CENGAGE

Tenth Edition

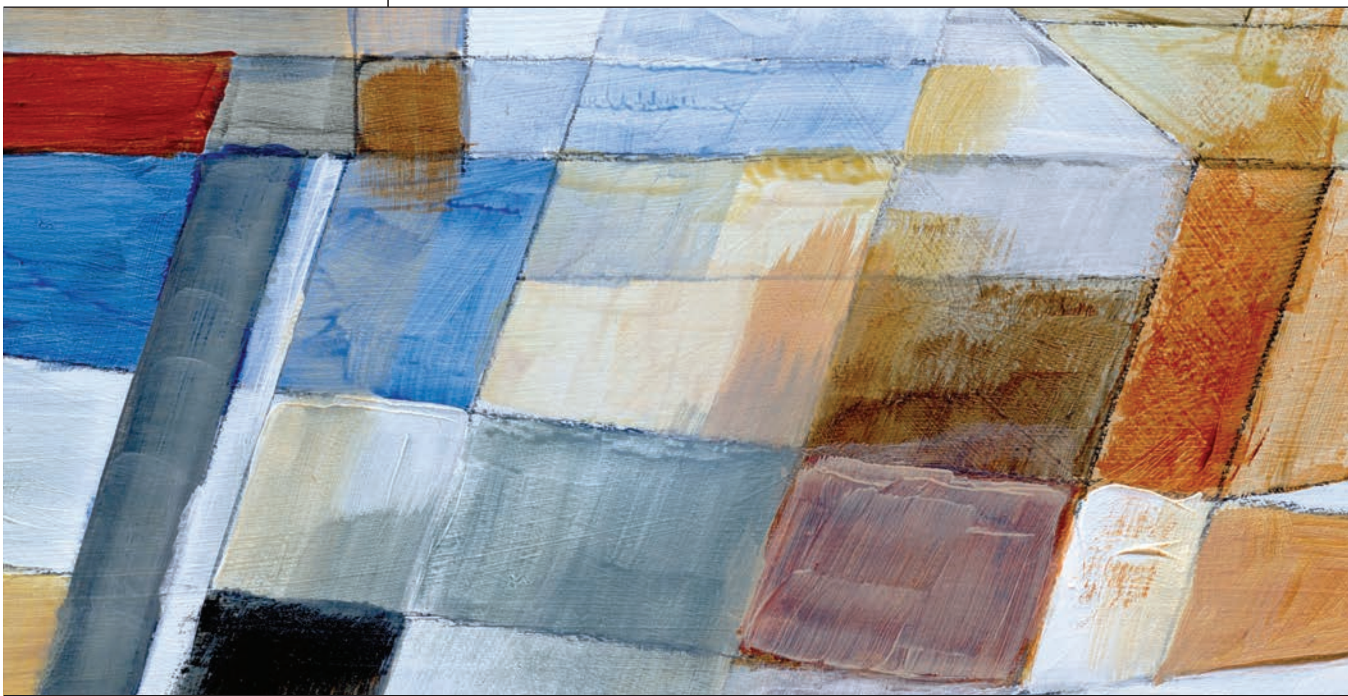
Essentials of Statistics for the Behavioral Sciences

Frederick J Gravetter | Larry B. Wallnau
Lori-Ann B. Forzano | James E. Witnauer

EDITION

10

Essentials of Statistics FOR THE Behavioral Sciences



clivewa/Shutterstock.com

FREDERICK J GRAVETTER
Late of The College at Brockport, State University of New York

LARRY B. WALLNAU
The College at Brockport, State University of New York

LORI-ANN B. FORZANO
The College at Brockport, State University of New York

JAMES E. WITNAUER
The College at Brockport, State University of New York

CENGAGE

Australia • Brazil • Mexico • Singapore • United Kingdom • United States

CENGAGE

Essentials of Statistics for the Behavioral Sciences, 10th Edition
Frederick J Gravetter, Larry B. Wallnau, Lori-Ann B. Forzano, James E. Witnauer

Senior Vice President, Higher Education & Skills Product: Erin Joyner

Product Director: Laura Ross

Product Manager: Josh Parrott

Content Manager: Tangelique Williams-Grayer, Brian Pierce

Product Assistant: Kat Wallace

Marketing Manager: Tricia Salata

Intellectual Property Analyst: Deanna Ettinger

Intellectual Property Project Manager: Nick Barrows

Production Service: Lori Hazzard, MPS Limited

Art Director: Bethany Bourgeois

Text Designer: Liz Harasymczuk

Cover Designer: Cheryl Carrington

Cover Image: clivewa/Shutterstock.com

Library of Congress Control Number: 2019911268

ISBN: 978-0-357-36529-8

Cengage
200 Pier 4 Boulevard
Boston, MA 02210
USA

Cengage is a leading provider of customized learning solutions with employees residing in nearly 40 different countries and sales in more than 125 countries around the world. Find your local representative at **www.cengage.com.**

Cengage products are represented in Canada by Nelson Education, Ltd.

To learn more about Cengage platforms and services, register or access your online learning solution, or purchase materials for your course, visit **www.cengage.com.**

Printed in the United States of America
Print Number: 01 Print Year: 2019

BRIEF CONTENTS

CONTENTS

CHAPTER 1 Introduction to Statistics 1


clivewa/Shutterstock.com

CHAPTER 2 Frequency Distributions 43


clivewa/Shutterstock.com

CHAPTER 3 Central Tendency 73


clivewa/Shutterstock.com

CHAPTER 8 Introduction to Hypothesis Testing 243


clivewa/Shutterstock.com

CHAPTER 9 Introduction to the *t* Statistic 291


clivewa/Shutterstock.com

CHAPTER 10 The *t* Test for Two Independent Samples 323

clivewa/Shutterstock.com

CHAPTER 11 The *t* Test for Two Related Samples 359

clivewa/Shutterstock.com

CHAPTER 15 The Chi-Square Statistic: Tests for Goodness of Fit and Independence 533

clivewa/Shutterstock.com

APPENDIXES

PREFACE

Many students in the behavioral sciences view the required statistics course as an intimidating obstacle that has been placed in the middle of an otherwise interesting curriculum. They want to learn about psychology and human behavior—not about math and science. As a result, the statistics course is seen as irrelevant to their education and career goals. However, as long as psychology and the behavioral sciences in general are founded in science, knowledge of statistics will be necessary. Statistical procedures provide researchers with objective and systematic methods for describing and interpreting their research results. Scientific research is the system that we use to gather information, and statistics are the tools that we use to distill the information into sensible and justified conclusions. The goal of this book is not only to teach the methods of statistics, but also to convey the basic principles of objectivity and logic that are essential for the behavioral sciences and valuable for decision making in everyday life.

Essentials of Statistics for the Behavioral Sciences, *Tenth Edition*, is intended for an undergraduate statistics course in psychology or any of the related behavioral sciences. The overall learning objectives of this book include the following, which correspond to some of the learning goals identified by the American Psychological Association (Noland and the Society for the Teaching of Psychology Statistical Literacy Taskforce, 2012).

1. Calculate and interpret the meaning of basic measures of central tendency and variability.
2. Distinguish between causal and correlational relationships.
3. Interpret data displayed as statistics, graphs, and tables.
4. Select and implement an appropriate statistical analysis for a given research design, problem, or hypothesis.
5. Identify the correct strategy for data analysis and interpretation when testing hypotheses.
6. Select, apply, and interpret appropriate descriptive and inferential statistics.
7. Produce and interpret reports of statistical analyses using APA style.
8. Distinguish between statistically significant and chance findings in data.
9. Calculate and interpret the meaning of basic tests of statistical significance.
10. Calculate and interpret the meaning of confidence intervals.
11. Calculate and interpret the meaning of basic measures of effect size statistics.
12. Recognize when a statistically significant result may also have practical significance.

The book chapters are organized in the sequence that we use for our own statistics courses. We begin with descriptive statistics (Chapters 1–4), then lay the foundation for inferential statistics (Chapters 5–8), and then we examine a variety of statistical procedures focused on sample means and variance (Chapters 9–13) before moving on to correlational methods and nonparametric statistics (Chapters 14 and 15). Information about modifying this sequence is presented in the "To the Instructor" section for individuals who prefer a different organization. Each chapter contains numerous examples (many based on actual research studies), learning objectives and learning checks for each section, a summary and list of key terms, instructions for using SPSS®, detailed problem-solving tips and demonstrations, and a set of end-of-chapter problems.

Those of you who are familiar with previous editions of *Statistics for the Behavioral Sciences* and *Essentials of Statistics for the Behavioral Sciences* will notice that some changes have been made. These changes are summarized in the "To the Instructor" section. Students who are using this edition should read the section of the preface entitled "To the Student." In revising this text, our students have been foremost in our minds. Over the years, they have provided honest and useful feedback, and their hard work and perseverance has made our writing and teaching most rewarding. We sincerely thank them.

To the Instructor

Previous users of any of the Gravetter-franchise textbooks should know that we have maintained all the hallmark features of our *Statistics* and *Essentials of Statistics* textbooks: the organization of chapters and content within chapters; the student-friendly, conversational tone; and the variety of pedagogical aids, including, Tools You Will Need, chapter outlines, and section-by-section Learning Objectives and Learning Checks, as well as end-of-chapter Summaries, Key Terms lists, Focus on Problem Solving tips, Demonstrations of problems solved, SPSS sections, and end-of-chapter Problems (with solutions to odd-numbered problems provided to students in Appendix C).

New to This Edition

Those of you familiar with the previous edition of *Statistics for the Behavioral Sciences* will be pleased to see that *Essentials of Statistics for the Behavioral Sciences* has the same "look and feel" and includes much of its content. For those of you familiar with *Essentials*, the following are highlights of the changes that have been made:

- Every chapter begins with a Preview, which highlights an example of a published study. These have been selected for level of interest so that they will draw the student in. The studies are used to illustrate the purpose and rationale of the statistical procedure presented in the chapter.
- There has been extensive revision of the end-of-chapter Problems. Many old problems have been replaced with new examples that cite research studies. As an enhanced instructional resource for students, the odd-numbered solutions in Appendix C now show the work for intermediate answers for problems that require more than one step. The even-numbered solutions are available online in the instructor's resources.
- The sections on research design and methods in Chapter 1 have been revised to be consistent with Gravetter and Forzano, *Research Methods for the Behavioral Sciences, Sixth Edition.* The interval and ratio scales discussion in Chapter 1 has been refined and includes a new table distinguishing scales of measurement.
- In Chapter 2, a new section on stem and leaf displays describes this exploratory data analysis as a simple alternative to a frequency distribution table or graph. A basic presentation of percentiles and percentile ranks has been added to the coverage of frequency distribution tables in Chapter 2. The topic is revisited in Chapter 6 (Section 6-4, Percentiles and Percentile Ranks), showing how percentiles and percentile ranks can be determined with normal distributions.

- Chapter 3 (Central Tendency) has added coverage for the median when there are tied scores in the middle of the distribution. It includes a formula for determining the median with interpolation.
- The coverage of degrees of freedom in Chapter 4 (Variability) has been revised, including a new box feature (Degrees of Freedom, Cafeteria-Style) that provides an analogy for the student. Rounding and rounding rules are discussed in a new paragraph in Section 4-2, Defining Variance and Standard Deviation. It was presented in this section because Example 4.2 is the first instance where the answer is an irrational number. A section on quartiles and the interquartile range has been added.
- Coverage of the distribution of sample means (Chapter 7) has been revised to provide more clarification. The topic is revisited in Chapter 9, where the distribution of sample means is more concretely compared and contrasted with the distribution of *z*-scores, along with a comparison between the unit normal table and the *t* distribution table. Chapter 7 also includes a new box feature that depicts the law of large numbers using an illustration of online shopping (The Law of Large Numbers and Online Shopping).
- In Chapter 8 (Introduction to Hypothesis Testing), the section on statistical power has been completely rewritten. It is now organized and simplified into steps that the student can follow. The figures for this section have been improved as well.
- A new box feature has been added to Chapter 10 demonstrating how the *t* statistic for an independent-measures study can be calculated from sample means, standard deviations, and sample sizes in a published research paper. There is an added section describing the role of individual differences in the size of standard error.
- The comparison of independent- and repeated-measures designs has been expanded in Chapter 11, and includes the issue of power.
- In Chapter 12 the section describing the numerator and denominator in the *F*-ratio has been expanded to include a description of the sources of the random and unsystematic differences.
- Chapter 13 now covers only the two-factor, independent-measures ANOVA. The single-factor, repeated-measures ANOVA was dropped because repeated-measures designs are typically performed in a mixed design that also includes one (or more) between-subject factors. As a result, Chapter 13 now has expanded coverage of the two-factor, independent-measures ANOVA.
- For Chapter 14, three graphs have been redrawn to correct minor inaccuracies and improve clarity. As with other chapters, there is a new SPSS section with figures and end-of-chapter Problems have been updated with current research examples.
- Chapter 15 has minor revisions and an updated SPSS section with four figures. As with other chapters, the end-of-chapter Problems have been extensively revised and contain current research examples.
- Many research examples have been updated with an eye toward selecting examples that are of particular interest to college students and that cut across the domain of the behavioral sciences.
- Learning Checks have been revised.
- All SPSS sections have been revised using SPSS® 25 and new examples. New screenshots of analyses are presented. Appendix D, General Instructions for Using SPSS®, has been significantly expanded.
- A summary of statistics formulas has been added.

- This edition of *Essentials of Statistics for the Behavioral Sciences* has been edited to align with Gravetter and Forzano, *Research Methods*, providing a more seamless transition from statistics to research methods in its organization and terminology. Taken together, the two books provide a smooth transition for a two-semester sequence of Statistics and Methods, or, even an integrated Statistics/Methods course.

Matching the Text to Your Syllabus

The book chapters are organized in the sequence that we use for our own statistics courses. However, different instructors may prefer different organizations and probably will choose to omit or deemphasize specific topics. We have tried to make separate chapters, and even sections of chapters, completely self-contained, so that they can be deleted or reorganized to fit the syllabus for nearly any instructor. Instructors using MindTap® can easily control the inclusion and sequencing of chapters to match their syllabus exactly. Following are some common examples:

- It is common for instructors to choose between emphasizing analysis of variance (Chapters 12 and 13) or emphasizing correlation/regression (Chapter 14). It is rare for a one-semester course to complete coverage of both topics.
- Although we choose to complete all the hypothesis tests for means and mean differences before introducing correlation (Chapter 14), many instructors prefer to place correlation much earlier in the sequence of course topics. To accommodate this, Sections 14-1, 14-2, and 14-3 present the calculation and interpretation of the Pearson correlation and can be introduced immediately following Chapter 4 (Variability). Other sections of Chapter 14 refer to hypothesis testing and should be delayed until the process of hypothesis testing (Chapter 8) has been introduced.
- It is also possible for instructors to present the chi-square tests (Chapter 15) much earlier in the sequence of course topics. Chapter 15, which presents hypothesis tests for proportions, can be presented immediately after Chapter 8, which introduces the process of hypothesis testing. If this is done, we also recommend that the Pearson correlation (Sections 14-1, 14-2, and 14-3) be presented early to provide a foundation for the chi-square test for independence.

To the Student

A primary goal of this book is to make the task of learning statistics as easy and painless as possible. Among other things, you will notice that the book provides you with a number of opportunities to practice the techniques you will be learning in the form of Examples, Learning Checks, Demonstrations, and end-of-chapter Problems. We encourage you to take advantage of these opportunities. Read the text rather than just memorizing the formulas. We have taken care to present each statistical procedure in a conceptual context that explains why the procedure was developed and when it should be used. If you read this material and gain an understanding of the basic concepts underlying a statistical formula, you will find that learning the formula and how to use it will be much easier. In the "Study Hints" that follow, we provide advice that we give our own students. Ask your instructor for advice as well; we are sure that other instructors will have ideas of their own.

Study Hints

You may find some of these tips helpful, as our own students have reported.

- The key to success in a statistics course is to keep up with the material. Each new topic builds on previous topics. If you have learned the previous material, then the new topic is just one small step forward. Without the proper background, however, the new topic can be a complete mystery. If you find that you are falling behind, get help immediately.

- You will learn (and remember) much more if you study for short periods several times a week rather than try to condense all of your studying into one long session. Distributed practice is best for learning. For example, it is far more effective to study and do problems for half an hour every night than to have a single three-and-a-half-hour study session once a week. We cannot even work on *writing* this book without frequent rest breaks.
- Do some work before class. Stay a little bit ahead of the instructor by reading the appropriate sections before they are presented in class. Although you may not fully understand what you read, you will have a general idea of the topic, which will make the lecture easier to follow. Also, you can identify material that is particularly confusing and then be sure the topic is clarified in class.
- Pay attention and think during class. Although this advice seems obvious, often it is not practiced. Many students spend so much time trying to write down every example presented or every word spoken by the instructor that they do not actually understand and process what is being said. Check with your instructor—there may not be a need to copy every example presented in class, especially if there are many examples like it in the text. Sometimes, we tell our students to put their pens and pencils down for a moment and just listen.
- Test yourself regularly. Do not wait until the end of the chapter or the end of the week to check your knowledge. As you are reading the textbook, stop and do the examples. Also, stop and do the Learning Checks at the end of each section. After each lecture, work on solving some of the end-of-chapter Problems and check your work for odd-numbered problems in Appendix C . Review the Demonstration problems, and be sure you can define the Key Terms. If you are having trouble, get your questions answered *immediately*—reread the section, go to your instructor, or ask questions in class. By doing so, you will be able to move ahead to new material.
- Do not kid yourself! Avoid denial. Many students observe their instructor solving problems in class and think to themselves, "This looks easy, I understand it." Do you really understand it? Can you really do the problem on your own without having to read through the pages of a chapter? Although there is nothing wrong with using examples in the text as models for solving problems, you should try working a problem with your book closed to test your level of mastery.
- We realize that many students are embarrassed to ask for help. It is our biggest challenge as instructors. You must find a way to overcome this aversion. Perhaps contacting the instructor directly would be a good starting point, if asking questions in class is too anxiety-provoking. You could be pleasantly surprised to find that your instructor does not yell, scold, or bite! Also, your instructor might know of another student who can offer assistance. Peer tutoring can be very helpful.

Contact Us

Over the years, the students in our classes and other students using our book have given us valuable feedback. If you have any suggestions or comments about this book, you can write to Professor Emeritus Larry Wallnau, Professor Lori-Ann Forzano, or Associate Professor James Witnauer at the Department of Psychology, The College at Brockport, SUNY, 350 New Campus Drive, Brockport, New York 14420. You can also contact us directly at: lforzano@brockport.edu or jwitnaue@brockport.edu or lwallnau@brockport.edu.

Ancillaries

Ancillaries for this edition include the following.

- **MindTap® Psychology** *MindTap® Psychology for Gravetter/Wallnau/Forzano/Witnauer's Essentials of Statistics for the Behavioral Sciences, Tenth Edition,* is the digital learning solution that helps instructors engage and transform today's students into critical thinkers. Through paths of dynamic assignments and applications that you can personalize, real-time course analytics, and an accessible reader, MindTap helps you turn cookie cutter into cutting edge, apathy into engagement, and memorizers into higher-level thinkers. As an instructor using MindTap, you have at your fingertips the right content and unique set of tools curated specifically for your course, such as video tutorials that walk students through various concepts and interactive problem tutorials that provide students opportunities to practice what they have learned, all in an interface designed to improve workflow and save time when planning lessons and course structure. The control to build and personalize your course is all yours, focusing on the most relevant material while also lowering costs for your students. Stay connected and informed in your course through real-time student tracking that provides the opportunity to adjust the course as needed based on analytics of interactivity in the course.
- **Online Instructor's Manual** The manual includes learning objectives, key terms, a detailed chapter outline, a chapter summary, lesson plans, discussion topics, student activities, "What If" scenarios, media tools, a sample syllabus, and an expanded test bank. The learning objectives are correlated with the discussion topics, student activities, and media tools.
- **Online PowerPoints** Helping you make your lectures more engaging while effectively reaching your visually oriented students, these handy Microsoft PowerPoint® slides outline the chapters of the main text in a classroom-ready presentation. The PowerPoint slides are updated to reflect the content and organization of the new edition of the text.
- **Cengage Learning Testing, powered by Cognero®** Cengage Learning Testing, powered by Cognero®, is a flexible online system that allows you to author, edit, and manage test bank content. You can create multiple test versions in an instant and deliver tests from your LMS in your classroom.

Acknowledgments

It takes a lot of good, hard-working people to produce a book. Our friends at Cengage have made enormous contributions to this textbook. We thank: Laura Ross, Product Director; Josh Parrott, Product Manager; Kat Wallace, Product Assistant; and Bethany Bourgeois, Art Director. Special thanks go to Brian Pierce and Tangelique Williams-Grayer, our Content Managers, and to Lori Hazzard, who led us through production at MPS Limited.

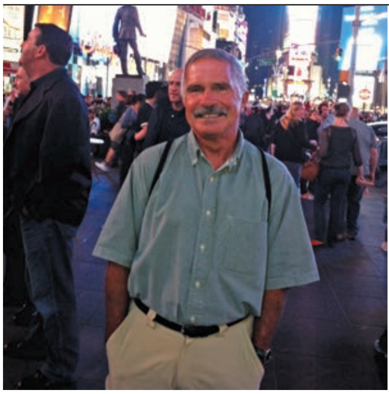
Reviewers play an important role in the development of a manuscript. Accordingly, we offer our appreciation to the following colleagues for their assistance: Kara Moore, Knox College; Tom Williams, Mississippi College; Stacey Todaro, Adrian College; Dave Matz, Augsburg University; Barbara Friesth, Indiana University-Purdue University Indianapolis; Bethany Jurs, Transylvania University; Ben Denkinger, Augsburg University; Sara Festini, University of Tampa; Lindsey Johnson, University of Southern Mississippi; Lawrence Preiser, York College CUNY; Stephen Blessing, University of Tampa; Pamela A MacLaughlin, Indiana University.

We must give our heartfelt thanks to our families: Naomi and Nico Wallnau; Charlie, Ryan, and Alex Forzano; and Beth, JJ, Nate, and Ben Witnauer. This book could not have been written without their patience and support.

Finally, it is with great sorrow that we acknowledge Fred Gravetter's passing. His expertise in statistics, teaching experience, and years of assisting students are woven into the fabric of every edition of this book. His students had the utmost praise for his courses and his teaching ability. Fred was appreciated as a mentor to students and faculty alike, including his fellow authors. Yet, he was modest despite his accomplishments, and he was approachable and engaging. We were reminded of his contributions as we worked on each chapter during the course of this revision and were guided by his vision during this process. He has, no doubt, left a lasting legacy for his students and colleagues. We were most fortunate to benefit from his friendship, and he is sorely missed.

Larry B. Wallnau
Lori-Ann B. Forzano
James E. Witnauer

ABOUT THE AUTHORS



FREDERICK J GRAVETTER was Professor Emeritus of Psychology at The College at Brockport, State University of New York. While teaching at Brockport, Dr. Gravetter specialized in statistics, experimental design, and cognitive psychology. He received his bachelor's degree in mathematics from M.I.T. and his Ph.D. in psychology from Duke University. In addition to publishing this textbook and several research articles, Dr. Gravetter coauthored all editions of the best-selling *Statistics for the Behavioral Sciences, Essentials of Statistics for the Behavioral Sciences,* and *Research Methods for the Behavioral Sciences*. Dr. Gravetter passed away in November 2017.

LARRY B. WALLNAU is Professor Emeritus of Psychology at The College at Brockport, State University of New York. While teaching at Brockport, his research has been published in journals such as *Pharmacology Biochemistry and Behavior, Physiology and Behavior, Journal of Human Evolution, Folia Primatologica,* and *Behavior Research Methods and Instrumentation*. He also has provided editorial consultation. His courses have included statistics, biopsychology, animal behavior, psychopharmacology, and introductory psychology. With Dr. Gravetter, he coauthored all previous editions of *Statistics for the Behavioral Sciences* and *Essentials of Statistics for the Behavioral Sciences*. Dr. Wallnau received his bachelor's degree from the University of New Haven and his Ph.D. in psychology from the State University of New York at Albany. In his leisure time, he is an avid runner with his canine companion, Gracie.

LORI-ANN B. FORZANO is Professor of Psychology at The College at Brockport, State University of New York, where she regularly teaches undergraduate and graduate courses in research methods, statistics, learning, animal behavior, and the psychology of eating. She earned a Ph.D. in experimental psychology from the State University of New York at Stony Brook, where she also received her B.S. in psychology. Dr. Forzano's research examines impulsivity and self-control in adults and children. Her research has been published in the *Journal of the Experimental Analysis of Behavior, Learning and Motivation,* and *The Psychological Record.*

Dr. Forzano has also coauthored *Essentials of Statistics for the Behavioral Sciences, Ninth Edition,* and all previous editions of *Research Methods for the Behavioral Sciences,* now in its sixth edition.

JAMES E. WITNAUER is Associate Professor of Psychology at The College at Brockport, State University of New York, where he teaches undergraduate courses in experimental psychology and graduate courses in statistics and biopsychology. He earned a Ph.D. in cognitive psychology from State University of New York, Binghamton, and a B.A. in psychology from State University of New York, Buffalo State College. Dr. Witnauer's research aims to test mathematical models of learning and behavior and has been published in *Behavioural Processes*, *Journal of Experimental Psychology: Animal Behavior Processes*, and *Neurobiology of Learning and Memory.*

CHAPTER 1

Introduction to Statistics

clivewa/Shutterstock.com

PREVIEW

PREVIEW

Before we begin our discussion of statistics, we ask you to take a few moments to read the following paragraph, which has been adapted from a classic psychology experiment reported by Bransford and Johnson (1972).

> The procedure is actually quite simple. First you arrange things into different groups depending on their makeup. Of course, one pile may be sufficient, depending on how much there is to do. If you have to go somewhere else due to lack of facilities, that is the next step; otherwise you are pretty well set. It is important not to overdo any particular endeavor. That is, it is better to do too few things at once than too many. In the short run this may not seem important, but complications from doing too many can easily arise. A mistake can be expensive as well. The manipulation of the appropriate mechanisms should be self-explanatory, and we need not dwell on it here. At first the whole procedure will seem complicated. Soon, however, it will become just another facet of life. It is difficult to foresee any end to the necessity for this task in the immediate future, but then one never can tell.

You probably find the paragraph a little confusing, and most of you probably think it is describing some obscure statistical procedure. Actually, the paragraph describes the everyday task of doing laundry. Now that you know the topic (or context) of the paragraph, try reading it again—it should make sense now.

Why did we begin a statistics textbook with a paragraph about washing clothes? Our goal is to demonstrate the importance of context—when not in the proper context, even the simplest material can appear difficult and confusing. In the Bransford and Johnson (1972) experiment, people who knew the topic before reading the paragraph were able to recall 73% more than people who did not know that it was about laundry. When you are given the appropriate background, it is much easier to fit new material into your memory and recall it later. In this book each chapter begins with a preview that provides the background context for the new material in the chapter. As you read each preview section, you should gain a general overview of the chapter content. Similarly, we begin each section within each chapter with clearly stated learning objectives that prepare you for the material in that section. Finally, we introduce each new statistical procedure by explaining its purpose. Note that all statistical methods were developed to serve a purpose. If you understand why a new procedure is needed, you will find it much easier to learn and remember the procedure.

The objectives for this first chapter are to provide an introduction to the topic of statistics and to give you some background for the rest of the book. We will discuss the role of statistics in scientific inquiry, and we will introduce some of the vocabulary and notation that are necessary for the statistical methods that follow. In some respects, this chapter serves as a preview section for the rest of the book.

As you read through the following chapters, keep in mind that the general topic of statistics follows a well-organized, logically developed progression that leads from basic concepts and definitions to increasingly sophisticated techniques. Thus, the material presented in the early chapters of this book will serve as a foundation for the material that follows, even if those early chapters seem basic. The content of the first seven chapters provides an essential background and context for the statistical methods presented in Chapter 8. If you turn directly to Chapter 8 without reading the first seven chapters, you will find the material incomprehensible. However, if you learn the background material and practice the statistics procedures and methods described in early chapters, you will have a good frame of reference for understanding and incorporating new concepts as they are presented in each new chapter.

Finally, we cannot promise that learning statistics will be as easy as washing clothes. But if you begin each new topic with the proper context, you should eliminate some unnecessary confusion.

1-1 Statistics and Behavioral Sciences

LEARNING OBJECTIVES

1. Define the terms population, sample, parameter, and statistic, and describe the relationships between them; identify examples of each.
2. Define the two general categories of statistics, descriptive and inferential statistics, and describe how they are used to summarize and make decisions about data.
3. Describe the concept of sampling error and explain how sampling error creates the fundamental problem that inferential statistics must address.

Definitions of Statistics

By one definition, *statistics* consist of facts and figures such as the average annual snowfall in Buffalo or the average yearly income of recent college graduates. These statistics are usually informative and time-saving because they condense large quantities of information into a few simple figures. Later in this chapter we return to the notion of calculating statistics (facts and figures) but, for now, we concentrate on a much broader definition of statistics. Specifically, we use the term statistics to refer to a general field of mathematics. In this case, we are using the term *statistics* as a shortened version of *statistical methods* or *statistical procedures*. For example, you are probably using this book for a statistics course in which you will learn about the statistical procedures that are used to summarize and evaluate research results in the behavioral sciences.

Research in the behavioral sciences (and other fields) involves gathering information. To determine, for example, whether college students learn better by reading material on printed pages or on a computer screen, you would need to gather information about students' study habits and their academic performance. When researchers finish the task of gathering information, they typically find themselves with pages and pages of measurements such as preferences, personality scores, opinions, and so on. In this book, we present the statistics that researchers use to analyze and interpret the information that they gather. Specifically, statistics serve two general purposes:

1. Statistics are used to organize and summarize the information so that the researcher can see what happened in the study and can communicate the results to others.
2. Statistics help the researcher to answer the questions that initiated the research by determining exactly what general conclusions are justified based on the specific results that were obtained.

The term **statistics** refers to a set of mathematical procedures for organizing, summarizing, and interpreting information.

Statistical procedures help ensure that the information or observations are presented and interpreted in an accurate and informative way. In somewhat grandiose terms, statistics help researchers bring order out of chaos. In addition, statistics provide researchers with a set of standardized techniques that are recognized and understood throughout the scientific community. Thus, the statistical methods used by one researcher will be familiar to other researchers, who can accurately interpret the statistical analysis with a full understanding of how it was done and what the results signify.

Populations and Samples

Research in the behavioral sciences typically begins with a general question about a specific group (or groups) of individuals. For example, a researcher may want to know what factors are associated with academic dishonesty among college students. Or a researcher may want to determine the effect of lead exposure on the development of emotional problems in school-age children. In the first example, the researcher is interested in the group of college students. In the second example the researcher is studying school-age children. In statistical terminology, a *population* consists of all possible members of the group a researcher wishes to study.

> A **population** is the set of all the individuals of interest in a particular study.

As you can well imagine, a population can be quite large—for example, the entire set of all registered voters in the United States. A researcher might be more specific, limiting the study's population to people in their twenties who are registered voters in the United States. A smaller population would be first-time voter registrants in Burlington, Vermont. Populations can be extremely small too, such as those for people with a rare disease or members of an endangered species. The Siberian tiger, for example, has a population of roughly only 500 animals.

Thus, populations can obviously vary in size from extremely large to very small, depending on how the investigator identifies the population to be studied. The researcher should always specify the population being studied. In addition, the population need not consist of people—it could be a population of laboratory rats, North American corporations, engine parts produced in an automobile factory, or anything else an investigator wants to study. In practice, however, populations are typically very large, such as the population of college sophomores in the United States or the population of coffee drinkers that patronize a major national chain of cafés.

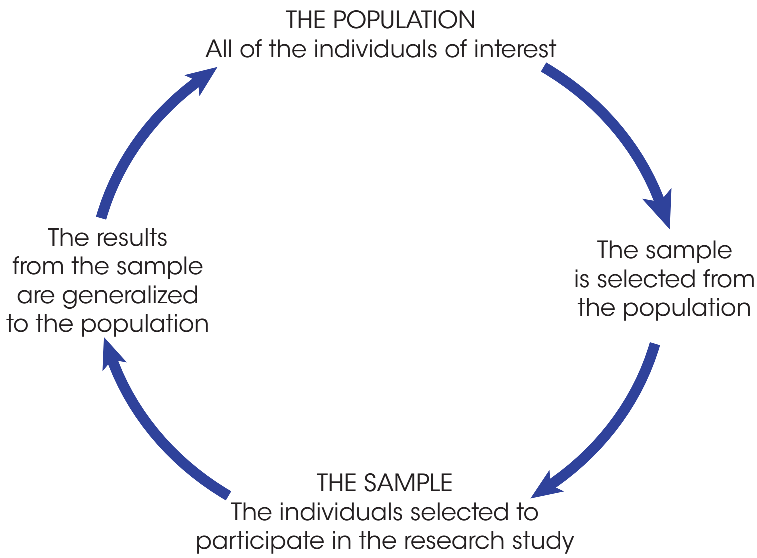
Because populations tend to be very large, it usually is impossible for a researcher to examine every individual in the population of interest. Therefore, researchers typically select a smaller, more manageable group from the population and limit their studies to the individuals in the selected group. In statistical terms, a set of individuals selected from a population is called a *sample*. A sample is intended to be representative of its population, and a sample should always be identified in terms of the population from which it was selected. We shall see later that one way to ensure that a sample is representative of a population is to select a *random sample*. In random sampling every individual has the same chance of being selected from the population.

> A **sample** is a set of individuals selected from a population, usually intended to represent the population in a research study. In a **random sample** everyone in the population has an equal chance of being selected.

Just as we saw with populations, samples can vary in size. For example, one study might examine a sample of only 20 middle-school students in an experimental reading program, and another study might use a sample of more than 2,000 people who take a new cholesterol medication.

So far, we have talked about a sample being selected from a population. However, this is actually only half of the full relationship between a sample and its population. Specifically, when a researcher finishes examining the sample, the goal is to generalize the results

FIGURE 1.1
The relationship between a population and a sample.



back to the entire population. Remember that the researcher started with a general question about the population. To answer the question, a researcher studies a sample and then generalizes the results from the sample to the population. The full relationship between a sample and a population is shown in Figure 1.1.

Variables and Data

Typically, researchers are interested in specific characteristics of the individuals in the population (or in the sample), or they are interested in outside factors that may influence behavior of the individuals. For example, Bakhshi, Kanuparthy, and Gilbert (2014) wanted to determine if the weather is related to online ratings of restaurants. As the weather changes, do people's reviews of restaurants change too? Something that can change or have different values is called a *variable.*

> A **variable** is a characteristic or condition that changes or has different values for different individuals.

In the case of the previous example, both weather and people's reviews of restaurants are variables. By the way, in case you are wondering, the authors did find a relationship between weather and online reviews of restaurants. Reviews were worse during bad weather (for example, during extremely hot or cold days).

Once again, variables can be characteristics that differ from one individual to another, such as weight, gender identity, personality, or motivation and behavior. Also, variables can be environmental conditions that change, such as temperature, time of day, or the size of the room in which the research is being conducted.

To demonstrate changes in variables, it is necessary to make measurements of the variables being examined. The measurement obtained for each individual is called a *datum*, or more commonly, a *score* or *raw score.* The complete set of scores is called the *data set* or simply the *data.*

Data (plural) are measurements or observations. A **data set** is a collection of measurements or observations. A **datum** (singular) is a single measurement or observation and is commonly called a **score** or **raw score**.

Before we move on, we should make one more point about samples, populations, and data. Earlier, we defined populations and samples in terms of *individuals*. For example, we previously discussed a population of registered voters and a sample of middle-school children. Be forewarned, however, that we will also refer to populations or samples of *scores*. Research typically involves measuring each individual to obtain a score, therefore every sample (or population) of individuals produces a corresponding sample (or population) of scores.

Parameters and Statistics

When describing data it is necessary to distinguish whether the data come from a population or a sample. A characteristic that describes a population—for example, the average score for the population—is called a *parameter*. A characteristic that describes a sample is called a *statistic*. Thus, the average score for a sample is an example of a statistic. Typically, the research process begins with a question about a population parameter. However, the actual data come from a sample and are used to compute sample statistics.

A **parameter** is a value, usually a numerical value, that describes a population. A parameter is usually derived from measurements of the individuals in the population.

A **statistic** is a value, usually a numerical value, that describes a sample. A statistic is usually derived from measurements of the individuals in the sample.

Every population parameter has a corresponding sample statistic, and most research studies involve using statistics from samples as the basis for answering questions about population parameters. As a result, much of this book is concerned with the relationship between sample statistics and the corresponding population parameters. In Chapter 7, for example, we examine the relationship between the mean obtained for a sample and the mean for the population from which the sample was obtained.

Descriptive and Inferential Statistical Methods

Although researchers have developed a variety of different statistical procedures to organize and interpret data, these different procedures can be classified into two general categories. The first category, *descriptive statistics*, consists of statistical procedures that are used to simplify and summarize data.

Descriptive statistics are statistical procedures used to summarize, organize, and simplify data.

Descriptive statistics are techniques that take raw scores and organize or summarize them in a form that is more manageable. Often the scores are organized in a table or graph so that it is possible to see the entire set of scores. Another common technique is to summarize a set of scores by computing an average. Note that even if the data set has hundreds of scores, the average provides a single descriptive value for the entire set.

The second general category of statistical techniques is called *inferential statistics*. Inferential statistics are methods that use sample data to make general statements about a population.

Inferential statistics consist of techniques that allow us to study samples and then make generalizations about the populations from which they were selected.

Because populations are typically very large, it usually is not possible to measure everyone in the population. Therefore, researchers select a sample that represents the population. By analyzing the data from the sample, we hope to make general statements about the population. Typically, researchers use sample statistics as the basis for drawing conclusions about population parameters or relationships between variables that might exist in the population. One problem with using samples, however, is that a sample provides only limited information about the population. Although samples are generally *representative* of their populations, a sample is not expected to give a perfectly accurate picture of the whole population. There usually is some discrepancy between a sample statistic and the corresponding population parameter. This discrepancy is called *sampling error*, and it creates the fundamental problem inferential statistics must always address.

Sampling error is the naturally occurring discrepancy, or error, that exists between a sample statistic and the corresponding population parameter.

The concept of sampling error is illustrated in Figure 1.2. The figure shows a population of 1,000 college students and two samples, each with five students who were selected from the population. Notice that each sample contains different individuals who have different characteristics. Because the characteristics of each sample depend on the specific people in the sample, statistics will vary from one sample to another. For example, the five students in sample 1 have an average age of 19.8 years and the students in sample 2 have an average age of 20.4 years. It is unlikely that the statistics for a sample will be identical to the parameter for the entire population. Both of the statistics in the example vary slightly from the population parameter (21.3 years) from which the samples were drawn. The difference between these sample statistics and the population parameter illustrate sampling error.

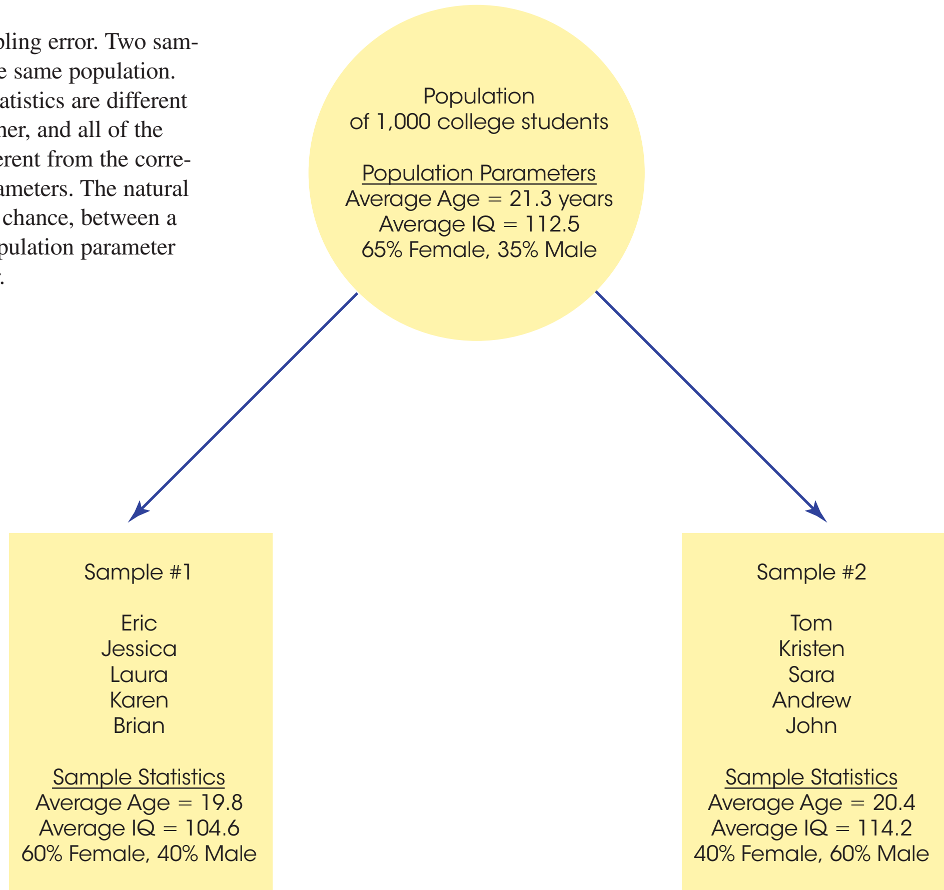
You should also realize that Figure 1.2 shows only two of the hundreds of possible samples. Each sample would contain different individuals and would produce different statistics. This is the basic concept of sampling error: sample statistics vary from one sample to another and typically are different from the corresponding population parameters.

One common example of sampling error is the error associated with a sample proportion (or percentage). For instance, in newspaper articles reporting results from political polls, you frequently find statements such as this:

> Candidate Brown leads the poll with 51% of the vote. Candidate Jones has 42% approval, and the remaining 7% are undecided. This poll was taken from a sample of registered voters and has a margin of error of plus or minus 4 percentage points.

The "margin of error" is the sampling error. In this case, the reported percentages were obtained from a sample and are being generalized to the whole population of potential voters.

FIGURE 1.2
A demonstration of sampling error. Two samples are selected from the same population. Notice that the sample statistics are different from one sample to another, and all of the sample statistics are different from the corresponding population parameters. The natural differences that exist, by chance, between a sample statistic and a population parameter are called sampling error.



As always, you do not expect the statistics from a sample to be a perfect reflection of the population. There always will be some "margin of error" when sample statistics are used to represent population parameters.

As a further demonstration of sampling error, imagine that your statistics class is separated into two groups by drawing a line from front to back through the middle of the room. Now imagine that you compute the average age (or height, or GPA) for each group. Will the two groups have exactly the same average? Almost certainly they will not. No matter what you choose to measure, you will probably find some difference between the two groups. However, the difference you obtain does not necessarily mean that there is a systematic difference between the two groups. For example, if the average age for students on the right-hand side of the room is higher than the average for students on the left, it is unlikely that some mysterious force has caused the older people to gravitate to the right side of the room. Instead, the difference is probably the result of random factors such as chance. The unpredictable, unsystematic differences that exist from one sample to another are an example of sampling error. Inferential statistics tell us whether the differences between samples (e.g., a difference in age, height, or GPA) are the result of random factors (sampling error) or the result of some meaningful relationship in the population.

Statistics in the Context of Research

The following example shows the general stages of a research study and demonstrates how descriptive statistics and inferential statistics are used to organize and interpret the data. At the end of the example, note how sampling error can affect the interpretation of experimental results, and consider why inferential statistical methods are needed to deal with this problem.

EXAMPLE 1.1

Figure 1.3 shows an overview of a general research situation and demonstrates the roles that descriptive and inferential statistics play. The purpose of the research study is to address a question that we posed earlier: do college students learn better by studying text on printed pages or on a computer screen? Two samples of six students each are selected from the population of college students. The students in sample A read text on a computer

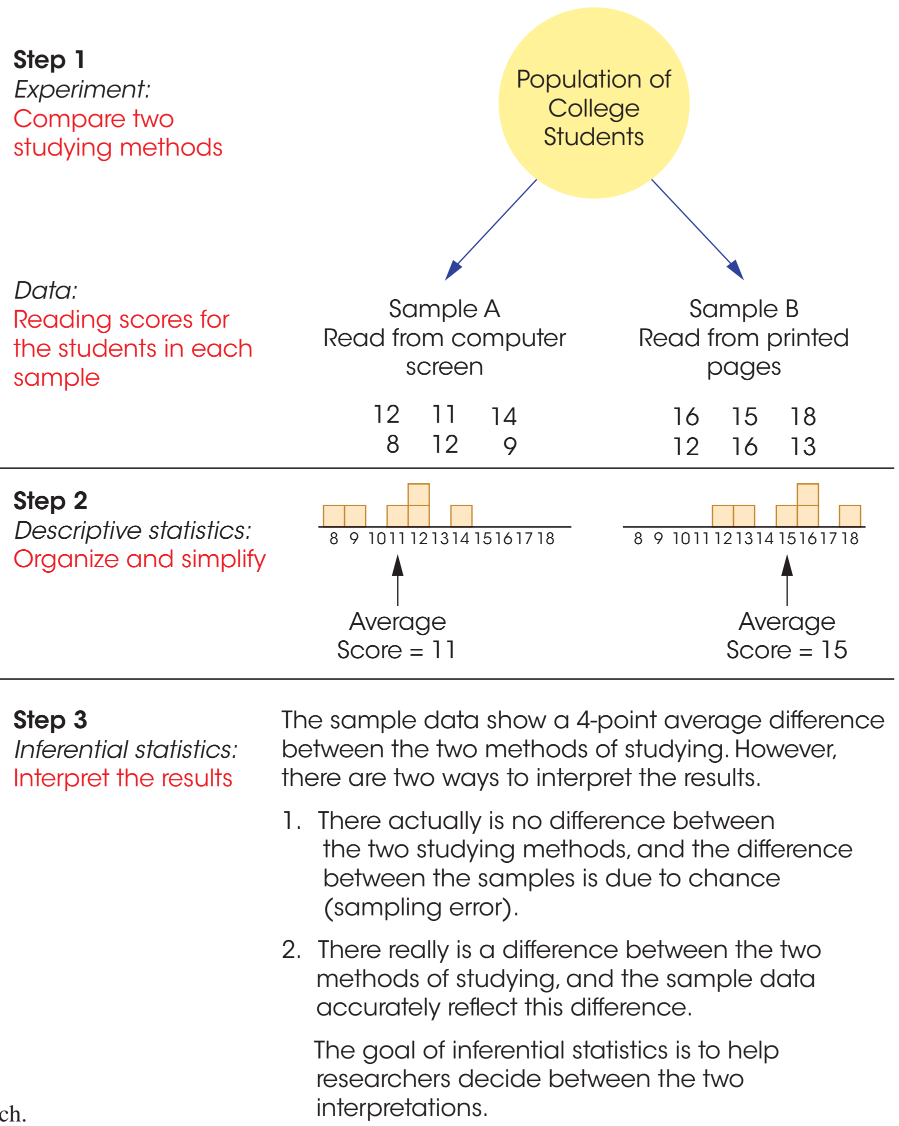


FIGURE 1.3
The role of statistics in research.

screen to study for 30 minutes and the students in sample B are given printed pages. Next, all of the students are given a multiple-choice test to evaluate their knowledge of the material. At this point, the researcher has two groups of data: the scores for sample A and the scores for sample B (see the figure). Now is the time to begin using statistics.

First, descriptive statistics are used to simplify the pages of data. For example, the researcher could draw a graph showing the scores for each sample or compute the average score for each group. Note that descriptive methods provide a simplified, organized description of the scores. In this example, the students who studied text on the computer screen averaged 11 on the test, and the students who studied printed pages had an average score of 15. These descriptive statistics efficiently summarize—with only two values—the two samples containing six scores each.

Once the researcher has described the results, the next step is to interpret the outcome. This is the role of inferential statistics. In this example, the researcher has found a difference of 4 points between the two samples (sample A averaged 11 and sample B averaged 15). The problem for inferential statistics is to differentiate between the following two interpretations:

1. There is no real difference between the printed page and a computer screen, and the 4-point difference between the samples is just an example of sampling error (like the samples in Figure 1.2).
2. There really is a difference between the printed page and a computer screen, and the 4-point difference between the samples was caused by the different methods of studying.

In simple English, does the 4-point difference between samples provide convincing evidence of a difference between the two studying methods, or is the 4-point difference just chance? Inferential statistics attempt to answer this question. ■

LEARNING CHECK

Note that each chapter section begins with a list of Learning Objectives (see page 3 for an example) and ends with a Learning Check to test your mastery of the objectives. Each Learning Check question is preceded by its corresponding Learning Objective number.

LO1 **1.** A researcher is interested in the Netflix binge-watching habits of American college students. A group of 50 students is interviewed and the researcher finds that these students stream an average of 6.7 hours per week. For this study, the average of 6.7 hours is an example of a(n) ____________.

- **a.** parameter
- **b.** statistic
- **c.** population
- **d.** sample

LO2 **2.** Researchers are interested in how robins in New York State care for their newly hatched chicks. The team measures how many times per day the adults visit their nests to feed their young. The entire group of robins in the state is an example of a ____________.

- **a.** sample
- **b.** statistic
- **c.** population
- **d.** parameter

LO2 **3.** Statistical techniques that use sample data to draw conclusions about the population are ______________.

a. population statistics

b. sample statistics

c. descriptive statistics

d. inferential statistics

LO3 **4.** The SAT is standardized so that the population average score on the verbal test is 500 each year. In a sample of 100 graduating seniors who have taken the verbal SAT, what value would you expect to obtain for their average verbal SAT score?

a. 500

b. Greater than 500

c. Less than 500

d. Around 500 but probably not equal to 500

ANSWERS **1. b 2. c 3. d 4. d**

1-2 Observations, Measurement, and Variables

LEARNING OBJECTIVES

4. Explain why operational definitions are developed for constructs and identify the two components of an operational definition.

5. Describe discrete and continuous variables and identify examples of each.

6. Define real limits and explain why they are needed to measure continuous variables.

7. Compare and contrast the four scales of measurement (nominal, ordinal, interval, and ratio) and identify examples of each.

Observations and Measurements

Science is *empirical.* This means it is based on observation rather than intuition or conjecture. Whenever we make a precise observation we are taking a measurement, either by assigning a numerical value to observations or by classifying them into categories. Observation and measurement are part and parcel of the scientific method. In this section, we take a closer look at the variables that are being measured and the process of measurement.

Constructs and Operational Definitions

The scores that make up the data from a research study are the result of observing and measuring variables. For example, a researcher may obtain a set of memory recall scores, personality scores, or reaction-time scores when conducting a study. Some variables, such as height, weight, and eye color are well-defined, concrete entities that can be observed and measured directly. On the other hand, many variables studied by behavioral scientists are internal characteristics that people use to help describe and explain behavior. For example, we say that a student does well in school because the student has strong *motivation* for achievement. Or we say that someone is *anxious* in social situations, or that someone seems to be *hungry*. Variables like motivation, anxiety, and hunger are called *constructs*, and because they are intangible and cannot be directly observed, they are often called *hypothetical constructs*.

Although constructs such as intelligence are internal characteristics that cannot be directly observed, it is possible to observe and measure behaviors that are representative of the construct. For example, we cannot "see" high self-esteem but we can see examples of behavior reflective of a person with high self-esteem. The external behaviors can then be used to create an operational definition for the construct. An *operational definition* defines a construct in terms of external behaviors that can be observed and measured. For example, your self-esteem is measured and operationally defined by your score on the Rosenberg Self-Esteem Scale, or hunger can be measured and defined by the number of hours since last eating.

Constructs are internal attributes or characteristics that cannot be directly observed but are useful for describing and explaining behavior.

An **operational definition** identifies a measurement procedure (a set of operations) for measuring an external behavior and uses the resulting measurements as a definition and a measurement of a hypothetical construct. Note that an operational definition has two components. First, it describes a set of operations for measuring a construct. Second, it defines the construct in terms of the resulting measurements.

Discrete and Continuous Variables

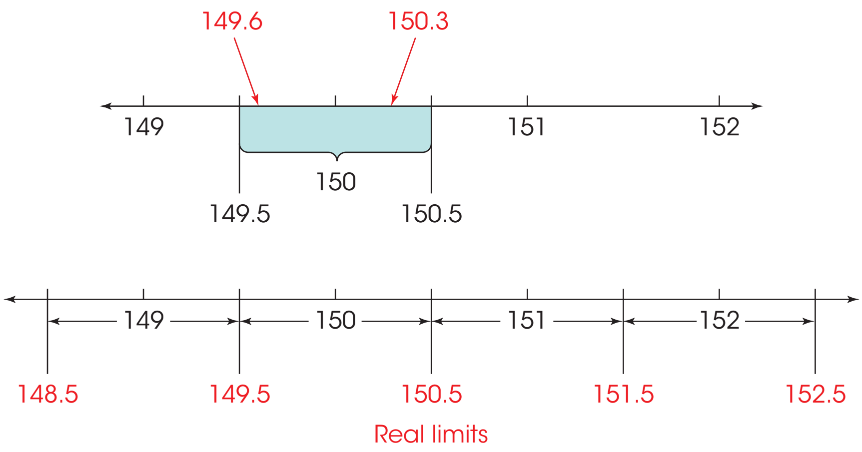
The variables in a study can be characterized by the type of values that can be assigned to them and, as we will discuss in later chapters, the type of values influences the statistical procedures that can be used to summarize or make inferences about those values. A *discrete variable* consists of separate, indivisible categories. For this type of variable, there are no intermediate values between two adjacent categories. Consider the number of questions that each student answers correctly on a 10-item multiple-choice quiz. Between neighboring values—for example, seven correct and eight correct—no other values can ever be observed.

A **discrete variable** consists of separate, indivisible categories. No values can exist between two neighboring categories.

Discrete variables are commonly restricted to whole, countable numbers (i.e., integers)—for example, the number of children in a family or the number of students attending class. If you observe class attendance from day to day, you may count 18 students one day and 19 students the next day. However, it is impossible ever to observe a value between 18 and 19. A discrete variable may also consist of observations that differ qualitatively. For example, people can be classified by birth order (first-born or later-born), by occupation (nurse, teacher, lawyer, etc.), and college students can be classified by academic major (art, biology, chemistry, etc.). In each case, the variable is discrete because it consists of separate, indivisible categories.

On the other hand, many variables are not discrete. Variables such as time, height, and weight are not limited to a fixed set of separate, indivisible categories. You can measure time, for example, in hours, minutes, seconds, or fractions of seconds. These variables are called *continuous* because they can be divided into an infinite number of fractional parts.

FIGURE 1.4
When measuring weight to the nearest whole pound, 149.6 and 150.3 are assigned the value of 150 (top). Any value in the interval between 149.5 and 150.5 is given the value of 150.



For a **continuous variable**, there are an infinite number of possible values that fall between any two observed values. A continuous variable is divisible into an infinite number of fractional parts.

Suppose, for example, that a researcher is measuring weights for a group of individuals participating in a diet study. Because weight is a continuous variable, it can be pictured as a continuous line (Figure 1.4). Note that there are an infinite number of possible points on the line without any gaps or separations between neighboring points. For any two different points on the line, it is always possible to find a third value that is between the two points.

Two other factors apply to continuous variables:

1. When measuring a continuous variable, it should be very rare to obtain identical measurements for two different individuals. Because a continuous variable has an infinite number of possible values, it should be almost impossible for two people to have exactly the same score. If the data show a substantial number of tied scores, then you should suspect either the variable is not really continuous or that the measurement procedure is very crude—meaning the continuous variable is divided into widely separated discrete numbers.
2. When measuring a continuous variable, researchers must first identify a series of measurement categories on the scale of measurement. Measuring weight to the nearest pound, for example, would produce categories of 149 pounds, 150 pounds, and so on. However, each measurement category is actually an *interval* that must be defined by boundaries. To differentiate a weight of 150 pounds from the surrounding values of 149 and 151, we must set up boundaries on the scale of measurement. These boundaries are called *real limits* and are positioned exactly halfway between adjacent scores. Thus, a score of 150 pounds is actually an interval bounded by a *lower real limit* of 149.5 at the bottom and an *upper real limit* of 150.5 at the top. Any individual whose weight falls between these real limits will be assigned a score of 150. As a result, two people who both claim to weigh 150 pounds are probably not *exactly* the same weight. One person may actually weigh 149.6 and the other 150.3, but they are both assigned a weight of 150 pounds (see Figure 1.4).

Students often ask whether a measurement of exactly 150.5 should be assigned a value of 150 or a value of 151. The answer is that 150.5 is the *boundary* between the two intervals and is not necessarily in one or the other. Instead, the placement of 150.5 depends on the rule that you are using for rounding numbers. If you are rounding up, then 150.5 goes in the higher interval (151) but if you are rounding down, then it goes in the lower interval (150).

Real limits are the boundaries of intervals for scores that are represented on a continuous number line. The real limit separating two adjacent scores is located exactly halfway between the scores. Each score has two real limits. The **upper real limit** is at the top of the interval, and the **lower real limit** is at the bottom.

The concept of real limits applies to any measurement of a continuous variable, even when the score categories are not whole numbers. For example, if you were measuring time to the nearest tenth of a second, the measurement categories would be 31.0, 31.1, 31.2, and so on. Each of these categories represents an interval on the scale that is bounded by real limits. For example, a score of $X = 31.1$ seconds indicates that the actual measurement is in an interval bounded by a lower real limit of 31.05 and an upper real limit of 31.15. Remember that the real limits are always halfway between adjacent categories.

Later in this book, real limits are used for constructing graphs and for various calculations with continuous scales. For now, however, you should realize that real limits are a necessity whenever you make measurements of a continuous variable.

Finally, we should warn you that the terms *continuous* and *discrete* apply to the variables that are being measured and not to the scores that are obtained from the measurement. For example, measuring people's heights to the nearest inch produces scores of 60, 61, 62, and so on. Although the scores may appear to be discrete numbers, the underlying variable is continuous. One key to determining whether a variable is continuous or discrete is that a continuous variable can be divided into any number of fractional parts. Height can be measured to the nearest inch, the nearest 0.5 inch, or the nearest 0.1 inch. Similarly, a professor evaluating students' knowledge could use a pass/fail system that classifies students into two broad categories. However, the professor could choose to use a 10-point quiz that divides student knowledge into 11 categories corresponding to quiz scores from 0 to 10. Or the professor could use a 100-point exam that potentially divides student knowledge into 101 categories from 0 to 100. Whenever you are free to choose the degree of precision or the number of categories for measuring a variable, the variable must be continuous.

■ Scales of Measurement

It should be obvious by now that data collection requires that we make measurements of our observations. Measurement involves assigning individuals or events to categories. The categories can simply be names such as introvert/extrovert or employed/unemployed, or they can be numerical values such as 68 inches or 175 pounds. The categories used to measure a variable make up a *scale of measurement*, and the relationships between the categories determine different types of scales. The distinctions among the scales are important because they identify the limitations of certain types of measurements and because certain statistical procedures are appropriate for scores that have been measured on some scales but not on others. If you were interested in people's heights, for example, you could measure a group of individuals by simply classifying them into three categories: tall, medium, and short. However, this simple classification would not tell you much about the actual heights of the individuals, and these measurements would not give you enough information to calculate an average height for the group. Although the simple classification would be adequate for some purposes, you would need more sophisticated measurements before you could answer more detailed questions. In this section, we examine four different scales of measurement, beginning with the simplest and moving to the most sophisticated.

The Nominal Scale The word *nominal* means "having to do with names." Measurement on a nominal scale involves classifying individuals into categories that have different names but are not quantitative or numerically related to each other. For example, if you were measuring the academic majors for a group of college students, the categories would be art, biology, business, chemistry, and so on. Each student would be classified in one category according to his or her major. The measurements from a nominal scale allow us to determine whether two individuals are different, but they do not identify either the direction or the size of the difference. If one student is an art major and another is a biology major we can say that they are different, but we cannot say that art is "more than" or "less than" biology and we cannot specify how much difference there is between art and biology. Other examples of nominal scales include classifying people by race, gender, or occupation.

> A **nominal scale** consists of a set of categories that have different names. Measurements on a nominal scale label and categorize observations, but do not make any quantitative distinctions between observations.

Although the categories on a nominal scale are not quantitative values, they are occasionally represented by numbers. For example, the rooms or offices in a building may be identified by numbers. You should realize that the room numbers are simply names and do not reflect any quantitative information. Room 109 is not necessarily bigger than Room 100 and certainly not 9 points bigger. It also is fairly common to use numerical values as a code for nominal categories when data are entered into computer programs for analysis. For example, the data from a political opinion poll may code Democrats with a 0 and Republicans with a 1 as a group identifier. Again, the numerical values are simply names and do not represent any quantitative difference. The scales that follow do reflect an attempt to make quantitative distinctions.

The Ordinal Scale The categories that make up an *ordinal scale* not only have different names (as in a nominal scale) but also are organized in a fixed order corresponding to differences of magnitude.

> An **ordinal scale** consists of a set of categories that are organized in an ordered sequence. Measurements on an ordinal scale rank observations in terms of size or magnitude.

Often, an ordinal scale consists of a series of ranks (first, second, third, and so on) like the order of finish in a horse race. Occasionally, the categories are identified by verbal labels (like small, medium, and large drink sizes at a fast-food restaurant). In either case, the fact that the categories form an ordered sequence means that there is a directional relationship between categories. With measurements from an ordinal scale you can determine whether two individuals are different, and you can determine the direction of difference. However, ordinal measurements do not allow you to determine the size of the difference between two individuals. For example, suppose in the Winter Olympics you watch the medal ceremony for the women's downhill ski event. You know that the athlete receiving the gold medal had the fastest time, the silver medalist had the second fastest time, and the bronze medalist had the third fastest time. This represents an ordinal scale of measurement and reflects no more information than first, second, and third place. Note that it does not provide information about how much time difference there was between competitors. The first-place skier might have won the event by a mere one one-hundredth of a second—or perhaps by as much as one second. Other examples of ordinal scales include socioeconomic class (upper, middle, lower) and T-shirt sizes (small, medium, large). In addition, ordinal scales are often used to measure variables for which it is difficult to assign numerical scores. For example, people can rank their food preferences but might have trouble explaining "how much more" they prefer chocolate ice cream to cheesecake.

The Interval and Ratio Scales Both an *interval scale* and a *ratio scale* consist of a series of ordered categories (like an ordinal scale) with the additional requirement that the categories form a series of intervals that are all exactly the same size. Thus, the scale of measurement consists of a series of equal intervals, such as inches on a ruler. Examples of interval scales are the temperature in degrees Fahrenheit or Celsius and examples of ratio scales are the measurement of time in seconds or weight in pounds. Note that, in each case, the difference between two adjacent values (1 inch, 1 second, 1 pound, 1 degree) is the same size, no matter where it is located on the scale. The fact that the differences between adjacent values are all the same size makes it possible to determine both the size and the direction of the difference between two measurements. For example, you know that a measurement of 80° Fahrenheit is higher than a measure of 60°, and you know that it is exactly 20° higher.

The factor that differentiates an interval scale from a ratio scale is the nature of the zero point. An interval scale has an arbitrary zero point. That is, the value 0 is assigned to a particular location on the scale simply as a matter of convenience or reference. In particular, a value of zero does not indicate a total absence of the variable being measured. The two most common examples are the Fahrenheit and Celsius temperature scales. For example, a temperature of 0º Fahrenheit does not mean that there is no temperature, and it does not prohibit the temperature from going even lower. Interval scales with an arbitrary zero point are not common in the physical sciences or with physical measurements.

A ratio scale is anchored by a zero point that is not arbitrary but rather is a meaningful value representing none (a complete absence) of the variable being measured. The existence of an absolute, non-arbitrary zero point means that we can measure the absolute amount of the variable; that is, we can measure the distance from 0. This makes it possible to compare measurements in terms of ratios. For example, a fuel tank with 10 gallons of gasoline has twice as much gasoline as a tank with only 5 gallons because there is a true absolute zero value. A completely empty tank has 0 gallons of fuel. Ratio scales are used in the behavioral sciences, too. A reaction time of 500 milliseconds is exactly twice as long as a reaction time of 250 milliseconds and a value of 0 milliseconds is a true absolute zero. To recap, with a ratio scale, we can measure the direction and the size of the difference between two measurements and we can describe the difference in terms of a ratio. Ratio scales are common and include physical measurements such as height and weight, as well as measurements of variables such as reaction time or the number of errors on a test. The distinction between an interval scale and a ratio scale is demonstrated in Example 1.2 and in Table 1.1.

An **interval scale** consists of ordered categories that are all intervals of exactly the same size. Equal differences between numbers on a scale reflect equal differences in magnitude. However, the zero point on an interval scale is arbitrary and does not indicate a zero amount of the variable being measured.

A **ratio scale** is an interval scale with the additional feature of an absolute zero point. With a ratio scale, ratios of numbers do reflect ratios of magnitude.

TABLE 1.1
Scales of Measurement for a Marathon

Scale	Information	Example
Nominal	Category only	Country of athlete (U.S., U.K., Ethiopia, Japan, Kenya, etc.)
Ordinal	*Ordered* category	Finishing position in a race (1st, 2nd, 3rd, etc.)
Interval	Ordered category *with equal intervals separating adjacent scores and arbitrary (not absolute) zero*	Time difference (above or below) from the course record, an arbitrary zero point (Example: a person who finishes the Boston Marathon 4 minutes slower than the course record takes 3 minutes longer to finish the race than a person who was 1 minute slower than the course record, but does not take four times longer.)
Ratio	Ordered category with equal amounts separating adjacent scores, *and a true absolute zero*	Amount of time to complete a marathon (Example: a person who finishes the Boston Marathon in 4 hours, 30 minutes takes 2 times longer than one who finishes in 2 hours, 15 minutes.)

EXAMPLE 1.2

A researcher obtains measurements of height for a group of 8-year-old boys. Initially, the researcher simply records each child's height in inches, obtaining values such as 44, 51, 49, and so on. These initial measurements constitute a ratio scale. A value of zero represents no height (absolute zero). Also, it is possible to use these measurements to form ratios. For example, a child who is 60 inches tall is one-and-a-half times taller than a child who is 40 inches tall.

Now suppose that the researcher converts the initial measurement into a new scale by calculating the difference between each child's actual height and the average height for this age group. A child who is 1 inch taller than average now gets a score of +1; a child 4 inches taller than average gets a score of +4. Similarly, a child who is 2 inches shorter than average gets a score of −2. On this scale, a score of zero corresponds to average height. Because zero no longer indicates a complete absence of height, the new scores constitute an interval scale of measurement.

Notice that original scores and the converted scores both involve measurement in inches, and you can compute differences, or distances, on either scale. For example, there is a 6-inch difference in height between two boys who measure 57 and 51 inches tall on the first scale. Likewise, there is a 6-inch difference between two boys who measure +9 and +3 on the second scale. However, you should also notice that ratio comparisons are not possible on the second scale. For example, a boy who measures +9 is not three times taller than a boy who measures +3. ■

Statistics and Scales of Measurement For our purposes, scales of measurement are important because they help determine the statistics that are used to evaluate the data. Specifically, there are certain statistical procedures that are used with numerical scores from interval or ratio scales and other statistical procedures that are used with non-numerical scores from nominal or ordinal scales. The distinction is based on the fact that numerical scores are compatible with basic arithmetic operations (adding, multiplying, and so on) but non-numerical scores are not. For example, in a memory experiment a researcher might record how many words participants can recall from a list they previously studied. It is possible to add the recall scores together to find a total and then calculate the average score for the group. On the other hand, if you measure the academic major for each student, you cannot add the scores to obtain a total. (What is the total for three psychology majors plus an English major plus two chemistry majors?) The vast majority of the statistical techniques presented in this book are designed for numerical scores from interval or ratio scales. For most statistical applications, the distinction between an interval scale and a ratio scale is not important because both scales produce numerical values that permit us to compute differences between scores, add scores, and calculate mean scores. On the other hand, measurements from nominal or ordinal scales are typically not numerical values, do not measure distance, and are not compatible with many basic arithmetic operations. Therefore, alternative statistical techniques are necessary for data from nominal or ordinal scales of measurement (for example, the median and the mode in Chapter 3, the Spearman correlation in Chapter 14, and the chi-square tests in Chapter 15).

LEARNING CHECK

LO4 **1.** An operational definition is used to ________ a hypothetical construct.

a. define
b. measure
c. measure and define
d. None of the other choices is correct.

LO5 **2.** A researcher studies the factors that determine the length of time a consumer stays on a website before clicking off. The variable, length of time, is an example of a ________ variable.

a. discrete
b. continuous
c. nominal
d. ordinal

LO5 **3.** A researcher records the number of bites a goat takes of different plants. The variable, number of bites, is an example of a ________ variable.

a. discrete
b. continuous
c. nominal
d. ordinal

LO6 **4.** When measuring height to the nearest inch, what are the real limits for a score of 68.0 inches?

a. 67 and 69
b. 67.5 and 68.5
c. 67.75 and 68.75
d. 67.75 and 68.25

LO7 **5.** The professor in a communications class asks students to identify their favorite reality television show. The different television shows make up a ________ scale of measurement.

a. nominal
b. ordinal
c. interval
d. ratio

LO7 **6.** Ranking jobs, taking into account growth potential, work-life balance, and salary, would be an example of measurement on a(n) ____________ scale.

a. nominal
b. ordinal
c. interval
d. ratio

ANSWERS **1. c 2. b 3. a 4. b 5. a 6. b**

1-3 Three Data Structures, Research Methods, and Statistics

LEARNING OBJECTIVES

8. Describe, compare, and contrast correlational, experimental, and nonexperimental research, and identify the data structures associated with each.

9. Define independent, dependent, and quasi-independent variables and recognize examples of each.

Data Structure 1. One Group with One or More Separate Variables Measured for Each Individual: Descriptive Research

Some research studies are conducted simply to describe individual variables as they exist naturally. For example, a college official may conduct a survey to describe the eating, sleeping, and study habits of a group of college students. Table 1.2 shows an example of data from this type of research. Although the researcher might measure several different variables, the goal of the study is to describe each variable separately. In particular, this type of research is not concerned with relationships between variables.

A study that produces the kind of data shown in Table 1.2 and is focused on describing individual variables rather than relationships is an example of *descriptive research* or the *descriptive research strategy*.

Descriptive research or the **descriptive research strategy** involves measuring one or more separate variables for each individual with the intent of simply describing the individual variables.

TABLE 1.2
Three separate variables measured for each individual in a group of eight students.

	Weekly Number of Student Fast-Food Meals	Number of Hours Sleeping Each Day	Number of Hours Studying Each Day
A	0	9	3
B	4	7	2
C	2	8	4
D	1	10	3
E	0	11	2
F	0	7	4
G	5	7	3
H	3	8	2

When the results from a descriptive research study consist of numerical scores—such as the number of hours spent studying each day—they are typically described by the statistical techniques that are presented in Chapters 3 and 4. For example, a researcher may want to know the average number of meals eaten at fast-food restaurants each week for students at the college. Non-numerical scores are typically described by computing the proportion or percentage in each category. For example, a recent newspaper article reported that 34.9% of American adults are obese, which is roughly 35 pounds over a healthy weight.

■ Relationships Between Variables

Most research, however, is intended to examine relationships between two or more variables. For example, is there a relationship between the amount of violence in the video games played by children and the amount of aggressive behavior they display? Is there a relationship between vocabulary development in childhood and academic success in college? To establish the existence of a relationship, researchers must make observations—that is, measurements of the two variables. The resulting measurements can be classified into two distinct data structures that also help to classify different research methods and different statistical techniques. In the following section we identify and discuss these two data structures.

■ Data Structure 2. One Group with Two Variables Measured for Each Individual: The Correlational Method

One method for examining the relationship between variables is to observe the two variables as they exist naturally for a set of individuals. That is, simply measure the two variables for each individual. For example, research results tend to find a relationship between Facebook™ use and academic performance, especially for freshmen (Junco, 2015). Figure 1.5 shows an example of data obtained by measuring time on Facebook and academic performance for eight students. The researchers then look for consistent patterns in the data to provide evidence for a relationship between variables. For example, as Facebook time changes from one student to another, is there also a tendency for academic performance to change?

Student	Facebook Time	Academic Performance
A	4	2.4
B	2	3.6
C	2	3.2
D	5	2.2
E	0	3.8
F	3	2.2
G	3	3.0
H	1	3.0

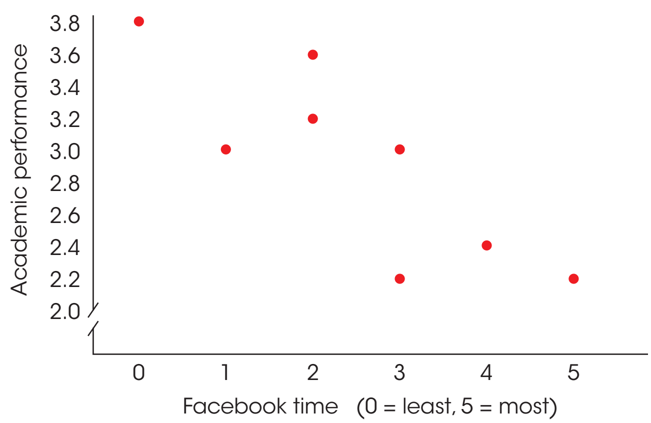


FIGURE 1.5
One of two data structures for studies evaluating the relationship between variables. Note that there are two separate measurements for each individual (Facebook time and academic performance). The same scores are shown in a table and a graph.

Consistent patterns in the data are often easier to see if the scores are presented in a graph. Figure 1.5 also shows the scores for the eight students in a graph called a scatter plot. In the scatter plot, each individual is represented by a point so that the horizontal position corresponds to the student's Facebook time and the vertical position corresponds to the student's academic performance score. The scatter plot shows a clear relationship between Facebook time and academic performance: as Facebook time increases, academic performance decreases.

A research study that simply measures two different variables for each individual and produces the kind of data shown in Figure 1.5 is an example of the *correlational method*, or the *correlational research strategy.*

In the **correlational method**, two different variables are observed to determine whether there is a relationship between them.

Statistics for the Correlational Method When the data from a correlational study consist of numerical scores, the relationship between the two variables is usually measured and described using a statistic called a correlation. Correlations and the correlational method are discussed in detail in Chapter 14. Occasionally, the measurement process used for a correlational study simply classifies individuals into categories that do not correspond to numerical values. For example, a researcher could classify study participants by age (40 years of age and over, or under 40 years) and by preference for smartphone use (talk or text). Note that the researcher has two scores for each individual (age category and phone use preference) but neither of the scores is a numerical value. These types of data are typically summarized in a table showing how many individuals are classified into each of the possible categories. Table 1.3 shows an example of this kind of summary table. The table shows, for example, that 15 of the people 40 and over in the sample preferred texting and 35 preferred talking. The pattern is quite different for younger participants—45 preferred texting and only 5 preferred talking. Note that by presenting the data in a table, one can see the difference in preference for age at a quick glance. The relationship between categorical variables (such as the data in Table 1.3) is usually evaluated using a statistical technique known as a *chi-square test.* Chi-square tests are presented in Chapter 15.

Limitations of the Correlational Method The results from a correlational study can demonstrate the existence of a relationship between two variables, but they do not provide an explanation for the relationship. In particular, a correlational study cannot demonstrate a cause-and-effect relationship. For example, the data in Figure 1.5 show a systematic relationship between Facebook time and academic performance for a group of college students; those who spend more time on Facebook tend to have lower grades. However, there are many possible explanations for the relationship and we do not know exactly what factor (or factors) is responsible for Facebook users having lower grades. For example,

TABLE 1.3 Correlational data consisting of non-numerical scores. Note that there are two measurements for each individual: age and smartphone preference. The numbers indicate how many people fall into each category.

	Smartphone Preference		
	Text	Talk	
40 years and over	15	35	50
Under 40	45	5	50

many students report that they multitask with Facebook while they are studying. In this case, their lower grades might be explained by the distraction of multitasking while studying. Another possible explanation is that there is a third variable involved that produces the relationship. For example, perhaps level of interest in the course material accounts for the relationship. That is, students who have less interest in the course material might study it less and spend more time on interesting pursuits like Facebook. In particular, we cannot conclude that simply reducing time on Facebook would cause their academic performance to improve. To demonstrate a cause-and-effect relationship between two variables, researchers must use the experimental method, which is discussed next.

■ Data Structure 3. Comparing Two (or More) Groups of Scores: Experimental and Nonexperimental Methods

The second method for examining the relationship between two variables compares two or more groups of scores. In this situation, the relationship between variables is examined by using one of the variables to define the groups, and then measuring the second variable to obtain scores for each group. For example, Polman, de Castro, and van Aken (2008) randomly divided a sample of 10-year-old boys into two groups. One group then played a violent video game and the second played a nonviolent game. After the game-playing session, the children went to a free play period and were monitored for aggressive behaviors (hitting, kicking, pushing, frightening, name-calling, fighting, quarreling, or teasing another child). An example of the resulting data is shown in Figure 1.6. The researchers then compared the scores for the violent-video group with the scores for the nonviolent-video group. A systematic difference between the two groups provides evidence for a relationship between playing violent video games and aggressive behavior for 10-year-old boys.

Statistics for Comparing Two (or More) Groups of Scores Most of the statistical procedures presented in this book are designed for research studies that compare groups of scores like the study in Figure 1.6. Specifically, we examine descriptive statistics that summarize and describe the scores in each group and we use inferential statistics to determine whether the differences between the groups can be generalized to the entire population.

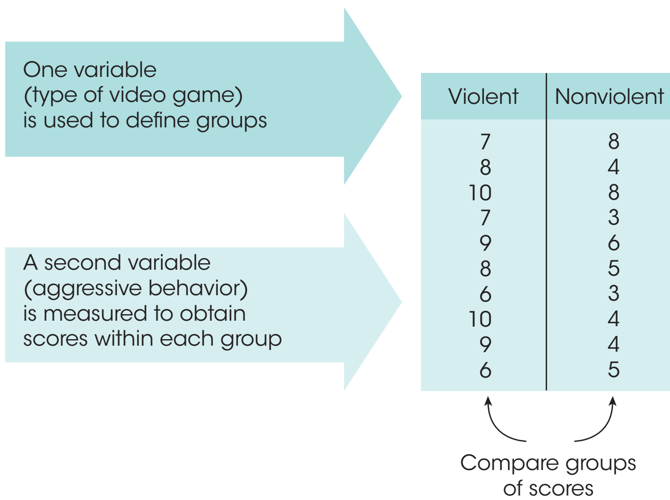


FIGURE 1.6
Evaluating the relationship between variables by comparing groups of scores. Note that the values of one variable are used to define the groups and the second variable is measured to obtain scores within each group.

When the measurement procedure produces numerical scores, the statistical evaluation typically involves computing the average score for each group and then comparing the averages. The process of computing averages is presented in Chapter 3, and a variety of statistical techniques for comparing averages are presented in Chapters 8–13. If the measurement process simply classifies individuals into non-numerical categories, the statistical evaluation usually consists of computing proportions for each group and then comparing proportions. Previously, in Table 1.3, we presented an example of non-numerical data examining the relationship between age and smartphone preference. The same data can be used to compare the proportions for participants age 40 and over with the proportions for those under 40 years of age. For example, 30% of people 40 and over prefer to text compared to 90% of those under 40. As before, these data are evaluated using a chi-square test, which is presented in Chapter 15.

Experimental and Nonexperimental Methods

There are two distinct research methods that both produce groups of scores to be compared: the experimental and the nonexperimental strategies. These two research methods use exactly the same statistics and they both demonstrate a relationship between two variables. The distinction between the two research strategies is how the relationship is interpreted. The results from an experiment allow a cause-and-effect explanation. For example, we can conclude that changes in one variable are responsible for causing differences in a second variable. A nonexperimental study does not permit a cause-and effect explanation. We can say that changes in one variable are accompanied by changes in a second variable, but we cannot say why. Each of the two research methods is discussed in the following sections.

The Experimental Method

One specific research method that involves comparing groups of scores is known as the *experimental method* or the *experimental research strategy.* The goal of an experimental study is to demonstrate a cause-and-effect relationship between two variables. Specifically, an experiment attempts to show that changing the value of one variable causes changes to occur in the second variable. To accomplish this goal, the experimental method has two characteristics that differentiate experiments from other types of research studies:

In more complex experiments, a researcher may systematically manipulate more than one variable and may observe more than one variable. Here we are considering the simplest case, in which only one variable is manipulated and only one variable is observed.

1. **Manipulation** The researcher manipulates one variable by changing its value from one level to another. In the Polman et al. (2008) experiment examining the effect of violence in video games on aggressive behavior (Figure 1.6), the researchers manipulate the amount of violence by giving one group of boys a violent game to play and giving the other group a nonviolent game. A second variable is observed (measured) to determine whether the manipulation causes changes to occur. In the Polman et al. (2008) experiment, aggressive behavior was measured.
2. **Control** The researcher must exercise control over the research situation to ensure that other, extraneous variables do not influence the relationship being examined. Control usually involves matching different groups as closely as possible on those variables that we don't want to manipulate.

To demonstrate these two characteristics, consider the Polman et al. (2008) study examining the effect of violence in video games on aggression (see Figure 1.6). To be able to say that the difference in aggressive behavior is caused by the amount of violence in the game, the researcher must rule out any other possible explanation for the difference. That is, any other variables that might affect aggressive behavior must be controlled. Two of the general categories of variables that researchers must consider:

1. **Environmental Variables** These are characteristics of the environment such as lighting, time of day, and weather conditions. A researcher must ensure that the

individuals in treatment A are tested in the same environment as the individuals in treatment B. Using the video game violence experiment (see Figure 1.6) as an example, suppose that the individuals in the nonviolent condition were all tested in the morning and the individuals in the violent condition were all tested in the evening. It would be impossible to determine if the results were due to the type of video game the children played or the time of day they were tested because an uncontrolled environmental variable (time of day) is allowed to vary with the treatment conditions. Whenever a research study allows more than one explanation for the results, the study is said to be *confounded* because it is impossible to reach an unambiguous conclusion.

According to APA convention, the term *participants* is used when referring to research with humans and the term *subjects* is used when referring to research with animals.

2. **Participant Variables** These are characteristics such as age, gender, motivation, and personality that vary from one individual to another. Because no two people (or animals) are identical, the individuals who participate in research studies will be different on a wide variety of participant variables. These differences, known as *individual differences*, are a part of every research study. Whenever an experiment compares different groups of participants (one group in treatment A and a different group in treatment B), the concern is that there may be consistent differences between groups for one or more participant variables. For the experiment shown in Figure 1.6, for example, the researchers would like to conclude that the violence in the video game causes a change in the participants' aggressive behavior. In the study, the participants in both conditions were 10-year-old boys. Suppose, however, that the participants in the violent video game condition, just by chance, had more children who were bullies. In this case, there is an alternative explanation for the difference in aggression that exists between the two groups. Specifically, the difference between groups may have been caused by the amount of violence in the game, but it also is possible that the difference was caused by preexisting differences between the groups. Again, this would produce a confounded experiment.

Researchers typically use three basic techniques to control other variables. First, the researcher could use *random assignment*, which means that each participant has an equal chance of being assigned to each of the treatment conditions. The goal of random assignment is to distribute the participant characteristics evenly between the two groups so that neither group is noticeably smarter (or older, or faster) than the other. Random assignment can also be used to control environmental variables. For example, participants could be assigned randomly for testing either in the morning or in the afternoon. A second technique for controlling variables is to use *matching* to ensure groups are equivalent in terms of participant variables and environmental variables. For example, the researcher could match groups by ensuring that each group has exactly 60% females and 40% males. Finally, the researcher can control variables by *holding them constant*. For example, in the video game violence study discussed earlier (Polman et al., 2008), the researchers used only 10-year-old boys as participants (holding age and gender constant). In this case the researchers can be certain that one group is not noticeably older or has a larger proportion of females than the other.

The matched-subject is a method to prevent preexisting participant differences between groups and is covered in Chapter 11.

> In the **experimental method**, one variable is manipulated while another variable is observed and measured. To establish a cause-and-effect relationship between the two variables, an experiment attempts to control all other variables to prevent them from influencing the results.
>
> The individuals in a research study differ on a variety of participant variables such as age, weight, skills, motivation, and personality. The differences from one participant to another are known as **individual differences**.

Terminology in the Experimental Method Specific names are used for the two variables that are studied by the experimental method. The variable that is manipulated by the experimenter is called the *independent variable*. It can be identified as the treatment conditions to which participants are assigned. For the example in Figure 1.6, the amount of violence in the video game is the independent variable. The variable that is observed and measured to obtain scores within each condition is the *dependent variable*. In Figure 1.6, the level of aggressive behavior is the dependent variable.

> The **independent variable** is the variable that is manipulated by the researcher. In behavioral research, the independent variable usually consists of the two (or more) treatment conditions to which subjects are exposed. The independent variable is manipulated *prior* to observing the dependent variable.
>
> The **dependent variable** is the one that is observed to assess the effect of the treatment. The dependent variable is the variable that is measured in the experiment and its value changes in a way that depends on the status of the independent variable.

An experimental study evaluates the relationship between two variables by manipulating one variable (the independent variable) and measuring one variable (the dependent variable). Note that in an experiment only one variable is actually measured. You should realize that this is different from a correlational study, in which all variables are measured and the data consist of at least two separate scores for each individual.

Control Conditions in an Experiment Often an experiment will include a condition in which the participants do not receive any experimental treatment. The scores from these individuals are then compared with scores from participants who do receive the treatment. The goal of this type of study is to demonstrate that the treatment has an effect by showing that the scores in the treatment condition are substantially different from the scores in the no-treatment condition. In this kind of research, the no-treatment condition is called the *control condition*, and the treatment condition is called the *experimental condition*.

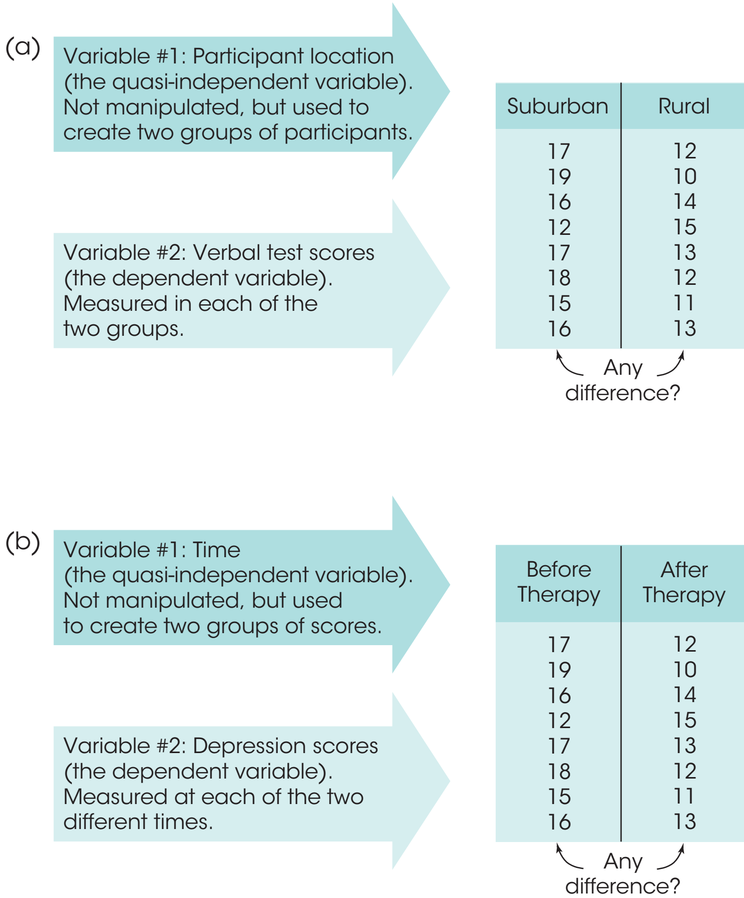
> Individuals in a **control condition** do not receive the experimental treatment. Instead, they either receive no treatment or they receive a neutral, placebo treatment. The purpose of a control condition is to provide a baseline for comparison with the experimental condition.
>
> Individuals in the **experimental condition** do receive the experimental treatment.

Note that the independent variable always consists of at least two values. (Something must have at least two different values before you can say that it is “variable.”) For the video game violence experiment (see Figure 1.6), the independent variable is the amount of violence in the video game. For an experiment with an experimental group and a control group, the independent variable is treatment versus no treatment.

■ Nonexperimental Methods: Nonequivalent Groups and Pre-Post Studies

In informal conversation, there is a tendency for people to use the term *experiment* to refer to any kind of research study. You should realize, however, that the term applies only to studies that satisfy the specific requirements outlined earlier. In particular, a real

FIGURE 1.7
Two examples of nonexperimental studies that involve comparing two groups of scores. In (a), the study uses two preexisting groups (suburban/rural) and measures a dependent variable (verbal scores) in each group. In (b), the study uses time (before/after) to define the two groups and measures a dependent variable (depression) in each group.



experiment must include manipulation of an independent variable and rigorous control of other, extraneous variables. As a result, there are a number of other research designs that are not true experiments but still examine the relationship between variables by comparing groups of scores. Two examples are shown in Figure 1.7 and are discussed in the following paragraphs. This type of research study is classified as nonexperimental.

The top part of Figure 1.7 shows an example of a *nonequivalent groups study* comparing third-grade students from suburban communities to those from rural communities. Notice that this study involves comparing two groups of scores (like an experiment). However, the researcher has no ability to control which participants go into which group—group assignment for the children is determined by where they live, not by the researcher. Because this type of research compares preexisting groups, the researcher cannot control the assignment of participants to groups and cannot ensure equivalent groups. Other examples of nonequivalent group studies include comparing 8-year-old children and 10-year-old children, people diagnosed with an eating disorder and those not diagnosed with a disorder, and comparing children from a single-parent home and those from a two-parent home. Because it is impossible to use techniques like random assignment to control participant variables and ensure equivalent groups, this type of research is not a true experiment.

Correlational studies are also examples of nonexperimental research. In this section, however, we are discussing nonexperimental studies that compare two or more groups of scores.

The bottom part of Figure 1.7 shows an example of a *pre-post study* comparing depression scores before therapy and after therapy. A pre-post study uses the passage of time (before/after) to create the groups of scores. In Figure 1.7 the two groups of scores are obtained by measuring the same variable (depression) twice for each participant; once before therapy and again after therapy. In a pre-post study, however, the researcher has no control over the passage of time. The "before" scores are always measured earlier than the "after" scores. Although a difference between the two groups of scores may be caused by the treatment, it is always possible that the scores simply change as time goes by. For example, the depression scores may decrease over time in the same way that the symptoms of a cold disappear over time. In a pre-post study the researcher also has no control over other variables that change with time. For example, the weather could change from dark and gloomy before therapy to bright and sunny after therapy. In this case, the depression scores could improve because of the weather and not because of the therapy. Because the researcher cannot control the passage of time or other variables related to time, this study is not a true experiment.

Terminology in Nonexperimental Research Although the two research studies shown in Figure 1.7 are not true experiments, you should notice that they produce the same kind of data that are found in an experiment (see Figure 1.6). In each case, one variable is used to create groups, and a second variable is measured to obtain scores within each group. In an experiment, the groups are created by manipulation of the independent variable, and the participants' scores are the dependent variable. The same terminology is often used to identify the two variables in nonexperimental studies. That is, the variable that is used to create groups is the independent variable and the scores are the dependent variable. For example, the top part of Figure 1.7, the child's location (suburban/rural), is the independent variable and the verbal test scores are the dependent variable. However, you should realize that location (suburban/rural) is not a true independent variable because it is not manipulated. For this reason, the "independent variable" in a nonexperimental study is often called a *quasi-independent variable.*

In a nonexperimental study, the "independent variable" that is used to create the different groups of scores is often called the **quasi-independent variable**.

LEARNING CHECK

LO8 **1.** Which of the following is most likely to be a purely correlational study?

a. One variable and one group
b. One variable and two groups
c. Two variables and one group
d. Two variables and two groups

LO8 **2.** A research study comparing alcohol use for college students in the United States and Canada reports that more Canadian students drink but American students drink more (Kuo, Adlaf, Lee, Gliksman, Demers, & Wechsler, 2002). What research design did this study use?

a. Correlational
b. Experimental
c. Nonexperimental
d. Noncorrelational

LO9 **3.** Stephens, Atkins, and Kingston (2009) found that participants were able to tolerate more pain when they shouted their favorite swear words over and over than when they shouted neutral words. For this study, what is the independent variable?

a. The amount of pain tolerated

b. The participants who shouted swear words

c. The participants who shouted neutral words

d. The kind of word shouted by the participants

ANSWERS **1. c** **2. c** **3. d**

1-4 Statistical Notation

LEARNING OBJECTIVES

10. Identify what is represented by each of the following symbols: X, Y, N, n, and Σ.

11. Perform calculations using summation notation and other mathematical operations following the correct order of operations.

The measurements obtained in research studies provide the data for statistical analysis. Most of the statistical analyses use the same general mathematical operations, notation, and basic arithmetic that you have learned during previous years of schooling. In case you are unsure of your mathematical skills, there is a mathematics review section in Appendix A at the back of this book. The appendix also includes a skills assessment exam (p. 570) to help you determine whether you need the basic mathematics review. In this section, we introduce some of the specialized notation that is used for statistical calculations. In later chapters, additional statistical notation is introduced as it is needed.

Scores

Quiz Scores
X
37
35
35
30
25
17
16

Height	Weight
X	Y
72	165
68	151
67	160
67	160
68	146
70	160
66	133

Measuring a variable in a research study yields a value or a score for each individual. Raw scores are the original, unchanged scores obtained in the study. Scores for a particular variable are typically represented by the letter X. For example, if performance in your statistics course is measured by tests and you obtain a 35 on the first test, then we could state that $X = 35$. A set of scores can be presented in a column that is headed by X. For example, a list of quiz scores from your class might be presented as shown in the margin (the single column on the left).

When observations are made for two variables, there will be two scores for each individual. The data can be presented as two lists labeled X and Y for the two variables. For example, observations for people's height in inches (variable X) and weight in pounds (variable Y) can be presented as shown in the double column in the margin. Each pair X, Y represents the observations made of a single participant.

The letter N is used to specify how many scores are in a set. An uppercase letter N identifies the number of scores in a population and a lowercase letter n identifies the number of scores in a sample. Throughout the remainder of the book you will notice that we often use notational differences to distinguish between samples and populations. For the height and weight data in the preceding table, $n = 7$ for both variables. Note that by using a lowercase letter n, we are implying that these data are a sample.

Summation Notation

Many of the computations required in statistics involve adding a set of scores. Because this procedure is used so frequently, a special notation is used to refer to the sum of a set of scores. The Greek letter *sigma*, or Σ, is used to stand for summation. The expression ΣX means to add all the scores for variable X. The summation sign Σ can be read as "the sum of." Thus, ΣX is read "the sum of the scores." For the following set of quiz scores, 10, 6, 7, 4,

$$\Sigma X = 27 \quad \text{and} \quad N = 4.$$

To use summation notation correctly, keep in mind the following two points:

1. The summation sign, Σ, is always followed by a symbol or mathematical expression. The symbol or expression identifies exactly which values are to be added. To compute ΣX, for example, the symbol following the summation sign is X, and the task is to find the sum of the X values. On the other hand, to compute $\Sigma(X - 1)$, the summation sign is followed by a relatively complex mathematical expression, so your first task is to calculate all the $(X - 1)$ values and then add those results.
2. The summation process is often included with several other mathematical operations, such as multiplication or squaring. To obtain the correct answer, it is essential that the different operations be done in the correct sequence. Following is a list showing the correct *order of operations* for performing mathematical operations. Most of this list should be familiar, but you should note that we have inserted the summation process as the fourth operation in the list.

More information on the order of operations for mathematics is available in the Math Review Appendix A, Section A.1.

Order of Mathematical Operations

1. Any calculation contained within parentheses is done first.
2. Squaring (or raising to other exponents) is done second.
3. Multiplying and/or dividing is done third. A series of multiplication and/or division operations should be done in order from left to right.
4. Summation using the Σ notation is done next.
5. Finally, any other addition and/or subtraction is done.

The following examples demonstrate how summation notation is used in most of the calculations and formulas we present in this book. Notice that whenever a calculation requires multiple steps, we use a computational table to help demonstrate the process. The table simply lists the original scores in the first column and then adds columns to show the results of each successive step. Notice that the first three operations in the order-of-operations list all create a new column in the computational table. When you get to summation (number 4 in the list), you simply add the values in the last column of your table to obtain the sum.

EXAMPLE 1.3

A set of four scores consists of values 3, 1, 7, and 4. We will compute ΣX, ΣX^2, and $(\Sigma X)^2$ for these scores. To help demonstrate the calculations, we will use a computational table showing the original scores (the X values) in the first column. Additional columns can then be added to show additional steps in the series of operations. You should notice that the first three operations in the list (parentheses, squaring, and multiplying) all create a new column of values. The last two operations, however, produce a single value corresponding to the sum.

X	X^2
3	9
1	1
7	49
4	16

The table to the left shows the original scores (the X values) and the squared scores (the X^2 values) that are needed to compute ΣX^2.

The first calculation, ΣX, does not include any parentheses, squaring, or multiplication, so we go directly to the summation operation. The X values are listed in the first column of the table, and we simply add the values in this column:

$$\Sigma X = 3 + 1 + 7 + 4 = 15$$

To compute ΣX^2, the correct order of operations is to square each score and then find the sum of the squared values. The computational table shows the original scores and the results obtained from squaring (the first step in the calculation). The second step is to find the sum of the squared values, so we simply add the numbers in the X^2 column:

$$\Sigma X^2 = 9 + 1 + 49 + 16 = 75$$

The final calculation, $(\Sigma X)^2$, includes parentheses, so the first step is to perform the calculation inside the parentheses. Thus, we first find ΣX and then square this sum. Earlier, we computed $\Sigma X = 15$, so

$$(\Sigma X)^2 = (15)^2 = 225$$

■

EXAMPLE 1.4 Use the same set of four scores from Example 1.3 and compute $\Sigma(X - 1)$ and $\Sigma(X - 1)^2$. The following computational table will help demonstrate the calculations.

X	$(X - 1)$	$(X - 1)^2$
3	2	4
1	0	0
7	6	36
4	3	9

The first column lists the original scores. A second column lists the $(X - 1)$ values, and a third column shows the $(X - 1)^2$ values.

To compute $\Sigma(X - 1)$, the first step is to perform the operation inside the parentheses. Thus, we begin by subtracting one point from each of the X values. The resulting values are listed in the middle column of the table. The next step is to add the $(X - 1)$ values, so we simply add the values in the middle column.

$$\Sigma(X - 1) = 2 + 0 + 6 + 3 = 11$$

The calculation of $\Sigma(X - 1)^2$ requires three steps. The first step (inside parentheses) is to subtract 1 point from each X value. The results from this step are shown in the middle column of the computational table. The second step is to square each of the $(X - 1)$ values. The results from this step are shown in the third column of the table. The final step is to add the $(X - 1)^2$ values, so we add the values in the third column to obtain

$$\Sigma(X - 1)^2 = 4 + 0 + 36 + 9 = 49$$

Notice that this calculation requires squaring before adding. A common mistake is to add the $(X - 1)$ values and then square the total. Be careful! ■

EXAMPLE 1.5 In both the preceding examples, and in many other situations, the summation operation is the last step in the calculation. According to the order of operations, parentheses, exponents, and multiplication all come before summation. However, there are situations in which extra addition and subtraction are completed after the summation. For this example, use the same scores that appeared in the previous two examples, and compute $\Sigma X - 1$.

With no parentheses, exponents, or multiplication, the first step is the summation. Thus, we begin by computing ΣX. Earlier we found $\Sigma X = 15$. The next step is to subtract one point from the total. For these data,

$$\Sigma X - 1 = 15 - 1 = 14$$

■

EXAMPLE 1.6

For this example, each individual has two scores. The first score is identified as X, and the second score is Y. With the help of the following computational table, compute ΣX, ΣY, $\Sigma X \Sigma Y$, and ΣXY.

Person	X	Y	XY
A	3	5	15
B	1	3	3
C	7	4	28
D	4	2	8

To find ΣX, simply add the values in the X column.

$$\Sigma X = 3 + 1 + 7 + 4 = 15$$

Similarly, ΣY is the sum of the Y values in the middle column.

$$\Sigma Y = 5 + 3 + 4 + 2 = 14$$

To find $\Sigma X \Sigma Y$ you must add the X values and add the Y values. Then you multiply these sums.

$$\Sigma X \Sigma Y = 15(14) = 210$$

To compute ΣXY, the first step is to multiply X times Y for each individual. The resulting products (XY values) are listed in the third column of the table. Finally, we add the products to obtain

$$\Sigma XY = 15 + 3 + 28 + 8 = 54$$

■

The following example is an opportunity for you to test your understanding of summation notation.

EXAMPLE 1.7

Calculate each value requested for the following scores: 5, 2, 4, 2

a. ΣX^2 **b.** $\Sigma(X + 1)$ **c.** $\Sigma(X + 1)^2$

You should obtain answers of 49, 17, and 79 for a, b, and c, respectively. Good luck. ■

LEARNING CHECK

LO10 1. What value is represented by the lowercase letter n?
- **a.** The number of scores in a population
- **b.** The number of scores in a sample
- **c.** The number of values to be added in a summation problem
- **d.** The number of steps in a summation problem

LO11 2. What is the value of $\Sigma(X - 2)$ for the following scores: 6, 2, 4, 2?
- **a.** 12
- **b.** 10
- **c.** 8
- **d.** 6

LO11 3. What is the first step in the calculation of $(\Sigma X)^2$?
- **a.** Square each score.
- **b.** Add the scores.
- **c.** Subtract 2 points from each score.
- **d.** Add the $X - 2$ values.

ANSWERS **1.** b **2.** d **3.** b