CHAPTER 2

Understanding the Research Literature
You have selected your research topic. Now you need to discover what research has already been done on the topic. It is time for you to peruse the research literature. You will probably find some research on your topic, but it is unlikely that anyone has done exactly the same study that you have in mind. We often tell our students to think through their topic before going to the literature because it is easier to be creative about your project before you have looked at the other approaches. Indeed, once you have read a few articles in the area, it may be very difficult to think of an original idea!

Although it is unlikely that a study identical to yours has already been done, reading the literature will give you an idea about the kinds of problems that other researchers have had, and you can assess their solutions to those problems. The literature can provide a historical context for your study by describing what has been done and what remains to be explored. You will also find valuable information on how to measure variables, how to control variables, and, in experimental studies, how to manipulate variables. This sounds wonderful, but how do you start to search through all the research to find studies relevant to your topic? In the “good old days,” this required many hours (or days) of wading through heavy indexes of research (just imagine a dozen or more telephone directories). Today, there are a number of computer indexes that require no heavy lifting.

**SEARCHING THE LITERATURE**

There are a number of bibliographic databases for the psychology literature, including Proquest, ERIC, and PsycINFO. We prefer PsycINFO, an index that is produced by the
American Psychological Association and is probably the most widely used bibliographic search engine for English-language journals (http://www.apa.org/psycinfo/products/psycinfo.html).

What is the psychology literature? When we use this term, we are usually referring to original research published in peer-reviewed journals. These journals can be easily recognized by their common layout. They begin with an abstract (a short summary) and include an introduction, a method section, a results section, and a discussion. But the literature also includes review articles, books, chapters in books, edited volumes, and chapters in edited volumes. These sources are not usually where original research is published. Rather, they are usually a summary of a collection of research studies in a particular area. Although these sources can be useful by helping you to put the research in context, it is better to read the original research and draw your own conclusions.

Not included in the list above are newspapers, magazines (including Psychology Today), or Web sites. There are two reasons that scientific research is not published in these media: (1) the presence of advertising and (2) the lack of peer review. Let’s talk about advertising. Newspapers and magazines contain a lot of advertising of products and services. Indeed, these publications would not be viable without advertisers. Editors of newspapers and magazines must keep their advertisers happy; therefore, we, as consumers of the information in these publications, cannot be confident that reporting will be unbiased. Now, we are not saying that all newspaper and magazine reports are biased, but they could be. For example, imagine you have conducted a study that shows that drinking Brand X beer can lead to spontaneous uncontrollable hiccupping. You send your report to a magazine that has Brand X beer as a major advertiser. It’s pretty unlikely that the magazine editor will accept your paper for publication. Yes, your research might be terrific, but the potential damage to the advertiser in terms of sales of the product might lead the editor to reject your study for publication. This is why you will not see advertising in journals. Journal editors, by not permitting advertising, avoid the conflict of interest problem that advertising brings.

The second reason that original scientific research is not usually published in newspapers or magazines is that neither of these sources requires review of the research by expert peers. Some sources that you can search are peer reviewed, and some are not. It is important to understand the peer review process because it is a fundamental safeguard of quality in research. As the name implies, peer review is a process whereby the editor of a journal sends submitted manuscripts out to be reviewed by other researchers in the same field of study. The manuscript is read and critiqued by peers who have expertise in the area. The review is usually blind; this means that the name(s) of the author(s) of the manuscript is removed from the manuscript before the copies are sent to the peer reviewers. Blind review also means that the editor does not reveal the reviewers’ names to the author(s) of the manuscript. Blind review helps to guard against any personal conflicts that may be present among researchers and to facilitate a fair review of the research. The editor receives the reviews and decides whether the paper should be accepted as submitted, accepted with minor changes, accepted with revisions, or not accepted at all. If the manuscript is accepted but changes are required, the author(s) is given the opportunity to make the changes necessary to satisfy the editor and to address the concerns of the reviewers. If the study has major flaws, it may have to be redone and resubmitted (of course, the researchers can submit the manuscript to another journal and hope for a more positive review).
Peer review helps to maintain a high standard of quality in research. Without peer review, shoddy, or even fraudulent, research might be published that could send other researchers off on a wild-goose chase that could last for years. Keep in mind that books and magazines (e.g., *Psychology Today*) are usually not peer reviewed.

**NOTE:** In our courses, we do not permit the use of *Psychology Today* as a source. Students often question this because a lot of psychological research is reported in *Psychology Today*. Along with the advertising and peer review issues, which we discussed above, there are two other reasons for our decision to not permit this magazine as a source for student papers. One reason is simple—we do not want our students to use secondary sources of any kind. Second, the writers for *Psychology Today* are not scientists; they may have some academic background in the area, but they are writing for the magazine as reporters. *Psychology Today*, like any other magazine, needs to sell magazines. Articles written by reporters will not be published unless the editor believes that the report will help sales.

**FYI**

There are two notable exceptions to the “no ads in journals” rule. Ironically, the two most prestigious scientific journals worldwide are full of advertisements. *Science* and *Nature* are peer reviewed, highly respected, and widely circulated.

When you are ready to use a bibliographic database, your first step is to select the appropriate search terms. Often students will complain that they cannot find anything in the literature on a particular topic even though a lot of research is there. Usually, the problem is that they have not used the correct terms when doing the search. Using the correct terms is crucial; fortunately, the databases have a thesaurus of keywords to help you.

Imagine you want to find articles on treating seasonal affective disorder (SAD) with light therapy. You can go directly to the thesaurus in PsycINFO and look for “seasonal affective disorder.” You will find that the term is used in the database; you will see a brief definition of SAD and the information that the term was introduced in 1991. PsycINFO also provides broader terms that would include SAD and narrower terms that are relevant. One of the narrower terms is *phototherapy*. That sounds useful. You can select “phototherapy” by clicking on the box, and you can similarly select “seasonal affective disorder.” The default is to connect these two phrases with “or,” but if you want articles that contain both phrases, you should choose the “and” option. When you click “add” and search (DE “Seasonal Affective Disorder”) and (DE “Phototherapy”), you will find 276 publications with those keywords. You might want to limit your search to only peer-reviewed articles in English by clicking “refine search” and selecting *peer reviewed journals* under *publication type*. Below that is a box for selecting English language. If you do this and search again, you will get 235 hits. That’s a more manageable number, but probably still more than you want to read. Let’s limit the search to only *original journal articles* under *document type*. This time the search finds only 25 articles. That seems like a reasonable number. Now you can click on each one and read the abstract of the article. The abstract is a short but comprehensive summary of the
article, and based on your perusal of the abstracts, you can decide which articles you want to read in their entirety. If you had done the search the way we have described, you should find that the second article is titled “Light Therapy for Seasonal Affective Disorder With Blue ‘Narrow-Band’ Light-Emitting Diodes (LEDS).” This might be one you choose to read. If your library has access to the full-text electronic version of the paper, you will be able to download it immediately. If your library does not have access, then you will have to order a copy, called a reprint. You can also write to the principal investigator and request a reprint. This may take a while, so you should begin your literature search early. If you leave reprint requests until the last minute, you may not be able to get the articles that are most relevant to your research topic.

Here is another example. Suppose we are interested in how the color of a room might influence mood. A search using the terms mood and room color produces nothing. Perhaps room is too restrictive. A search with mood and color gives 237 articles. Now let’s refine (limit) the search to just peer-reviewed articles in English. Wow, still 167 articles. PsycINFO has a bar that reads “Narrow results by subject:” and in the list is “color.” What happens if we click that? We get 31 articles that deal with color and mood—that’s perfect! Apparently, this search requires that color be a keyword and not just a word that appears in the title or abstract. The fifth article in the list looks interesting; it is titled “Effects of Colour of Light on Nonvisual Psychological Processes” by Igor Knez at the University of Gävle, Sweden. Again, your library may or may not have access to the full text of the article. Be sure to consult with your reference librarians. These people are well educated and typically underutilized by students. If you have a question, ask your librarian.

THE RESEARCH ARTICLE

Now that you have a research article in hand, let’s examine each section of the paper separately. When reading the literature, you need to understand that each section has a purpose and contains specific types of information. Once you know how research articles are written, you will know which section will contain the information you need. The following discussion is presented to help you read the literature. You will find information on writing a research article in Chapter 14 (“Communicating in Psychology”).

The Abstract

The abstract is a comprehensive summary of the article and describes what was done, to whom, and what was found. Online bibliographic search engines such as PsycINFO provide the title and the abstract. If the title seems relevant, you can read the abstract. It should provide you with enough information to decide whether you want to read the entire article.

The Introduction

The introduction directly follows the abstract. Here the author(s) provides background on the research problem. You will find a description of the relevant research (this is the researcher’s literature search) and how it logically leads to the research being reported in
the article. Usually, near the end of the introduction, you will find a description of the research hypothesis of the author. Again, for information on writing an introduction, see Chapter 14.

Knez (2001) has an introduction with a typical layout. He begins with a general statement about the state of knowledge in the research area. He then presents a discussion of the previous research organized by variables. He describes the various independent variables that have been identified as important and discusses various confounding variables and how these should be controlled. He also discusses a theoretical framework that is based largely on his own research. And finally, he defines the purpose of the present research, how it will solve some of the problems identified in the literature review, and why it is important.

Before we continue to a discussion of the method section, it is probably a good idea to review the types of variables you will read about in the introduction of many research articles.

**Independent Variable**

You will recall that an independent variable (IV) is the variable in an experiment that is manipulated by the researcher. The researcher chooses levels of the IV that he or she thinks will have effects on some response measure. The researcher then assigns participants to each level of the IV (or all levels in the case of repeated-measures designs) and compares differences in response measures to see if the IV had an effect. You will recall that some variables are not true IVs. The values of these participant variables may be inherent in the participants. Examples include gender, age, disability type, and so on. Or participants might have self-selected the value of the variable. For example, differences in school success between children attending private and public schools is a comparison on a participant variable where participants have, in effect, assigned themselves to the values of the variable. In either case, studies of group differences on participant variables are not true experiments; rather they are quasi-experiments. Remember that a true independent variable is under the direct control of the researcher. The researcher chooses the values of the variable and assigns participants to each and then looks to see if that manipulation has any effect on their responses, the dependent variable. In an experiment, the independent variable can be thought of as the cause in a cause-effect relationship.

**FYI**

Statistical packages such as SPSS do not distinguish between true independent variables and participant (or subject) variables. They refer to both as independent variables.

**Dependent Variable**

The dependent variable (DV) in psychological research is some response measure that we think will be influenced by our independent variable. Reading comprehension might be a dependent variable, and we might measure number correct on a comprehension quiz as our operational definition of reading comprehension. Or we might measure depression by having
participants rate how they feel on a scale. In an experiment, the dependent variable can be thought of as the effect in a cause-effect relationship.

When we are trying to determine patterns of responding by measuring variables, we are always concerned with the natural variability of participants’ responses. Of course, our goal in research is to explain some of this variability. For example, if your research question is “Do students who read a lot understand better what they read?” then you are in a sense trying to account for the variability in student reading comprehension by determining how much they read. This is the variability that you are interested in explaining with your relationship. However, some variability is outside our primary interest. For example, if we are trying to determine if classroom technology improves learning, we are not interested in variables such as temperature of the classroom, time of day of the class, or ability of the instructor. Rather, we want to control or account for these variables so that we can better assess the effect of our primary variable (i.e., technology).

If other variables that might have affected the DV have been controlled in some way, then the researcher can conclude that differences in the DV are a result of, or caused by, the IV manipulation. This is the core of the experimental design, and to the degree that other variables have been controlled, we can be more confident in making causal inferences with these designs than we can from nonexperimental research. We discuss the various ways to control these other variables in Chapter 4.

Conceptual Exercise 2A

1. Identify the independent and dependent variables for each of the following:
   a. Reaction time decreased when more practice trials were given.
   b. Amount of exercise had an effect on depression ratings.

Moderating Variables

Many cold remedies display warnings that they should never be taken with alcohol. It is often the case that these drugs can cause drowsiness, but this cause-and-effect relationship is increased with the consumption of alcohol. In this example, alcohol is acting as a moderating variable by amplifying the drowsiness effect of the drug. Moderating variables act to influence the relationship between the independent and dependent variables. A moderating variable can increase, decrease, or even reverse the relationship between the independent and dependent variables. If, as discussed above, the independent variable is the cause in the cause-and-effect relationship, and the dependent variable is the effect in the cause-and-effect relationship, then the moderating variable is a third influence that must be taken into account to clearly describe the cause-and-effect relationship. For example, in his famous studies on obedience, Milgram (1974) found that the actions of a confederate-companion (someone posing as a participant who is actually part of the study) could produce a strong moderating effect. When the companion-participant agreed to shock the learner, 93% of the
true participants continued administering shocks, but when the confederate disobeyed the order, only 10% of the true participants continued.

In the Knez (2001) study of the effect of light on mood, gender was identified as a moderating variable. Relative to cool light, warm lighting can produce a more positive emotional response in women than it does in men. Therefore, the influence that lighting has on mood is moderated by the gender of the participant (for more discussion of moderating variables, see the discussion of factorial designs in Chapter 7).

Mediating Variables

Sometimes the relationship between cause and effect is directly linked—your baseball strikes a window and it breaks. However, there are many instances when this relationship is anything but direct—you look at a bright light and your pupil constricts. Certainly, this is a cause-and-effect relationship, but there are many intervening steps. Suppose you just had an eye examination and the doctor used eyedrops to dilate your pupils. What will this do to the cause-and-effect relationship? You leave the office and go into the bright sunlight and... nothing, no pupil constriction. Clearly, the eyedrops are acting on some mediating variable between the light and the pupil constriction.

Identifying mediating variables may be centrally important to your research or entirely trivial depending on how the research fits into the particular theory. For much behavioral research, the mediating variables may be unimportant. Instead, the focus is on identifying and describing the environmental cues (cause) that elicit behavior (effect). Contrast this position with cognitive research, where much of the focus is on identifying mediating variables.

In the Knez (2001) study, the identification of a mediating variable was an important point. He was trying to show that the characteristics of light do not directly influence cognitive performance but, rather, that the light influences the participant’s mood, and that change in mood, in turn, affects the participant’s performance.

The introduction gives us a good understanding of the important variables in the study. In the method section, we will find detailed information about how the study was conducted.

The Method Section

Although the variables of the research are defined and discussed in the introduction, it is in the method section where you will read the details of exactly how these variables are measured, manipulated, or controlled. Indeed, if there is a theme to the method section, it would be details, details, details. You should find enough (dare we say it again) details in the method section to replicate the study on your own. That is, you should have all the necessary information to repeat the study as it was done by the authors, with different participants, of course.

The method section is typically divided into a number of subsections usually separated with subheads (for more detail on writing a method section, see Chapter 14). The first section, typically called Participants, if human, and Subjects, if animal, will provide information about who or what participated in the research. You will read how the participants were recruited into the study or how the animals were obtained. Demographic information will be included such as age, gender, race, and education level. If nonhuman subjects were
used, details of care and housing will be described. Of course, you will also read how many participants were included.

You may find a subsection called **Materials** and/or **Apparatus**. Here you will find manufacturers and model numbers of equipment and apparatus, details of tests and measures and the conditions of testing, and often a description of any stimulus materials used. Somewhere in the method section, you will read a description of the research design and the statistical tests. It does not necessarily make for good reading, but the purpose is to provide fine detail, not to entertain. It is in the method section where we find out exactly what was done. The procedure is described in a subsection called, well, **Procedure**. This is often written as a chronological sequence of what happened to the participants. Again, as with all the subsections of the method, it is written in painstaking detail.

In our example article, Knez (2001) tells us he had 54 women and 54 men, all 18 years old, and all in high school. The basic design is described as a factorial between-subjects (his word, not ours) design (see Chapter 7) with three different lights and two genders. He also describes the testing conditions, details of the lighting (the independent variable), and the various dependent measures. He describes the procedure of the experiment, providing the time of day of testing, the information participants were given, and order in which the tests were administered. At the end of the section, we have enough information that we could probably replicate the study exactly as he had done.

After reading the introduction and the method section, we now know quite a bit about the research area, the current researcher’s study, and how he or she carried it out. It is time to find out what happened.

### The Results

The results section is the part of the paper that is the most exciting. This is where we learn whether or not the data support the research hypothesis. Typically, the section begins with a general statement addressing that very point (i.e., did the data support or fail to support the researcher’s hypothesis?). As with the other sections, more detail on writing an article is presented in Chapter 14.

The results section is the most important section of a research paper. Unfortunately, students can become overwhelmed by all the statistics. Even students who have done very well in their statistics courses can find the results sections of most research papers impossible to understand. The problem is that basic statistics courses cover **basic statistics** such as measures of **central tendency** (the mean, median, and mode) and measures of **variability** (the range, variance, and standard deviation). Of course, these statistics will appear in the results section and are widely used for describing and summarizing the data.

The problem is often in the **tests of significance**. Your basic statistics course probably covered the **z test**, **t test**, analysis of variance (ANOVA), and the associated **F** test. You may have also learned about **correlation** and simple regression and perhaps chi-square tests. These are good statistics and are used sometimes in research, but unfortunately, when you go to read the literature, you will face statistical tests that you may have never heard of. We do not intend to teach advanced statistics here, but we do want to provide you with a conceptual understanding of these statistics so that when you read the literature, you will have
at least a basic understanding of these procedures. So as briefly as we can, we are going to review statistics. No, you do not need a calculator; this review is at a conceptual level only, but in Chapter 13, we provide more of the nitty-gritty of basic statistics that you may need to do a research project of your own.

In research, statistics are used for two purposes. The first is to summarize all the data and make it simpler to talk about the outcome of research. These are typically called descriptive statistics. The second purpose is to test research hypotheses, and these are called inferential statistics.

**Descriptive Statistics**

**Descriptive statistics** include measures of central tendency, variability, and the strength of the relationship between variables. The mean, median, and mode are the most common measures of central tendency. The **mean** (symbolized as $M$) is the arithmetic average. It is what we report when talking about the class average on a test, for example. The **median** (symbolized as $Mdn$) is the value that half the observations (or scores) exceeded and half were below. It is the middle score in a distribution of scores arranged from lowest to highest. The median is often reported when a distribution of scores is not bell shaped (i.e., not a normal distribution). The **mode** (symbolized as $Mo$) is the most frequently occurring score or value in your data. The mode gives us a measure of the typical value in the distribution. For example, if you were making a “one-size-fits-all” pair of eyeglasses, you would want the mode for head size. Each measure of central tendency uses a different approach to describe the average of a group of scores.

The most common statistics used for describing variability in data are the range, variance, and standard deviation. The **range** either is reported as the highest and lowest score or is reduced to a single value that is the distance between these two scores. On an exam, you may ask what was the highest score attained, and perhaps out of morbid curiosity, you may want to know the lowest score as well. The range is an appropriate measure of variability for some types of data, but it is quite crude. For example, there may be one very high score and one very low score, and the range will not indicate that perhaps all the other scores were concentrated very near the mean. Two related measures of variability do provide this information. The **variance** and its square root, the **standard deviation** (symbolized as $SD$), provide a measure of the average distance scores are from the mean. Particularly in data that are bell shaped or normally distributed, the standard deviation tells us that about 2/3 of the scores fall between one standard deviation above the mean and one standard deviation below the mean. More detail on the calculation and appropriate selection of these statistics is given in Chapter 15 and Chapter 5.

Often you will read research articles that describe the degree that variables are related to one another. The most common measure of association is the **Pearson product-moment correlation** (symbolized as $r$). This statistic describes how strongly (or weakly) variables are related to one another. For example, if two variables are perfectly correlated, the $r$ value will be 1 or –1. The sign of the number indicates the direction of the relationship. A **positive correlation** tells us that the variables are directly related; as one variable increases, so does the other, and as one variable decreases, so does the other. A **negative correlation** tells us that the variables are inversely related. That is, as one variable increases, the other decreases, and
as one variable decreases, the other increases. The magnitude of \( r \) tells us how strongly the variables are related. A zero correlation tells us that the variables are not related at all; as the value increases to +1 or decreases to –1, the strength of the relationship increases. A correlation of 1 (either positive or negative) is called a **perfect correlation.** Be aware that perfect correlations never actually occur in the real world. If they do, it usually means that you have inadvertently measured the same variable twice and correlated the data. For example, you would likely get a correlation of 1 if you measured reaction time in seconds and also in minutes. It would be no surprise to find that the values are correlated because they are the same measure only in different scales. Here is another example: Suppose you measured mood with two scales. It is likely that the measures will correlate highly. Again, this only indicates that you have two measures of the same thing.

These descriptive statistics are used to summarize what was observed in the research. But the idea of a lot of research is to generalize the findings beyond just the observations or participants in the study. We ultimately want to say something about behavior in general, not just the behavior that occurred in the study. To make these generalizations, we need inferential statistics. Before leaping into a list of the various inferential statistics you will likely come across in the literature, we would like to review some of the basic concepts of inference.

**Inferential Statistics**

**Inferential statistics** are used to generalize the findings of a study to a whole population. An **inference** is a general statement based on limited data. Statistics are used to attach a probability estimate to that statement. For example, a typical weather forecast does not tell you that it will rain tomorrow afternoon. Instead, the report will indicate the probability of rain tomorrow. Indeed, the forecast here for tomorrow is a 60% chance of rain. The problem with making an inference is that we might be wrong. No one can predict the future, but based on good meteorological information, an expert is able to estimate the probability of rain tomorrow. Similarly, in research, we cannot make generalized statements about everyone when we only include a sample of the population in our study. What we do instead is attach a probability estimate to our statements.

When you read the results of research papers, the two most common uses of inferential statistics will be **hypothesis testing** and **confidence interval estimation.**

“Does wearing earplugs improve test performance?”

“Is exercise an effective treatment for depression?”

“Is there a relationship between hours of sleep and ability to concentrate?”

“Are married couples happier than single individuals?”

These are all examples of research hypotheses that could be tested using inferential tests of significance. What about the following?

“Does the general public have confidence in its nation’s leader?”

“How many hours of sleep do most adults get?”

“At what age do most people begin dating?”
These are all examples of research with a focus on describing attitudes and/or behavior of a population. This type of research, which is more common in sociology than psychology, uses confidence interval estimation instead of tests of significance.

The vast majority of psychological research involves testing a research hypothesis. So let’s first look at the types of tests of significance you will likely see in the literature and then look at confidence intervals.

**Common Tests of Significance.** Results will be referred to as either statistically significant or not statistically significant. What does this mean? In hypothesis testing research, a straw-person argument is set up where we assume that a null hypothesis is true, and then we use the data to disprove the null and thus support our research hypothesis. **Statistical significance** means that it is unlikely that the null hypothesis is true given the data that were collected. Nowhere in the research article will you see a statement of the null hypothesis, but instead you will see statements about how the research hypothesis was supported or not supported. These statements will look like this:

> “With an alpha of .01, those wearing earplugs performed statistically significantly better \((M = 35, SD = 1.32)\) than those who were not \((M = 27, SD = 1.55)\), \(t(84) = 16.83, p = .002\).”

> “The small difference in happiness between married \((M = 231, SD = 9.34)\) and single individuals \((M = 240, SD = 8.14)\) was not statistically significant, \(t(234) = 1.23, p = .21\).”

These statements appear in the results section and describe the means and standard deviations of the groups and then a statistical test of significance (\(t\) test in both examples). In both statements, statistical significance is indicated by the italic \(p\). This value is the \(p\) value. It is an estimate of the probability that the null hypothesis is true. Because the null hypothesis is the opposite of the research hypothesis, we want this value to be low. The accepted convention is a \(p\) value lower than .05 or, better still, lower than .01. The results will support the research hypothesis when the \(p\) value is lower than .05 or .01. The results will not support the research hypothesis when the \(p\) value is greater than .05. You may see a nonsignificant result reported as NS with no \(p\) value included.

You will find a refresher on statistical inference, including a discussion of Type I and Type II errors, and statistical power in Chapter 4.

Researchers using inferential techniques draw inferences based on the outcome of a statistical significance test. There are numerous tests of significance, each appropriate to a particular research question and the measures used, as you will recall from your introductory statistics course. It is beyond the scope of our book to describe in detail all or even most of these tests. You might want to refresh your memory by perusing your stats text, which of course you have kept, haven’t you? We offer a brief review of some of the most common tests of significance used by researchers in psychology in the section called “Basic Tests of Significance” at the end of the Discussion section.

Going back to the results section of our example paper, we see that the author has divided that section into a number of subsections. The first section, with the heading “Mood,” reports the effect of light on mood. It is only one sentence: “No significant results were obtained” (Knez, 2001, p. 204). The results section is typically brief, but the author
could have provided the group means and statistical tests that were not statistically signif-
ificant. The next subsection, titled “Perceived Room Light Evaluation,” provides a statistically
significant effect. Knez reports a significant (meaning statistically significant) gender
difference. He reports Wilk’s lambda, which is a statistic used in multivariate ANOVA
(MANOVA; when there is more than one DV), and the associated F statistic and p value for
the gender difference, $F(7, 96) = 3.21, p = .04$. He also includes a figure showing the mean
evaluations by men and women of the four light conditions and separate $F$ statistics and
$p$ values for each condition.

In the subsections that follow, Knez (2001) reports the results and statistical tests for the
effect of light condition on the various dependent variables. One of the effects he reports
as a “weak tendency to a significant main effect” (p. 204) with a $p$ value of .12. We would
simply say that it was not statistically significant, NS. Indeed, many of his statistical tests pro-
duced $p$ values greater than .05. We bring this to your attention as a reminder that even peer-
reviewed journal articles need to be read with a critical eye. Don’t just accept everything you
read. You need to pay attention to the $p$ values and question when alpha levels are not .05
or .01. You also need to examine the numbers carefully to discern the effect size.

What is noticeably missing from the results section of Knez (2001), our example paper,
is a calculation of effect size. Effect size gives us some indication of the strength of the
effect (see Chapter 4 for more detail). Remember, statistical significance tells us that an
effect was likely not due to chance and is probably a reliable effect. What statistical sig-
nificance does not indicate is how large the effect is. If we inspect the numbers in Knez’s
paper, we can see that the effects were not very large. For example, on the short-term
recall task, the best performance was from the participants in the warm lighting condi-
tions. They had a mean score of 6.9 compared to the other groups, with a mean score of
about 6.25. A difference of only 0.65 of a word on a recall task seems like a pretty small
effect, but then again, one would hardly expect that lighting conditions would have a dra-
matic effect on performance.

Once you have finished reading the introduction, method, and results sections, you should
have a pretty good idea about what was done, to whom, and what was found. In the discus-
sion section, you will read the researcher’s interpretation of the research, comments about
unexpected findings, and speculations about the importance of the work or its application.

The Discussion

The dissertation adviser of one of the authors of this book told her that he never read the dis-
cussion section of research reports. He was not interested in the interpretation of the
authors. He interpreted the findings and their importance himself. We consider this good
advice for seasoned researchers but not good advice for students. The discussion section of
a research paper is where the author describes how the results fit into the literature. This is
a discussion of the theories that are supported by the research and the theories that are not.
It is also where you will find suggestions from the author as to where the research should
go in the future—what questions are left unanswered and what new questions the research
raises. Indeed, the discussion section may direct you in your selection of research project.
You may wish to contact the author to see if research is already being conducted on the ques-
tions posed in the discussion. Remember that it is important to be a critical consumer of
research. Do not simply accept what is said in the discussion. Ask yourself if the results really do support the author's conclusions. Are there other possible interpretations?

In the discussion section of our example paper, Knez (2001) relates the findings to his previous work and the research of others. He discusses the lack of effect of light on mood and questions the mood measure that was used. We think that another possibility, which he does not explore, is that lighting may not have an influence on mood. He also describes the effect of light on cognitive performance as being something new to the literature. We could speculate that this small effect might not be a reliable finding. Certainly, the weak p values reported in the results section would indicate either that the study should be replicated or that the results were a fluke. Again, as we said before, you need to be critical when reading the literature.

Basic Tests of Significance

*t Test.* The simplest experiment involves two groups, an experimental and a control group. The researchers treat the groups differently (the IV) and measure their performance (the DV). The question then is, “Did the treatment work?” Are the groups significantly different after receiving the treatment? If the research involves comparing means from two groups, the *t* test may be the appropriate test of significance. Be aware that the *t* test can also be used in nonexperimental studies. For example, a researcher who compares the mean performance of women with that of men might use a *t* test.

Typically, a researcher will report the group means, whether the difference was statistically significant, and the *t* test results. In essence, the *t* test is an evaluation of the difference between two means relative to the variability in the data. Simply reporting the group means is not enough because a large difference between two means might not be statistically significant when examined relative to the large variability of the scores of each group. Alternatively, a small difference between two means may be statistically significant if there is very little variation in scores within each group. The *t* test is a good test when you want to compare two groups, but what if you have more than two groups?

*F* Test. The *F* test of significance is used to compare means of more than two groups. There are numerous experimental (and quasi-experimental) designs, known as ANOVAs, that are analyzed with the *F* test. Indeed, when we were graduate students, we took entire courses in ANOVA. In general, the *F* test, like the *t* test, compares between-group variability with within-group variability.

As with the *t* test, the researcher will report the group means and whether the differences were statistically significant. From a significant *F* test, the researcher knows that at least two means were significantly different. To specify which groups were different from which others, the researcher must follow the *F* test with post hoc (after-the-fact) comparisons. For example, if there were three groups and the *F* test was statistically significant, a post hoc test might find that all three group means were statistically significantly different or perhaps that only one mean differed from the other two. There are a large number of post hoc tests (e.g., Scheffé, Tukey, LSD, Bonferroni) that have slightly different applications. What is common to all these tests is that each produces a *p* value that is used to indicate which means differ from which.
As indicated above, many designs are analyzed with an $F$ test, and they have names that indicate the number of independent variables. You will find a one-way ANOVA used when there is one independent variable, a two-way ANOVA when there are two independent variables, and a three-way ANOVA (you guessed it) when there are three. A null hypothesis is tested for each independent variable by calculating an $F$ statistic. The advantage of the two- and three-way ANOVAs is that an interaction effect can also be tested. An interaction occurs when different combinations of the levels of the independent variables have different effects on the dependent variable. For example, if we wanted to investigate the effect of environmental noise (silent vs. noisy) on reading comprehension and the effect of different-colored paper (white, yellow, pink) on reading comprehension, we could use a two-way ANOVA to evaluate the effect of each independent variable and also whether the color of paper might interact with the noise to influence reading comprehension. It may be that noise produces a reduction in reading comprehension for white paper but not for yellow or pink paper. The interaction effect is important because it indicates that a variable is acting as a moderating variable. In this example, the effect of environmental noise on reading comprehension is moderated by the color of paper.

There is another type of ANOVA that is used to control for a possible confounding variable. This procedure also uses the $F$ statistic and is called analysis of covariance, or ANCOVA. Using our paper color example, suppose we want to test whether the color of paper will influence reading comprehension, but our participants vary considerably in age. This could pose a serious confound because reading comprehension changes with age. If we measured age, we can use ANCOVA to remove variability in reading comprehension that is due to age and then test the effect of color. The procedure removes the variance due to age from the dependent variable before the $F$ is calculated for the effect of color. Consequently, we are testing the effect of color after we have taken into account the effect of age.

The statistics described above are useful for comparing group means, but you may come across research where the variables are categories and the data are summarized by counting the frequency of things. When there are frequency counts instead of scores, you may see a chi-square test.

**Chi-Square Test.** Do people prefer Coke or Pepsi? Perhaps we have offered both drinks and asked people to declare a preference. We count the number of people preferring each drink. These data are not measures, and means cannot be calculated. If people’s preference did not differ between the two drinks, we would expect about the same number of people to pick each, and we could use a chi-square test, called the goodness-of-fit test, to test our hypothesis. In chi-square, our null hypothesis is that our observed frequencies will not be different from those we would expect by chance.

In the literature, you will likely see the data summarized by reporting the frequencies of each category either as total counts or perhaps as percentages of the total. Then you may read a statement that the frequencies in the groups are statistically significant followed by a report of the chi-square statistic and $p$ value.

Chi-square is called a nonparametric or distribution-free test because the test does not make the assumption that the population is distributed normally. Indeed, hypotheses
about the *shape* of the population distribution are exactly what we are testing with chi-square.

There are two common chi-square tests: the **goodness-of-fit test** and the **test for independence**. The goodness-of-fit test is used when there are categorical data on one variable as we had in the soft drink preference example. Perhaps a researcher is interested in the relationship between two **categorical variables**. In this case, you might see the chi-square test for independence. Imagine our researcher has asked cola tasters to indicate their choice of cola and has also categorized them by age. The research hypothesis might be that preference for cola depends on age. The researcher might think that younger people prefer Pepsi, for example, and older people prefer Coke. Or perhaps older people have no preference. The chi-square statistic is the same for this test as for the goodness-of-fit test. The difference is in the hypothesis. The null is that the two variables are independent (i.e., there is no relationship between them). In a research article, you will likely see a table of frequencies (or percentages), a statement as to whether a relationship was found between the variables, and the chi-square statistic and *p* value.

### Conceptual Exercise 2B

For each of the following, decide if a *t* test, an *F* test, or a chi-square test might be appropriate:

1. A new teacher decides to put some of the principles he learned in school to the test. He randomly selects half of his class and consistently praises each student for being on task for a minimum period of time. With the other half of the class, he periodically gives praise for on-task behavior. He wants to know if periodic praise produces more on-task behavior than consistent praise.

2. Psychiatric walk-in clients are randomly assigned to five therapists for short-term counseling. One therapist specializes in psychoanalytic techniques, one in client-centered techniques, one in behavioral techniques, and one in cognitive techniques. The fifth therapist is eclectic, using techniques from each of the above therapies. All clients are rated on various scales designed to measure improvement. Mean improvement ratings of the clients for each therapist are compared.

3. A statistics professor wants to know if generally there are more or less equal numbers of psychology, sociology, and business students. She keeps a tally.

### Other Nonparametric Tests.

In addition to chi-square, there are numerous other nonparametric tests that you will see in the literature. We have not tried to present a complete list here, but instead we have included the more common tests.

A nonparametric alternative to a *t* test for independent groups is the **Mann-Whitney U test**, which detects differences in central tendency and differences in the entire distributions.
of rank-ordered data. The **Wilcoxon signed-ranks test** is an alternative to a \( t \) test for dependent groups for rank-order data on the same or matched participants.

A nonparametric alternative to the one-way ANOVA is the **Kruskal-Wallis \( H \) test**, used when the data are rank orders of three or more independent groups. When those groups are dependent (i.e., repeated measures), a nonparametric test is the Friedman test.

**Pearson’s \( r \) Test.** If you earned a lot of money, would you be happy? Is there a relationship between income and happiness? If a researcher were interested in investigating a linear relationship between two continuous variables, he or she would use the Pearson product-moment test to calculate the correlation, \( r \). If you are getting a sense of déjà vu, it is probably because we talked about \( r \) as a descriptive statistic, but here we are talking about it as an inferential statistic. The important distinction is that the \( r \) reported as an inferential statistic will have an associated \( p \) value. For example, in a research article, you will read that a positive relationship was found between a measure of need for achievement and years of education and that the relationship was statistically significant. If the relationship was statistically significant, then you will also see a \( p \) value reported.

**Regression.** Regression is related to correlation, but in regression, we are interested in using a **predictor variable** to predict a **criterion variable**. Continuing with the example of need for achievement and education, perhaps the researcher was also interested in predicting need for achievement from education level. If the correlation between the two variables is statistically significant, then it is a simple matter of fitting a line through the data and using the equation for the line to predict need for achievement from education level. We say simple matter because the calculations are all done by computer, but certainly the equation for a line is simple.

\[
Y = mX + b,
\]

(equation for a straight line)

where \( Y \) is the criterion variable, \( X \) is the predictor variable, \( m \) is the slope of the line, and \( b \) is the value of \( Y \) where the line intercepts the \( Y \)-axis. Be sure to keep in mind as you read the research that the accuracy of the predicted values will be as good as the correlation is. That is, the closer the correlation is to +1 (or −1), the better the predictions will be.

**Important note to students:**

If you’re reading this material and starting to get anxious, relax! Our intention here is to discuss these inferential statistics at a conceptual level. As we indicated earlier, when you begin reading the literature, it is unlikely that you will see research using \( t \) tests or simple ANOVAs. What you will see are complex statistics that may be completely new to you. Our intention here is to give you enough information to understand what is being described in the literature.
Multiple Regression. If predicting someone’s performance using one predictor variable is a good idea, then using more than one predictor variable is a better idea. Entire textbooks are devoted to multiple regression techniques, but the basic idea is to use more than one predictor variable, \( X \), to predict one criterion variable, \( Y \). As with simple regression, multiple regression requires the fitting of a line through your data, but first, all the predictor variables are combined, and then the linear combination of \( X \)s is correlated with \( Y \). An \( r \) value reflects how well the linear combination of \( X \)s predicts \( Y \). Some predictor variables are likely to be better predictors of \( Y \) than others, and the analysis produces weights that indicate how well each predictor variable predicts. These values can be used in a regression equation to predict \( Y \). Simply multiply the values of the predictor variables by their respective weights and you have your predicted value.

\[
Y(\text{predicted}) = B_1(X_1) + B_2(X_2) + B_3(X_3) + \text{etc.} + \text{Constant.}
\]

In addition to the weights used to predict criterion values, multiple regression analysis also provides standardized weights called beta (\( \beta \)) weights. These values tell us something about each individual predictor in the regression analysis. They can be interpreted much like an \( r \) value, with the sign indicating the relationship between the predictor variable and the criterion and the magnitude indicating the relative importance of the variable in predicting the criterion. Thus, in a multiple regression analysis, we can examine the relative contribution of each predictor variable in the overall analysis.

As you just learned, multiple regression is used to determine the influence of several predictor variables on a single criterion variable. Let’s look briefly at two useful concepts in multiple regression: partial and semipartial (also called part) correlation.

Partial Correlation. Sometimes we would like to measure the relationship between two variables when a third has an influence on them both. We can partial out the effects of that third variable by computing a partial correlation. Suppose there is a correlation between age and income. It seems reasonable that older people might make more money than younger people. Is there another variable that you think might be related to age and income? How about years of education? Older people are more likely to be better educated, having had more years to go to school, and it seems likely that better educated people earn more. So, what is the true relationship between age and income if the variable, years of education, is taken out of the equation? One solution would be to group people by years of education and then conduct a number of separate correlations between age and income for each education group. Partial correlation, however, provides a better solution by telling us what the true relationship is between age and income when years of education has been partialled out.

Semipartial Correlation. As we just discussed, in partial correlation, we remove the relationship of one variable from the other variables and then calculate the correlation. But what if we want to remove the influence of a variable from only one of the other variables? This is called a semipartial correlation. For example, at our school, we accept senior students into our applied psychology program based on their grades in the first and second years. We have found a strong positive correlation between previous grades and performance in our program. Suppose we could also administer an entrance exam to use as
another predictor, but the exam was expensive. We can use semipartial correlation to determine how much the entrance test will increase our predictive power over and above using previous grades.

How do we do this? Well, we correlate entrance test scores and performance in the program after first removing the influence of previous grades on program performance.

This correlation value then will tell us what relationship remains between entrance test and program performance when the correlation with previous grades has been partialed out of program performance but not entrance test scores. In our example, we could decide, based on this correlation, whether an expensive entrance test helped our predictive ability enough for us to go ahead and use it.

**Logistic Regression.** Suppose you were interested in predicting whether a young offender would reoffend. You measure a number of possible predictor variables, such as degree of social support, integration in the community, job history, and so on, and then follow your participants for 5 years and measure if they reoffend. The predictor variables may be continuous, but the criterion variable is discrete; they reoffend or they don’t. When we have a discrete criterion variable, we use **logistic regression.** Just as we used a combination of the predictor variables to predict the criterion variable in multiple regression, we do the same thing in logistic regression. The difference is that instead of predicting a value for the criterion variable, we predict the likelihood of the occurrence of the criterion variable. We express this as an **odds ratio**—that is, the odds of reoffense divided by the odds of not reoffending. If the probability of reoffending is .75 (or 75% chance of reoffending), then the probability of not reoffending is .25 (1 – .75). The odds of reoffending is .75/.25 or 3 to 1, and the odds of not reoffending is .25/.75 or .33. We calculate the odds ratio of reoffending versus not reoffending as .75/.33 or 2.25. In other words, the odds of reoffending is two and quarter times higher than not reoffending.

**Factor Analysis.** **Factor analysis** is a correlational technique we use to find simpler patterns of relationships among many variables. Factor analysis can tell us if a large number of variables can be explained by a much smaller number of uncorrelated constructs or factors.

About a hundred years ago, Charles Spearman (1904) thought that measures of mental ability such as mathematical skill, artistic talent, reasoning, language ability, and so on could all be explained by one underlying variable or factor, which he called \( g \) for general intelligence. He thought that general intelligence was the single common factor to all the various tests of mental ability. Although other theorists disagree with his notion of \( g \), his idea demonstrates what factor analysis is all about.

In factor analysis, the researcher is looking for underlying and independent factors that have not been directly measured to explain a lot of variables that have been measured. The procedure involves identifying the variables that are interrelated. Once the factors have been identified, it is up to the researcher to decide what construct this group of variables is measuring. These underlying factors are hypothetical, that is, inferred by the researcher. The researcher attempts to find the smallest number of factors that can adequately explain the observed variables and to determine the fundamental nature of those factors.

When you read a research report where factor analysis has been used, you will probably see a complicated-looking matrix called a correlation matrix. Don’t be discouraged.
Keep in mind that although the mathematics are complex and beyond the scope of this book, the concept is reasonably simple. Can an underlying variable such as general intelligence explain a whole lot of variation in measures of mental abilities?

**Cluster Analysis.** Cluster analysis includes a range of algorithms and methods used to group similar objects into categories, or clusters. The members of each cluster are thus more similar to each other than they are to members of other clusters. Unlike factor analysis, where the goal is to group similar variables together, in cluster analysis, the idea is to group similar members. Organizing data into meaningful structures or taxonomies is a task many researchers face. Cluster analysis is a method that can discover structure in data, but it does not in and of itself have any explanatory function. In other words, the analysis can find structure but does not explain it.

Imagine a hospital where patients are assigned to wards based on similar symptoms or perhaps similar treatments. Each ward could be considered a cluster. A cluster analysis might discover the similarities among the patients in each ward, and the researcher then has the job of determining why the cluster or ward is similar (i.e., symptoms, treatment, age, etc.).

Cluster analysis is often used when researchers have no a priori hypotheses and are in the beginning phase of their research. As such, statistical significance testing often has no role in such analyses.

**Structural Equation Modeling.** Structural equation modeling is a complex endeavor that can involve various techniques, including factor analysis, regression models, path analysis, and so on. We will just be able to give you an idea of the purpose of structural equation modeling here.

We hope you will remember what happens when we transform a set of numbers by adding a constant to each or multiplying each by a constant. Let’s say we multiply all the numbers in a list by a constant \( c \). The mean of that set of transformed numbers will be equal to the old mean times \( c \), the standard deviation of the new set of numbers will equal the old standard deviation times the absolute value (i.e., ignore the sign) of \( c \), and the variance will equal the old variance times \( c \) squared. Simple, right?

What is our point, you might be wondering? Well, bear with us. If we suspected that two sets of numbers were related, we could compare the variances of the two sets of numbers, for example, to confirm our suspicions. If one set was related to the other set by the equation \( Y = 2X \), then the variance of \( Y \) must be 4 times the variance of \( X \). So we could confirm our hypothesis about the relationship between the two sets of numbers by comparing their variances rather than the numbers themselves. We hope you are not too confused by this somewhat odd way of doing things, but we think it might help you understand structural equation modeling. Two sets of numbers could be related in much more mathematically complex ways than by \( Y = 2X \), but we hope you are getting the idea. You can determine if variables are related by looking at their variances and covariances.

Structural modeling is a way of determining whether a set of variances and covariances fits a specific structural model. In essence, the researcher hypothesizes that the variables are related in a particular way, often with something called a path diagram that shows the interrelationships. Then the researcher figures out what this model predicts about the values of the variances and covariances of the variables. This is the really complex part of the
process, and we just can’t go there! Then the researcher examines the variances and covari-
ces of the variables to see if they fit the predictions of the model.

As we said earlier, this is a complex procedure well beyond the scope of our book, but we hope our brief discussion gives you some idea of the purpose of structural equation modeling.

**Discriminant Function Analysis.** As we mentioned earlier, at our school, we offer an applied psychology degree program. One of our objectives is to prepare students for graduate work in applied areas. Imagine that we classified our graduates over the past 10 years into two groups: students who were accepted into graduate school and students who were not. We could use discriminant function analysis to predict acceptance into graduate school using grade point average (GPA) and workshop attendance, for example. Our analysis might help us determine how GPA and workshop attendance individually predict acceptance into graduate school and how a combination of both predicts acceptance.

This is the idea behind discriminant function analysis. Of course, we might have many more variables, and the analysis allows us to determine the predictive ability of each variable alone and in combination with other variables. If discriminant function analysis sounds like logistic regression, it is because they are related. They have similar applications, but discriminant function analysis is calculated as ANOVA with more than one dependent variable (MANOVA). The various dependent variables are used to predict group membership.

This analysis, like the others discussed in this section, is much more complex than this, but again we hope our brief discussion gives you an inkling about the use of these techniques so that when you read the literature, you will have some understanding about the research outcomes.

The statistical procedures we have been discussing in the last sections all involve an a priori hypotheses about the nature of the population. Hypothesis testing is used a lot in psychology. Some other disciplines tend to prefer post hoc procedures, and you will find confidence interval estimates quite often in the literature you will be reading.

**Confidence Intervals**

Confidence intervals are used when we are interested in estimating population parameters. We are still making an inference from a sample to a population, and because of that, we are using probability estimates. But instead of reporting a p value indicating the probability that the null is true, we report the probability that our estimate about the population is true. Pollsters describing political candidates often use confidence intervals. For example, you may have read reports that, based on a poll of 1,000 respondents, 83% say they would vote for X if there were an election tomorrow. These statements are typically followed with a statement such as, “These results are accurate to within 3 percentage points 19 times out of 20.” What does this mean? It means that, based on a sample of 1,000, the population support for the candidate is probably somewhere between 83% - 3% and 83% + 3% or somewhere between 80% and 86%. Are they absolutely sure? No, they say the estimate should be correct 19 times out of 20 or 95% of the time (19/20 = .95). So, a p value of .05 from hypothesis testing becomes a confidence interval of .95, and similarly, a p value of .01 becomes a confidence interval of .99 (reported as 99 times out of 100). Again, in hypothesis testing, we
report a significance test with a $p$ value that indicates the probability that the null is true. In confidence intervals, we report an interval within which we estimate the true population parameter to fall.

We hope that this chapter has prepared you, on a conceptual level, to understand the literature you will be reading as you continue in your social science studies. We now turn to a topic that is so important in social science research, we have devoted an entire chapter to it—research ethics.

**CHAPTER SUMMARY**

Once a general research topic has been selected, a literature search is necessary to determine what research has already been conducted in the area. Various databases are available for the psychology literature. One of the most useful is PsycINFO. Review articles, books, chapters in books, edited volumes, and chapters in edited volumes are also found in the research databases. Peer-reviewed journals are the best sources of original research. Searching the literature for relevant research will be more successful if appropriate search terms are used.

Original research journal articles generally include an abstract, an introduction, a method section, a results section, and a discussion section. The purpose of the abstract is to summarize the article. There should be enough information in the abstract for the reader to decide if he or she should read the entire research article. In the introduction, there will be a description of the relevant research and a description of the specific research hypotheses of the author(s). The independent and dependent variables are often described in the introduction, as well.

The method section is typically divided into subsections such as Participants or Subjects, Materials, and Apparatus. The method section always contains a subsection called Procedure. Enough details of the procedure must be included so that researchers elsewhere could replicate the research.

In the results section, the statistical data are presented. Both descriptive and inferential statistics will be reported. Descriptive measures of central tendency, variability, and the strength of the relationship between variables will be reported. Typically, the inferential statistics follow the descriptive statistics. A lot of psychology research involves testing hypotheses. Any tests of significance that were used to assess the research hypotheses will be reported in the results section. Basic tests of significance include $t$ tests, $F$ tests, chi-square tests, correlation and regression tests, and so on. The authors will indicate whether or not the hypotheses they put forth in the introduction were supported by the statistical analyses.

More complex analyses that are common in the research literature include multiple regression, partial correlation, semipartial correlation, logistic regression, factor analysis, cluster analysis, structural equation modeling, and discriminant function analysis.

Although hypothesis testing is more common in the psychology research, confidence interval estimation is also used. A confidence interval is a range of values with a known probability of containing a parameter.

The discussion section of a research article contains the authors’ interpretation of the statistical findings and suggestions about future research directions.
ANSWERS TO CONCEPTUAL EXERCISES

Conceptual Exercise 2A

1a. IV is amount of practice; DV is reaction time.
1b. IV is amount of exercise; DV is ratings of depression.

Conceptual Exercise 2B

1. Because there are two groups, a $t$ test might be appropriate.
2. There are five groups, and so an $F$ test might be appropriate.
3. A chi-square goodness-of-fit test would answer this question.

FAQ

Q1: What’s a dependent variable dependent on?
A1: We hope it is dependent on our manipulation (the independent variable).

Q2: I have taken an intro stats course and I can’t make heads or tails out of the research I am reading.
A2: We understand. The statistics used by most researchers today go well beyond what you learned in your intro stats course. You will need a graduate-level course under your belt to understand a lot of the statistics you will read about, but we hope you will be able to understand on a conceptual level a lot of what you read.

Q3: I just read a research article and they talked about a bunch of correlation stuff and validity and reliability. I have no clue.
A3: Go to Chapter 5 and Chapter 4.

Q4: I just read a paper that talked about stratified sampling. ????
A4: Go to Chapter 6.

Q5: My prof tells me to use APA style in my report, but the articles I have read don’t look anything like the APA manual.
A5: Yes, each specific journal uses its own style. Go to Chapter 14.

CHAPTER EXERCISES

1. Identify the IV and DV for each of the following. Indicate any participant variables being examined.
   a. Does the use of imagery enhance athletic performance?
   b. Are teens more concerned about their bodies than older adults?
c. Does repetition in advertising improve sales?
d. Does straight alley training improve the speed of rats in a maze?
e. Is there a difference in leadership style between men and women?

2. Why are peer-reviewed journals preferable? What does the term blind review mean?

3. List two things found in the introduction of a research article.

4. List and describe what is found in typical subsections of the method section of a research article.

5. List two kinds of statistics always found in the results section of a research article.

6. What is the purpose of the discussion section of a research article?

7. Describe the difference between a mediating and a moderating variable.

8. What is the general purpose of a significance test?

9. What is the general purpose of confidence interval estimation?

**CHAPTER PROJECTS**

1. Locate three research articles from peer-reviewed journals. Briefly summarize the article and describe the IV (or participant/subject variables), DV, control procedures employed (and why they were needed), and the descriptive statistics used. Discuss other ways that the variables could have been operationalized.

2. With a search term of your choice, find three empirical research articles. Describe the methods the researchers used to increase the power of their analysis. Can you think of other ways to increase power in each study?

**REFERENCES**


Visit the study site at www.sagepub.com/evansmprstudy for practice quizzes and other study resources.