (.025). We split the .05 we have to use. Unfortunately, splitting the .05 sacrifices power because each possible *r*-value direction gets only a .025 (2.5%) chance to occur. This is called a **nondirectional test**. A nondirectional test is also called a two-tailed test. If you have no educated guess about the direction of the relationship, conduct a two-tailed test. Notice that both tails of the distribution below are shaded, and each shaded region is smaller than a one-tailed test (below) because .05 has been divided between two areas. Likewise, if your instructor prefers a conservative approach, rely on a two-tailed test.

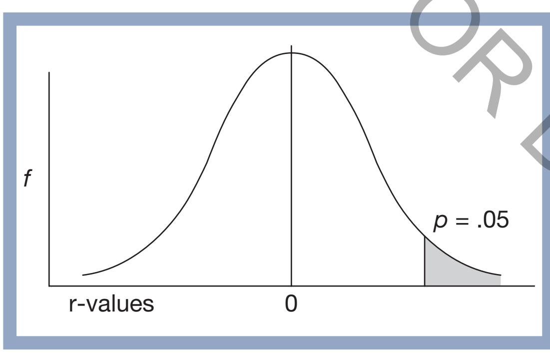
Nondirectional Test (Two-Tailed Test). If we have no educated guess about the direction of the relationship, we conduct a nondirectional or two-tailed test.



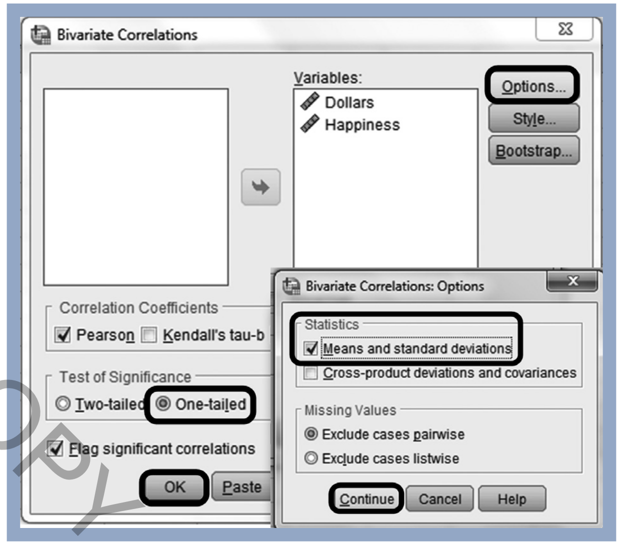
One-Tailed Test

Alternatively, if we have a specific idea (hypothesis) of whether the relationship will be positive or negative, we can gamble all of our .05 allowance on the single direction we expect. We call this approach a **directional test**. In our example, recall the research question: *Is the amount of money people have in their wallets related to happiness such that more money is related to greater happiness?* The research hypothesis specifically states what we hope to find: *Amount of money in wallets is related to happiness such that more money is related to greater happiness.* That is, we expect dollars and happiness to yield a *positive* *r*-value. Bet all .05 on a positive relationship. A directional test is also called a **one-tailed test**. As indicated in the following figure, this specific one-tailed test is based on expecting the *r*-value to fall in the upper tail of the normal distribution.

Directional Test (One-Tailed Test). If we have a specific idea (hypothesis) of whether the relationship will be positive or negative, we have a directional or one-tailed test.



In the money study, we expect money and happiness to show a positive relationship, so we conduct a one-tailed test in SPSS. Click the circle next to One-tailed under Test of Significance. Next, click Options. Options will open a second box that allows you to ask for descriptive statistics. In addition to testing for a significant relationship, you will need to report the mean and standard deviation for variables. After checking the box for these options under Statistics, click Continue.



When the box closes, click OK on the original box (Bivariate Correlations) to execute commands. See the output below.

Correlations

Descriptive Statistics

	Mean	Std. Deviation	N
Dollars	12.3000	8.00069	10
Happiness	2.9000	.99443	10

Correlations

		Dollars	Happiness
Dollars	Pearson Correlation	1	-.247
	Sig. (1-tailed)		.246
	N	10	10
Happiness	Pearson Correlation	-.247	1
	Sig. (1-tailed)	.246	
	N	10	10

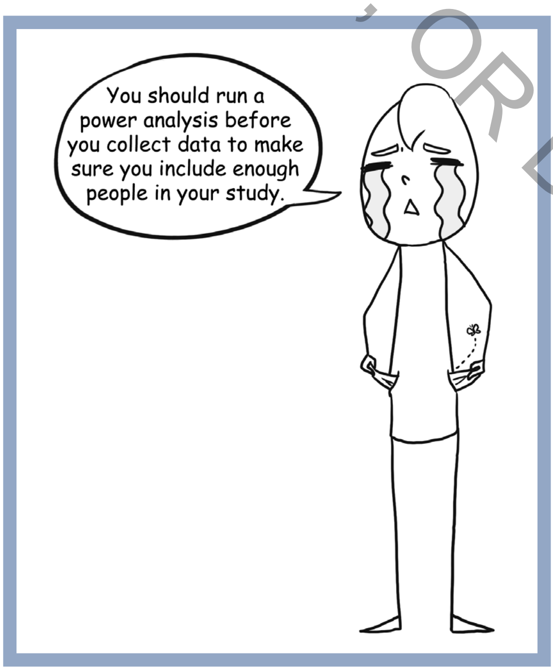
In the second box of the output, look at the correlation between Dollars and Happiness. The correlation is -.247, and based on a moderate correlation at +/–.30, we might be tempted to think our result shows a meaningful relationship between the two variables. But wait. The *p*-value must be rare, less than .05, to show an unusual relationship between amount of money in wallets and happiness. The significance value is .246, which is not equal to or less than .05. Therefore, the two variables are not related in our sample. Our research question asked, *Is the amount of money people have in their wallets related to happiness such that more money is related to greater happiness?* Now we can answer the research question: *No, the amount of money in the wallet and happiness were not related.* We can also discuss the outcome as failing to reject the null hypothesis: *The amount of money people had in their wallets was not related to their happiness.*

Power

Did we test enough participants to have a good chance of finding an effect if one existed? Recall power analysis from prior chapters. We need to know how many people are required to reveal a significant relationship, if one exists. Based on the table below, we should have included 85 participants for a good chance to show a significant outcome at $p \leq .05$ if a medium-sized effect exists. With our sample size of only 10 people, we had a poor chance of finding a significant outcome even if one was present. We lacked power due to a small sample size, and the potential for Type II error was high (review terms in Chapter 4 as needed). Of course, we used a small sample for the sake of a concise textbook example. Real life requires more.

	Small Effect Size	Medium Effect Size	Large Effect Size
Pearson's *r*	783	85	29

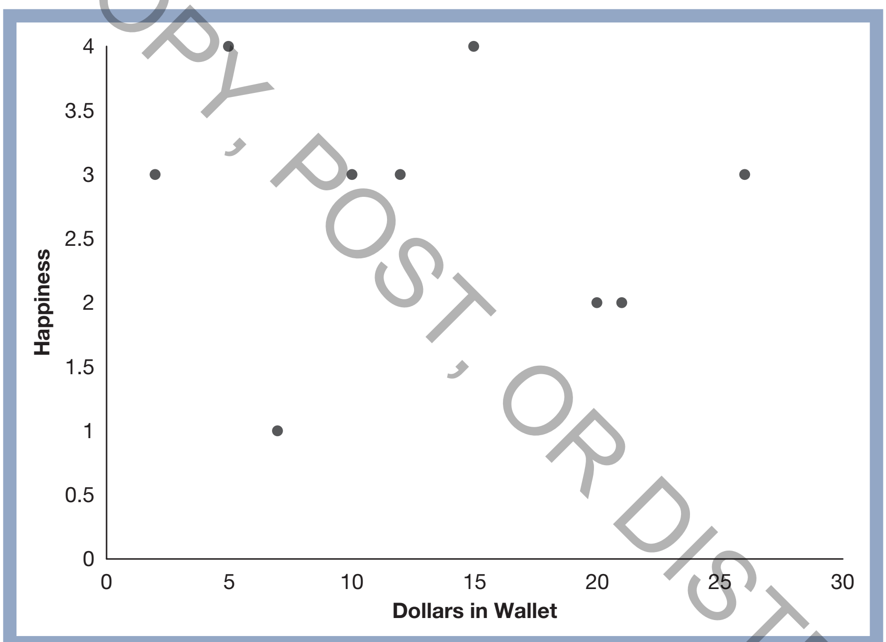
Note: Numbers in the table represent **total** sample size.



Graphs as Figures: Scatterplot

If you are going to include a scatterplot with your results, we prefer Excel for graphing. The final product in Excel is higher quality for publication. In Appendix B, we walk you through how to use Excel for graphing a scatterplot. Although Excel versions differ, the general ideas and formatting options remain fairly constant.

In our current example, if we wanted to include a figure to support our results, we would refer the reader to Figure 1 and offer a scatterplot. Below is a scatterplot for these data created in Excel.



APA Style for Pearson's *r*: Correlational Design

We rely on APA style to report our results, using all circled parts of the SPSS output for this example. As you know, we must report degrees of freedom (*df*) in the results section, and that piece of information is not on the output. We can calculate *df* using *N*, where *N* for correlation refers to number of participants in the sample. For Pearson's *r*, calculate *df* using $N - 2$, which in this case is $10 - 2 = 8$.

Method

Participants

Participants included 10 adults (5 men and 5 women) from a college sample. Ages ranged from 18 to 25 ($M = 21.07$, $SD = 2.35$), and ethnicities included 7 White and 3 Black individuals. We received IRB approval and treated all participants ethically.

Procedure

Students viewed a recruitment flyer placed on campus bulletin boards. Those interested in participating e-mailed the researcher with their availability and received an appointment time. At their appointment, participants reported the amount of money in their wallet as well as their age, gender, and ethnicity. Students also reported current happiness using a rating scale from 1 (*very unhappy*) to 4 (*very happy*). Participants received a $5 gift card for participating in the study.

Results

We examined the potential correlation between amount of money in wallets and ratings of happiness. Participants reported money as number of dollars ($M = 12.30$, $SD = 8.00$) and happiness ($M = 2.90$, $SD = 0.99$). Amount of money in wallets failed to relate to ratings of happiness, $r(8) = -.25$, $p = .246$.

A minor addition to the results section would allow you to include the scatterplot. We might indicate, "As shown in Figure 1, amount of money in wallets failed to relate to ratings of happiness, $r(8) = -.25$, $p = .246$."

In the APA-style results section above, we were careful to use the term *relate* rather than *cause* because neither variable was manipulated. We simply asked participants to report the number of dollars in their wallets and rate their happiness. Because the research design did not manipulate a variable, we could not establish cause and effect. Recall from Chapter 3 that researchers can only learn cause and effect with manipulation of what participants experience. That is, a true independent variable (IV) must be used in the design, and a dependent variable (DV) measures the outcome. With such a design, we can learn the cause (IV) and effect (DV).

PEARSON'S *r*: SEEKING CAUSE AND EFFECT

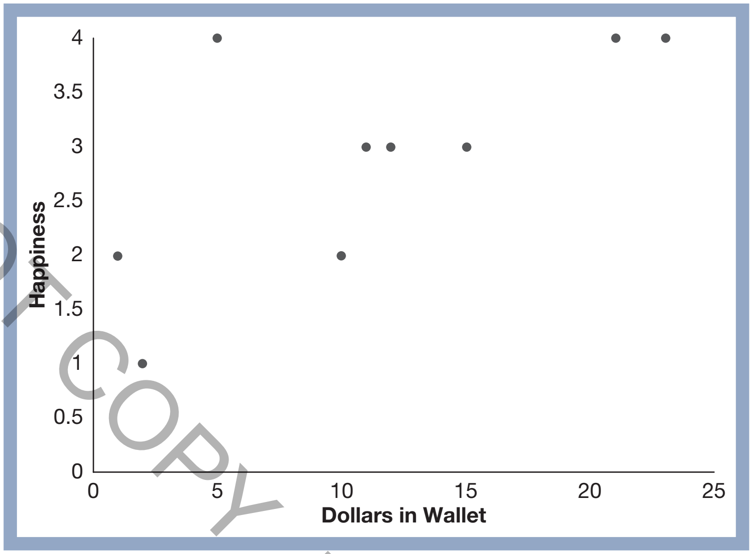
We could design a similar study with manipulation and a true IV. The research question would change slightly: *Does the amount of money people have in their wallets* affect *happiness such that more money* causes *greater happiness?* Notice that the research question indicates cause and effect by using the word "affect" as well as indicating that money "causes" happiness. In the prior example, with no manipulated variable, the research question only used the term "related" to avoid causation language: *Is the amount of money in a person's wallet* related *to happiness such that more money* relates *to greater happiness?* How might we change the correlational (not manipulated) study into an experimental study?

Let us assume participants entered the study with no money. Suppose we gave people different sums of money and then measured their happiness. We could pull some dollars from a basket and hand cash to each participant. Everyone would get a different amount, depending on what we pulled out. Then we could ask for ratings of happiness as the DV. Our null hypothesis would use cause-and-effect language: *The amount of money in people's wallets does not* affect *their happiness*. The research hypothesis states what we hope to find: *Money in wallets* affects *happiness such that more money* causes *more happiness*. Amount of money in wallets is still a ratio variable, and happiness using a rating scale still represents interval data. Our data might look like the following values. We again use Pearson's r to analyze the data, but this time our research design can tell us causation.

Dollars in Wallet	Happiness
5	4
10	2
15	3
21	4
23	4
12	3
1	2
11	3
2	1
17	3
12	2
20	4

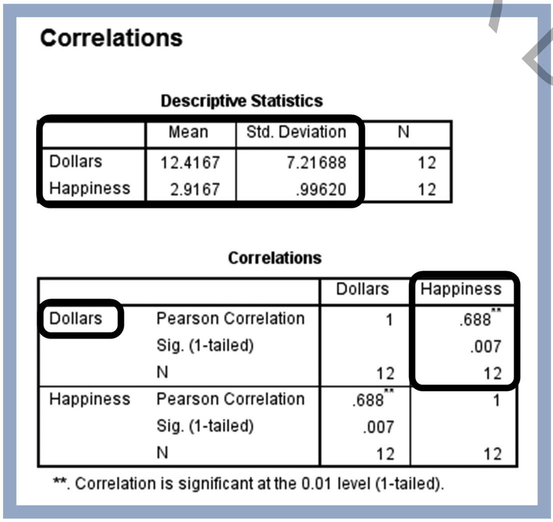
Previewing Data With a Scatterplot

We know that we can include a scatterplot in our APA results, but we can also create a scatterplot to provide a quick overview of our data. Given our research hypothesis of a positive correlation, we would expect to see a linear pattern of dots, with the approximate line increasing across the graph from left to right.



SPSS: Pearson's *r* (Seeking Cause and Effect)

Enter the data into SPSS as shown in the prior example. Remember that we expect a specific direction for the relationship (positive), allowing us to choose a one-tailed test and increase power. Be sure to ask for Descriptive Statistics as well by clicking Options and checking the box next to Means and standard deviations. Click Continue then OK for the output.



Correlations

Descriptive Statistics

	Mean	Std. Deviation	N
Dollars	12.4167	7.21688	12
Happiness	2.9167	.99620	12

Correlations

		Dollars	Happiness
Dollars	Pearson Correlation	1	.688**
	Sig. (1-tailed)		.007
	N	12	12
Happiness	Pearson Correlation	.688**	1
	Sig. (1-tailed)	.007	
	N	12	12

**. Correlation is significant at the 0.01 level (1-tailed).

On the output we can see that Dollars and Happiness correlate at .688 and a significance value below .05. Let us return to the research question: *Does the amount of money people have in their wallets affect happiness such that more money causes greater happiness?* Based on a *p*-value of .007 and a positive *r*-value, we have our answer: *Yes, the amount of money in the wallet affected happiness, with more money causing greater happiness.* We reject the null hypothesis of no effect in favor of the research hypothesis: *Money in wallets* affected *happiness such that more money* caused *more happiness*. Again, because we manipulated the amount of money participants put in their wallets, we know that the amount of money affected happiness ratings.

Effect Size: Coefficient of Determination

Coefficient of Determination. The coefficient of determination is the number that indicates the effect size of a significant *r*-value. To calculate the coefficient of determination, square the *r*-value.

All needed information for a results section is found on the SPSS output. The only exception is the value for effect size. Because this example yielded a significant effect, we must report effect size using the **coefficient of determination**. You already know that a correlation of near or beyond +/–.50 shows a strong relationship (Cohen, 1988), and the Pearson's *r* in this example is .688. Effect size is calculated by squaring the *r*-value. For r^2, .01 is considered a weak effect, .09 is a moderate effect, and .25 is a strong effect. You could have arrived at these values by squaring *r*-values of .10, .30, and .50, which you already knew indicated the strength of *r*.

Although Pearson's r can give you a good idea of the strength of a relationship, r^2 is the best measure for effect size because it is easier to interpret. In our example, the coefficient of determination is $.688^2 = .47$. An effect size of .47 means that the amount of money in wallets overlaps with happiness ratings by 47%. Another way of explaining effect size is to say that 47% of happiness ratings can be explained by the amount of money given to participants. Considering all other variables that likely influence happiness, explaining 47% of happiness by knowing the amount of money people have is impressive.

APA Style for Pearson's *r*: Experimental Design

Remember, APA style requires information about participants and procedure details related to variables in the research design. The method section operationalizes variables, and the results section reports values from the correlation analysis.

Method

Participants

Participants included 12 adults (6 men and 6 women) from a college sample. Ages ranged from 19 to 24 ($M = 22.17$, $SD = 1.55$), and ethnicities included 7 White, 3 Black, and 2 Latino individuals. We received IRB approval for the study and treated all participants ethically.

Procedure

Students viewed a recruitment flyer placed on bulletin boards on campus. Those interested in participating e-mailed the researcher with their availability and received an appointment time. At their appointments, participants received between $1 and $23, with the amount determined randomly. Students also reported current happiness using a rating scale from 1 (*very unhappy*) to 4 (*very happy*). Finally, participants reported their age, gender, and ethnicity, and they received a $5 gift card for participating in the study.

Results

We examined the potential correlation between amount of money given to participants and their ratings of happiness. Participants reported money as number of dollars ($M = 12.42$, $SD = 7.22$) and happiness on a rating scale from 1 to 4, with higher numbers indicating more happiness ($M = 2.92$, $SD = 1.00$). Amount of money in wallets affected ratings of happiness, $r(10) = .69$, $p = .007$, $r^2 = .47$. The more money people received to put in their wallets, the higher their ratings of happiness.

The APA-style results section above includes the word *affected* rather than *related* because manipulation occurred in this research design. Pearson's *r* is merely a statistical analysis. The way it is interpreted—relationship or cause and effect—depends on the research design. Although the SPSS analysis remains the same, the results section differs based on whether the design was a correlational study or an experiment, using wording specific to each design.

INACCURATE PEARSON'S *r*

Pearson's *r* is a statistic used to analyze two interval or ratio variables with many values. If neither variable is manipulated, Pearson's *r* quantifies a relationship between the two variables. On the other hand, if one variable is manipulated, a significant Pearson's *r* can show cause and effect. Your research question drives your research design and dictates whether you can look for cause and effect. Pearson's correlation coefficient analyzes the data you collect. It is a useful tool.

However, recall that Pearson's *r* only quantifies linear relationships. If the two interval or ratio variables in a research design are related in a linear way, *r* is a fine choice for analysis. Here we will explain several situations that can lead to erroneous results using the *r*-value. Three of the situations lead to an artificially low *r*-value, and two situations create an artificially high *r*-value.

WHEN PEARSON'S *r* FALSELY SHOWS NO RELATIONSHIP

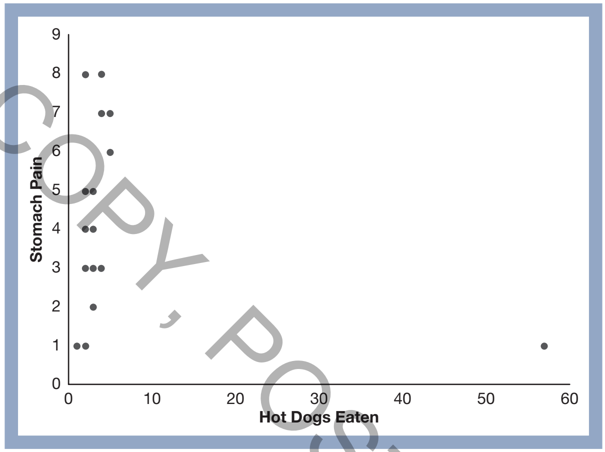
A low *r*-value is likely to be nonsignificant. If the sample size is quite large, a low *r*-value might be significant, but the effect size will be small. Either way, you will be disappointed that the two variables did not relate well to each other, or one variable did not cause changes in the other (with an experiment). But sometimes a low *r*-value is misleading. The three situations in which Pearson's *r* will be low even when a relationship clearly exists are based on (1) outliers, (2) a nonlinear relationship, and (3) restriction of range.

Outliers That Weaken Pearson's *r*

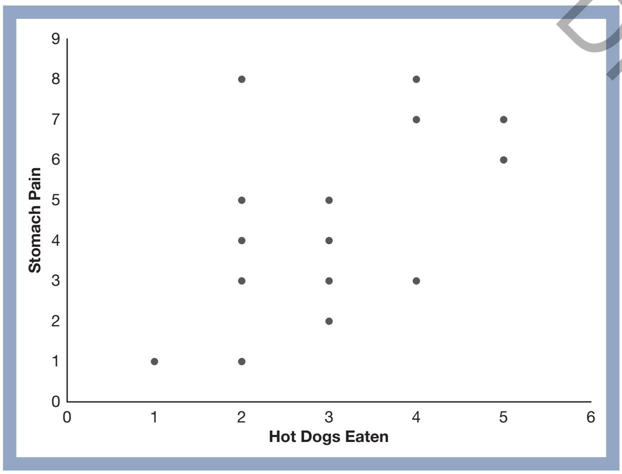
Outlier (for Ungrouped Data). An outlier is a data point that is out of the ordinary. In other words, an outlier is a data point outside of the general relationship pattern.

In psychology, we often want to know what would be true of most participants in a given situation. An **outlier (for ungrouped data)** is a data point based on a participant's response that is far from ordinary. For example, if we measured the potential relationship between number of hot dogs eaten and stomach pain on a scale from 1 (*I feel great*) to 8 (*I feel terrible*), we might find that most people in the sample choose to eat somewhere between 1 and 5 hot dogs. Imagine we happened to get a champion hot-dog eater in our sample who consumes 57 hot dogs and feels great. That participant would be far from ordinary and could obscure a fairly strong relationship by reducing Pearson's *r*.

Researchers should examine their data to see if the two variables of interest are linearly related by graphing the data. A graph will also show if a clear outlier or two exist in the data set. Consider the following graph of our hot-dog eaters.



Notice that the outlier obviously deviates from the rest of the data. With the outlier, a one-tailed correlational analysis reveals $r = -.31$, $p = .239$. In this situation, researchers may choose to remove the outlier and rerun Pearson's *r*. Look at a graph without the outlier.



In our example, the *r*-value is strong without the outlier, with $r = .53$, $p = .041$ (one-tailed test). Of course, ethical considerations require us to report the existence of an outlier, the removal, and the reanalysis. In fact, we recommend reporting the original *r*-value, providing a graph illustrating the outlier, and providing the subsequent *r*-value with the outlier removed. It is best to be transparent with readers about your approach to data analysis. An outlier should be visually obvious when graphing, and removing the outlier should be logical. But ultimately readers will decide if your approach was reasonable.

Nonlinear Relationship

Curvilinear Relationship. When the graph of a relationship is curved rather than a line, we call it a curvilinear relationship.

Because Pearson's *r* can only quantify linear relationships, a clear nonlinear relationship will yield a low *r*-value. For example, romantic partners should spend time together, but most of us also value time with other friends as well as time alone. Suppose we asked a sample of men to report average hours per day spent with their romantic partners and to rate their relationship satisfaction on a scale from 1 (*very unsatisfied*) to 9 (*very satisfied*). Take a look at the fictional graph of a **curvilinear relationship**.

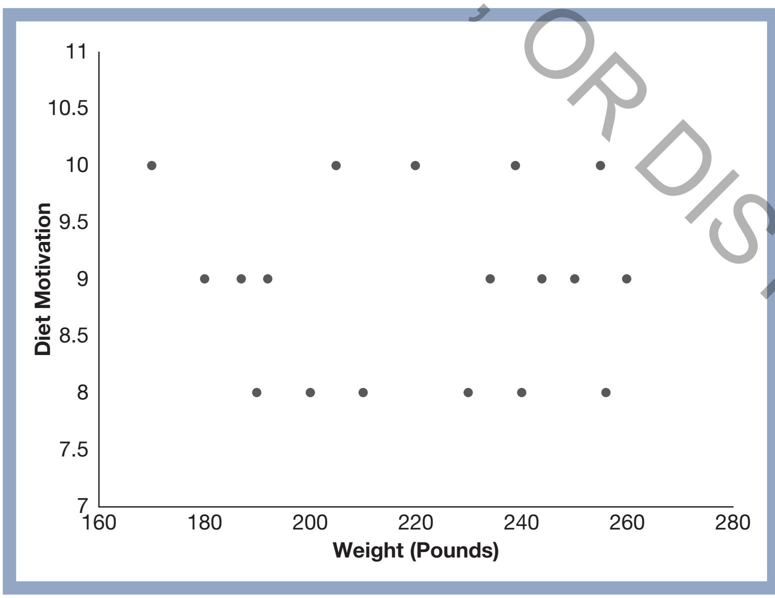
When a relationship is curved rather than linear, we call the shape curvilinear. Relationship satisfaction and time spent with a romantic partner are related in a curvilinear way. Both those who report very little time and those who report a great deal of

time spent with a romantic partner show less relationship satisfaction. Pearson's r will be low, suggesting no meaningful relation between the two variables. In fact, $r = -.03$ ($p = .882$) in the sample above. However, the graph illustrates a clear nonlinear relationship. Advanced statistics to analyze nonlinear relationships are beyond the scope of this book. But you should make a habit of graphing data to see if a low r-value can be explained by a nonlinear relationship. With a nonlinear pattern, a relationship indeed exists but cannot be captured by our analysis.

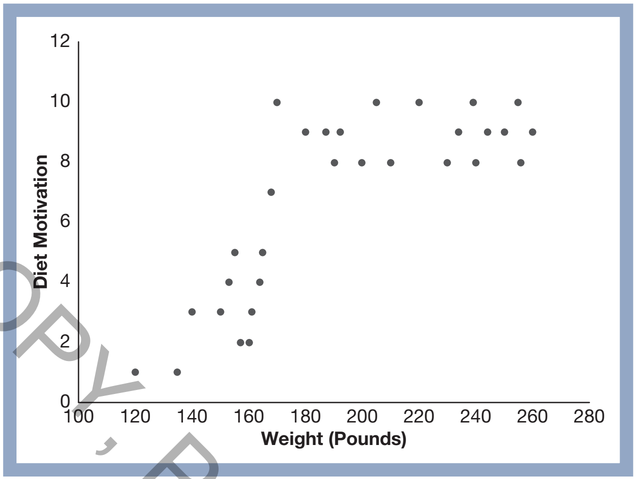
Restriction of Range

A final situation in which an r-value may be low even when a meaningful relationship exists is **restriction of range**. For interval and ratio data, variables have a range, such as adult male weights from 120 to 250 pounds and a diet-motivation scale from 1 to 10. Restriction of range occurs when one of your variables has a narrow range of values. For example, if all participants in your sample rated their diet motivation as 8, 9, or 10, the range on that variable would be restricted. The sample would contain no motivation values from 1 to 7. We might expect higher weight to be related to greater motivation to diet as indicated by a positive correlation and as tested using a one-tailed test. Based on a restriction of range, the relationship between diet-motivation ratings and weight among men might be illustrated by the following graph. The r-value for this sample is $-.40$ ($p = .433$), which shows no relationship between diet motivation and weight.

Restriction of Range. Restriction of range occurs when one variable in a data set has a narrow range of values. For example, if you are interested in sleep and grades, restriction of range would be indicated by a sample of participants who all sleep between 7 and 8 hours a night.

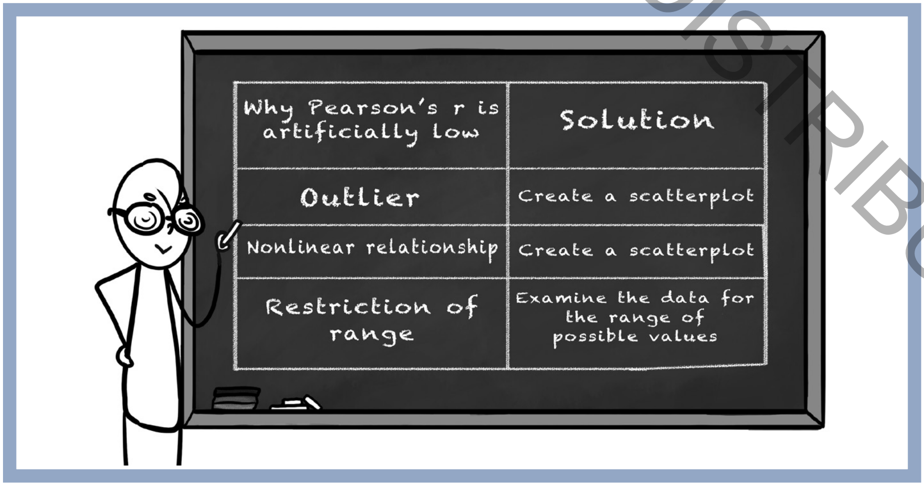


When we study interval or ratio variables, we want a wide range of possible values in the sample. If our motivation-weight example contained a wide range of possible motivation values, we might find a meaningful relationship as depicted in the following graph.



The *r*-value for the sample depicted in the prior graph is .79, with $p < .001$. Notice diet motivation ranges from 1 to 10, a wide range on the rating scale. It is also important that weight contains a wide range of values, or we would have restriction of range on weight.

When a variable is not represented by a wide range of values in our sample, Pearson's *r* may indicate no relationship. How will we know? In this situation, graphing the data will not show the problem. Instead, we must carefully review our variables, their possible ranges, and the spread of values on each (e.g., standard deviation). Logic is the antidote to restriction of range. If you think this problem exists in your sample, try to collect data from more participants until a wider range of values is represented.

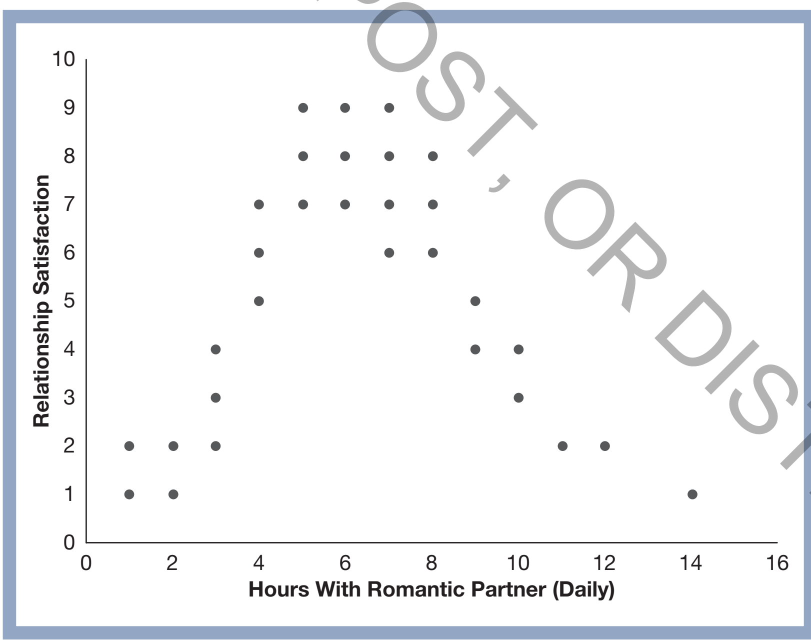


WHEN PEARSON'S *r* FALSELY SHOWS A RELATIONSHIP

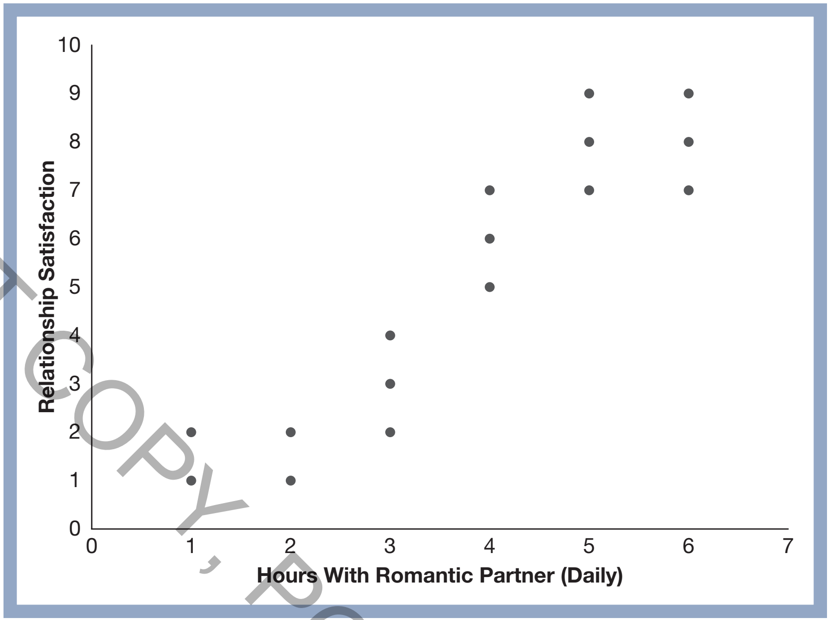
A strong *r*-value probably will indicate a significant relationship, even when the sample size is rather small. As hard-working researchers, a significant Pearson's *r* with a high absolute value makes us happy. We have something meaningful to report to the world. But we could be wrong. An *r*-value can be falsely high in two situations: (1) restriction of range on a curvilinear relationship and (2) outliers.

Restriction of Range on a Curvilinear Relationship

Remember that a curvilinear relationship cannot be quantified by Pearson's *r*. In our earlier example of daily time spent with a romantic partner and relationship satisfaction, we saw a fictional example in which satisfaction increased across hours spent together, but only to a point, and then satisfaction began to decrease. For a reminder, the graph follows, and the *r*-value was –.03.



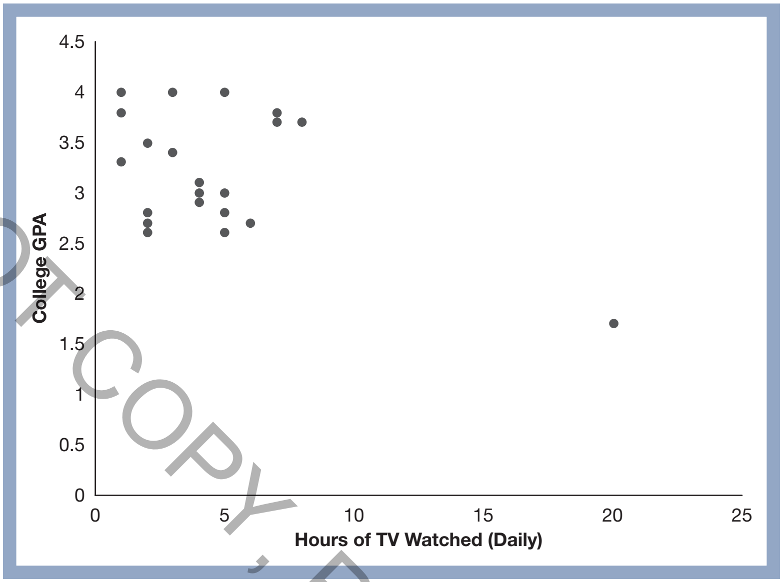
What if we restricted the range on a curvilinear relationship? We might collect a sample of men who see their romantic partners between one and six hours daily. A graph of the sample might look like the following.



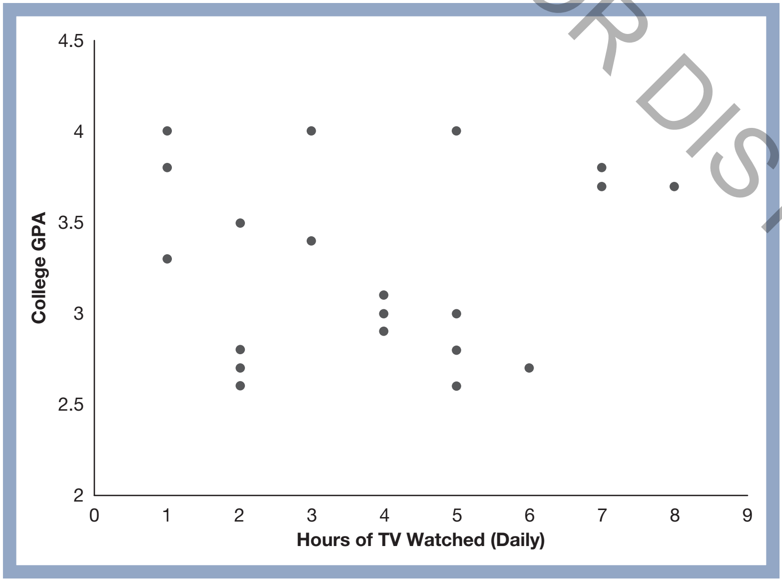
With restriction of range, we would only see a linear relationship between the two variables, and the *r*-value would be misleading ($r = .92$, $p < .001$). From our data, we might draw the conclusion that the more time people spend with each other, the more satisfied they are with the relationship. Indeed, we could publish this interesting result. Would we be wrong? Yes and no. Our data truly indicate a linear relationship, and over time, more research across the world would address the research question in many different ways, including longer durations of time spent together. Eventually the entire picture would emerge, showing a curvilinear relationship. That is the nature of science and the benefit of an entire scientific community sharing ideas.

Outliers That Strengthen Pearson's *r*

Just as outliers in the data can reduce Pearson's *r*, outliers can strengthen the *r*-value. Either way, researchers usually remove outliers. When unusual data points help to create a linear pattern, we can see the problem by graphing. Let us consider a hypothetical study of television (TV) hours watched daily and college grade point average (GPA). We likely would expect a negative relationship. Notice the outlier in the scatterplot below.



For these data, Pearson's $r = -.47$, $p = .016$, suggesting a significant negative relationship between number of TV hours and college GPA. The more TV watched, the lower the GPA of students in the sample. Most researchers would agree that removal of the outlier paints a clearer and more accurate picture of TV watching and college GPA. When Pearson's r is recalculated without the outlier, $r = .05$, $p = .419$. The scatterplot below illustrates the relationship more honestly.



Although it may seem disheartening to find a strong Pearson's *r* and have to abandon a significant result, always keep in mind that our job as researchers is to communicate the truth. An obvious outlier (or more than one outlier in a large sample) must be removed to show the general relationship pattern, even if that means revealing that no linear pattern exists. At this point in your education, remove only outliers that are obvious in a scatterplot. If you are unsure of whether or not a value is an outlier, talk with your instructor.

SUMMARY

If your research question asks about a relationship between two interval or ratio variables with many values, Pearson's *r* will analyze the data. The *r*-value can tell you if the two variables are related, as long as neither variable is manipulated. Alternatively, the *r*-value can tell you cause and effect as long as one variable is manipulated. Pearson's *r* does not dictate cause and effect. Pearson's *r* merely offers an analysis, and you must provide the explanation. At times, the *r*-value can be artificially inflated or low, and creating a scatterplot reveals the majority of problems that might occur, such as a nonlinear relationship or outliers. A restriction-of-range problem relies on logical evaluation of each variable's range. If no problems are revealed, and the *r*-value is significant, be sure to report effect size. Pearson's *r* offers a powerful tool to address a research question about a relationship or cause and effect.

REVIEW OF TERMS

Coefficient of Determination

Curvilinear Relationship

Directional Test (One-Tailed Test)

Nondirectional Test (Two-Tailed Test)

Outlier (for Ungrouped Data)

Pearson's *r* (Correlation Coefficient)

Restriction of Range

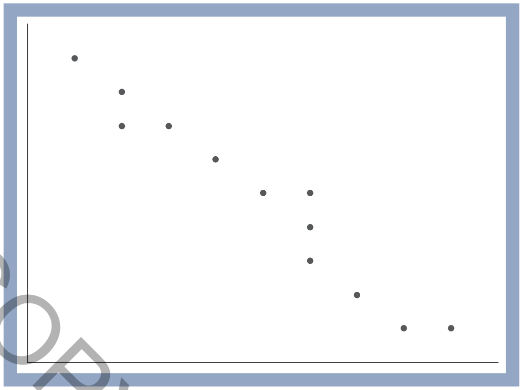
Scatterplot

PRACTICE ITEMS

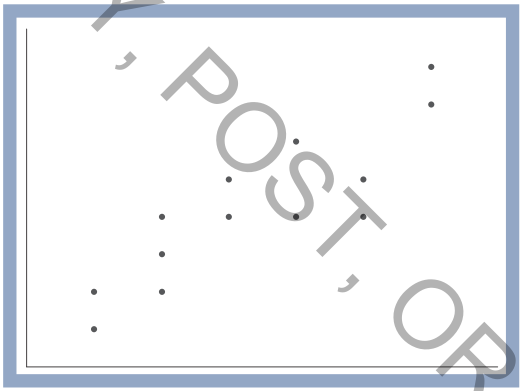
1. What is the difference between a study design analyzed by Pearson's *r* and a design analyzed by a two-way χ^2?
2. How does Pearson's *r* change based on testing for a relationship versus causation?

3. On the following graphs, indicate whether the linear relationship is positive, negative, or so scattered that you cannot determine a relationship.

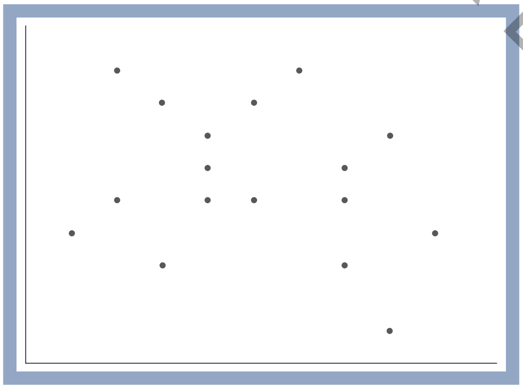
a.



b.



c.



4. Below are examples of r-values. Please indicate the direction (positive or negative) and strength (weak, moderate, or strong) of each relationship.

a. .90

b. −.31

c. .17

d. .02

e. −.57

f. .68

5. How do you calculate *df* for Pearson's *r*?

6. When can correlation convey causation?

7. What is the name of effect size for Pearson's *r*, and how is it calculated?

8. For each of the following, decide whether effect size should be calculated. If it should be calculated, do so, and indicate if the effect size is small, medium, or large.

   a. $r = .67, p = .002$

   b. $r = -.12, p = .062$

   c. $r = -.35, p = .049$

   d. $r = .22, p = .822$

* * *

For each of the following studies, (a) restate the research question as a research hypothesis, including whether it should be one- or two-tailed, and state the null hypothesis, (b) check for any problems with the data (e.g., an obvious outlier), (c) determine how many participants are needed for adequate power, and (d) enter and analyze the data as well as write an APA-style results section. Notice that we have written method sections for you to clarify study details.

9. In a review paper on the effects of sleep deprivation and attention (Lim & Dinges, 2008), you read about a positive relationship between sleep and attention such that those who get more sleep have better attention. You decide to study the relationship between these two variables using a sample of high-anxiety participants. Although prior research showed a positive relationship between sleep and attention, you suspect that amount of sleep and attention instead could be negatively related among people with high anxiety. In other words, the relationship might be positive or negative. You ask the research question: *Among people with high anxiety, is amount of sleep related to attention?* In the lab, you ask high-anxiety individuals to

estimate how much sleep they get at night, and then you give them a test of attention with 100 items to recognize. Their attention score will be the number of items they can recall. You collect the following data. Note that the three tables should be entered as two continuous columns in SPSS, and you will have 36 rows in the data file.

Hours Slept	Attention
4	54
4.5	47
5.5	62
5.5	48
6	47
6	60
6	52
6	70
6.5	68
7	70
7	98
7	53

(data continued)

Hours Slept	Attention
7.5	78
7.5	82
7.5	78
7.5	78
7.5	78
7.5	53
8	89
8	90
8	73
8.5	83
8.5	89
8.5	89

(data continued)

Hours Slept	Attention
9	80
9	80
9	94
9	93
9	62
9	76
9.5	76
9.5	82
9.5	92
10	95
10.5	97
11	87

Method

Participants

Participants included 36 college students (20 women and 16 men) recruited through the use of a flyer posted in college residence halls. The flyer specified that those who participated in the study must generally sleep with their televisions on at night. Ethnicities included 10 White, 21 Black, and 5 Asian individuals with a mean age of 21.15 years (*SD* = 2.33). All participants received ethical treatment, and the IRB approved the procedure.

Materials

Attention test. The attention test consisted of a series of 50 neutral pictures shown at 5-sec intervals. Next, participants saw 100 more pictures, 50 of which they saw previously. For each picture, participants indicated whether they had seen the picture before, with scores ranging from 0 correct to 100 correct answers.

Procedure

Participants e-mailed the examiner at the e-mail address provided on the flier. When they arrived at the lab, all participants read and signed informed consent, took the attention test, reported their average number of hours of sleep each night in the last 2 weeks, and provided demographic information.

10. Suppose after you ran a study on sleep and attention, you wanted to find out what other variables might be related to attention. You might read an article by Abdullaev, Posner, Nunnally, and Dishion (2010) linking chronic marijuana use as a teenager to problems with attention in young adulthood. You wonder, *Is use of marijuana positively related to attention problems in college students?* You collect the following data using an attention test based on number of items recalled out of 100 items. In this study, the memory test is provided online through a study link so people can participate anonymously.

Joints	Attention
3	54
3	47
3	45
3	57
3	62
3	48
3	47
2.5	60
2.5	62
2.5	60
2.5	32
2.5	78

(data continued)

Joints	Attention
2	78
2	82
2	88
2	90
2	78
2	78
2	78
2	53
1.5	82
1.5	85
1.5	62
1.5	82

(data continued)

Joints	Attention
1.5	80
1	80
1	81
1	72
1	94
1	93
.5	62
.5	76
.5	80
0	83
0	97
0	60

Method

Participants

Participants included 36 college students (18 men and 18 women) recruited through the use of a flyer posted in college residence halls. Ethnicities included 17 White, 5 Black, and 4 Asian individuals as well as 10 individuals who chose not to disclose their ethnicity, and 20.25 years (SD = 1.13) represented the mean age of participants. All participants received ethical treatment, and the IRB approved the procedure.

Materials

Attention test. The online attention test consisted of a series of 50 neutral pictures shown at 5-sec intervals. Next, participants saw 100 more pictures, 50 of which they viewed previously. For each picture, participants indicated whether they had seen the picture before by clicking on the Y key if they had previously seen the picture or the N key if they had not. Scores on this test ranged from 0 to 100 correct answers.

Procedure

Due to the sensitive nature of this project, participants completed the study online to allow anonymity. Participants e-mailed the examiner at the e-mail address provided on the flier and received a link to the study. First, all participants read and agreed to the informed consent by clicking on a button next to the statement, "I give my consent to participate in this study." Next, they took the attention test, reported the average number of joints smoked per week when they were between the ages of 16 and 18, and provided demographic information.

11. Your 25-year-old brother seems to be unable to keep more than a part-time job. He seems happy, but you recall reading a paper finding that unemployed men score lower on the personality trait of agreeableness than those who are employed (Boyce, Wood, Daly, & Sedikides, 2015). You wonder if men who work fewer hours have lower agreeableness than men who work more hours. Using the agreeableness scale of the Big Five Inventory (BFI; John & Srivastava, 1999), you ask men working either full or part time at various establishments in your community to complete the BFI and report about how many hours per week they have worked over the last six months. You ask the research question: *Is agreeableness related to number of hours worked such that more agreeableness is related with more hours worked?* You collect the following data.

Hours Worked	Agreeableness	Hours Worked (data continued)	Agreeableness	Hours Worked (data continued)	Agreeableness
12	39	20	15	31	41
12	30	20	22	32	16
13	25	20	45	35	38
14	25	20	30	35	26
15	14	23	15	36	20
15	21	25	31	36	41
15	29	26	27	38	30
16	35	27	17	40	41
17	17	29	32	41	27
18	37	29	24	42	31
18	28	30	37	45	23
20	45	30	40	45	12

Method

Participants

Participants included 36 adult men (mean age = 37.51, *SD* = 10.87) recruited at five local businesses (grocery store, bank, child-care facility, gym, bookstore) by approaching them and asking for participation. All participants received ethical treatment, and the IRB approved the procedure.

Materials

Big Five Inventory (BFI). The BFI is a 44-item scale that assesses personality through the lexical Big Five factors of personality (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism). Participants used a 5-point Likert scale to rate how strongly they agreed or disagreed with statements about their personality. Research on North American samples illustrates strong reliability of the BFI (α = .75 – .90; John & Srivastava, 1999). In this study, we only used the 9 Agreeableness items, with 45 items representing the maximum score.

Procedure

Researchers approached participants in the five locations listed above and asked if they worked at the establishment and would consider participating in a study. After agreement, participants received a premade packet with the consent form, BFI, and demographics form that asked their age, gender, and number of hours worked to complete and mail back to the examiner within 1 week. Of the 47 packets distributed, participants returned 36.

REFERENCES

Abdullaev, Y., Posner, M., Nunnally, R., & Dishion, T. (2010). Research report: Functional MRI evidence for inefficient attentional control in adolescent chronic cannabis abuse. *Behavioural Brain Research, 215*(1), 45–57. doi:10.1016/j.bbr.2010.06.023

Boyce, C. J., Wood, A. M., Daly, M., & Sedikides, C. (2015). Personality change following unemployment. *Journal of Applied Psychology, 4,* 991–1011. doi: 10.1037/a0038647

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. doi:10.1037/0033-2909.112.1.155

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (vol. 2, pp. 102–138). New York, NY: Guilford Press.

Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences, 1129*, 305–322. doi:10.1196/annals.1417.002