Ronet D. Bachman
Raymond Paternoster
Theodore H. Wilson

STATISTICS

for Criminology and Criminal Justice

FIFTH EDITION

Statistics for Criminology and Criminal Justice

Fifth Edition

This book is dedicated to Ray Paternoster, scholar and mentor extraordinaire, Emperor of Wyoming, and true Renaissance man. He is no longer on this planet, but he is always with us!

Sara Miller McCune founded SAGE Publishing in 1965 to support the dissemination of usable knowledge and educate a global community. SAGE publishes more than 1000 journals and over 600 new books each year, spanning a wide range of subject areas. Our growing selection of library products includes archives, data, case studies and video. SAGE remains majority owned by our founder and after her lifetime will become owned by a charitable trust that secures the company's continued independence.

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne

Statistics for Criminology and Criminal Justice

Fifth Edition

Ronet D. Bachman
University of Delaware

Raymond Paternoster

Theodore H. Wilson
University at Albany, State University of New York

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne

FOR INFORMATION:

SAGE Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

SAGE Publications Ltd.
1 Oliver's Yard
55 City Road
London EC1Y 1SP
United Kingdom

SAGE Publications India Pvt. Ltd.
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road, New Delhi 110 044
India

SAGE Publications Asia-Pacific Pte. Ltd.
18 Cross Street #10-10/11/12
China Square Central
Singapore 048423

Acquisitions Editor: Helen Salmon
Editorial Assistant: Kelsey Barkis
Production Editor: Rebecca Lee
Copy Editor: Tammy Giesmann
Typesetter: C&M Digitals (P) Ltd.
Indexer: Integra
Cover Designer: Rose Storey
Marketing Manager: Victoria Velasquez

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Names: Bachman, Ronet, author. | Paternoster, Raymond, author. | Wilson, Theodore H., author.

Title: Statistics for criminology and criminal justice / Ronet D. Bachman, University of Delaware, Raymond Paternoster, Theodore H. Wilson, The University at Albany, State University of New York.

Description: Fifth edition. | Los Angeles : SAGE, [2022] | Includes bibliographical references and index.

Identifiers: LCCN 2020042577 | ISBN 978-1-5443-7570-0 (paperback ; alk. paper) | ISBN 978-1-5443-7569-4 (epub) | ISBN 978-1-5443-7568-7 (epub) | ISBN 978-1-5443-7567-0 (pdf)

Subjects: LCSH: Criminal statistics.

Classification: LCC HV6208 .B33 2022 | DDC 364.01/5195—dc23
LC record available at https://lccn.loc.gov/2020042577

This book is printed on acid-free paper.

21 22 23 24 25 10 9 8 7 6 5 4 3 2 1

BRIEF CONTENTS

DETAILED CONTENTS

PREFACE

One of the most important aspects of teaching a statistics course is conveying to students the vital role that research and statistics play in the study of issues related to criminology and criminal justice. After years of teaching statistics courses, we have found that the best avenue for achieving this goal has been to link the teaching of "how to calculate and interpret statistics" with contemporary research examples from the field. By combining discussions of the "how to" in statistics with real data and research examples, students not only learn how to perform and understand statistical analyses but also to make the connection between how they are used and why they are so important.

In this new edition of *Statistics for Criminology and Criminal Justice* published by SAGE, our goal is to present a discussion of basic statistical procedures that is comprehensive in its coverage, yet accessible and readable for students. In view of this general goal, we have chosen to emphasize a practical approach to the use of statistics in research. We continue to stress the interpretation and understanding of statistical operations in answering research questions, be they theoretical or policy oriented in nature. Of course, this approach is at the expense of a detailed theoretical or mathematical treatment of statistics. Accordingly, we do not provide derivations of formulas, nor do we offer proofs of the underlying statistical theory behind the operations we present in this text. As you will see, however, we have not sacrificed statistical rigor.

Given the title, it is clear that we had the student majoring in criminology and criminal justice particularly in mind as a reader of this text. This can easily be seen in the nature of the research examples presented throughout the book. What are the causes of violence? What is the nature of hate crimes in the United States? Do different types of police patrolling activities affect rates of crime? These and many other research questions are examined in the examples provided in the book, which we believe not only makes the book more interesting to students but also makes the statistical material easier to understand and apply. If this book communicates the excitement of research and the importance of careful statistical analysis in research, then our endeavor has succeeded. We hope that students will enjoy learning how to investigate research questions related to criminal justice and criminology with statistics and that many will learn how to do some research of their own along the way.

In this edition, we continue to use our basic approach of describing each statistic's purpose and origins as we go. To facilitate learning, we present statistical formulas along with step-by-step instructions for calculation. The primary emphasis in our coverage of each statistical operation is on its interpretation and understanding. This edition updates all crime data and includes many new research examples. Each chapter sets up case studies from the research literature to highlight the concepts and statistical techniques under discussion. There are hand calculation practice problems at the end of each chapter that include examples from contemporary research in the field. This edition now includes Excel and Stata exercises along with SPSS exercises that correspond to the chapter material; these exercises use real data including subsets of data from the National Crime Victimization Survey, Monitoring the Future, the Youth Risk Behavior Survey, state-level crime data from the Uniform Crime Reports (UCR), and opinion data from the General Social Survey. In response to many reviewers requests, answers to these practice problems and computer output for all Excel, Stata, and IBM® SPSS® Statistics* exercises are available on the instructor's website, and the answers to odd questions are available to students in the back of the book.

*IBM® SPSS® Statistics was formerly called PASW® Statistics. SPSS is a registered trademark of International Business Machines Corporation.

ORGANIZATION OF THE BOOK

The book is organized sequentially into four parts. The first is titled "Univariate Analysis: Describing Variable Distributions" and begins with a basic discussion of research and data gathering. Chapters 1 and 2 discuss the research enterprise, sampling techniques, ways of presenting data, and levels of measurement. Chapter 3 offers an overview of interpreting data through the use of such graphical techniques as frequency distributions, pie charts, and bar graphs for qualitative data, as well as histograms, and frequency polygons for quantitative data. Chapter 4 provides an overview of measures of central tendency, and Chapter 5 discusses the various statistical techniques for measuring the variability of a variable, including the standard deviation as well as the exploratory data analysis technique of boxplots.

From this discussion of descriptive statistics, we move into the second section, "Making Inferences in Univariate Analysis: Generalizing From a Sample to the Population." Chapter 6 outlines the foundation of inferential statistics, probability theory, and sampling distributions (the normal distribution). In Chapter 6, the concept of hypothesis testing using the binomial distribution is also introduced. The remainder of the book concerns issues related to hypothesis testing and the search for a relationship between one or more independent variables and a dependent variable. Chapter 7 begins the journey into inferential statistics with confidence intervals. Chapter 8 reiterates the steps to formal hypothesis testing and focuses on testing the null hypothesis for one population mean. The steps to formal hypothesis testing are systematically repeated in each of the subsequent chapters.

The third section focuses on hypothesis testing using one independent variable to predict one dependent variable and is called "Bivariate Analysis: Relationships Between Two Variables." Chapter 9 is concerned with hypothesis testing when both independent and dependent variables are categorical using cross-tabulation and chi-square. In Chapter 10, you will examine hypothesis tests involving two population means or proportions, including tests for independent and matched groups. Chapter 11 discusses hypothesis testing involving three or more means using analysis of variance techniques. In Chapter 12, bivariate correlation and ordinary least-squares (OLS) regression analysis will be introduced. This chapter discusses the essential framework of linear regression, including the notion of "least squares," the importance of scatterplots, the regression line, and hypothesis tests with slopes and correlation coefficients.

The book concludes by highlighting the importance of controlling for other independent variables through "Multivariable Analysis: Predicting One Dependent Variable With Two or More Independent Variables." Chapter 13 extends OLS regression to two independent variables and one dependent variable. Chapter 14 provides a discussion of the essential components of logistic regression models and includes a discussion of multiple logistic regression analyses. Although logistic regression is seldom included in introductory statistics texts, these models have become so prominent in social science research that we felt their omission would have done a great disservice to those who want some degree of comprehensiveness in their first statistics course.

The fifth edition of *Statistics for Criminology and Criminal Justice* retains its distinctive feature of integrating current data and substantive research examples from the discipline that highlight how scholars selected the most appropriate statistical technique for their data. Examples from the literature are not simply dropped here and there to keep students' attention like other statistical texts in the field. Rather, each chapter presents a particular statistical method in the context of a substantive research story. This serves several purposes: it illustrates the process of research in the real world, it underscores why particular statistical techniques were selected over others, and it highlights the important role research and statistical analysis plays in policy decisions in our field. As such, this book's success is due in no small measure to the availability of so many excellent research examples in our discipline. The univariate chapters include all new data from the most recent publications from the Federal Bureau of Investigation and the Bureau of Justice Statistics. Other chapters that relied on real data have also been updated. For

example, Chapter 12 relies on the most recent data available to examine the bivariate relationships between state rates of murder and poverty, rates of robbery and rural population, and rates of robbery and divorce. In this way, students not only learn how to conduct appropriate statistical analyses, but also are simultaneously learning important substantive information related to the discipline.

LEARNING AIDS

Working together, the authors and editors have developed a format that makes *Statistics for Criminology and Criminal Justice* a readable, user-friendly text. In addition to all of the changes we have already mentioned, the Fifth Edition not only includes a host of new tables and figures to amplify text coverage, but also features the following student learning aids:

- Step-by-step lists and marginal key term and key formula boxes are included in every chapter to make mastery of statistical concepts and procedures easier.
- Each chapter closes with traditional practice problems to give students plenty of hands-on experience with important techniques, which incorporate research questions from contemporary published research from the discipline. Solutions to all end-of-chapter problems are also provided to instructors.
- Each chapter now includes Excel and Stata exercises along with the commands to obtain the output in addition to SPSS exercises that have been in previous editions. These provide students with the opportunity to obtain the statistics covered in each chapter using a computer software program of their choice. Output from all programs and all problems is also now provided on the accompanying website.

DIGITAL RESOURCES

A website to accompany the book at **edge.sagepub.com/bachmansccj5e** includes:

For instructors:

- Test bank, in Microsoft Word and LMS-ready versions
- Editable PowerPoint slides
- Answers to end-of-chapter problems from the book
- Software output for Excel, SPSS, and Stata organized by chapter
- Datasets for use with the book

For students:

- Datasets for use with the book
- Software output for Excel, SPSS, and Stata organized by chapter
- eFlashcards of Glossary key terms

ACKNOWLEDGMENTS

We must first acknowledge our heartfelt gratitude to our editor, Helen Salmon, whose hard work and guidance on this project are unrivaled. Her supervision of the project and meticulous attention to detail was herculean! We are also indebted to others on the SAGE team, including our editorial assistant, Kelsey Barkis, who provided invaluable pedagogical advice along with a very critical eye while shepherding the text through the publication process. Rebecca Lee, our production editor, kept all the trains moving on time through the publication process. And finally, our hats are off to Tammy Giesmann, who did a laborious and very impressive job of copy editing, Once again, despite three of us proofreading final drafts of chapters, many errors still remained!

We owe a huge debt of gratitude to those who provided thorough reviews and sage advice for this and earlier editions:

Viviana Andreescu
University of Louisville

Jeb A. Booth
Northeastern University

Bruce A. Carroll
Georgia Gwinnett College

Katherine A. Durante
Nevada State College

Kingsley Ejiogu
University of Maryland Eastern Shore

Gina A. Erickson
Hamline University

Sara Z. Evans
University of West Florida

David R. Forde
University of Alabama

Ni He
University of Texas-San Antonio

David Holleran
The College of New Jersey

Wesley G. Jennings
University of South Florida

Brian Johnson
University of Maryland

William E. Kelly
Auburn University

Brian A. Lawton
Sam Houston University

Alan Lehman
University of Maryland, College Park

Shelly A. McGrath
University of Alabama at Birmingham

Binneh Minteh
Rutgers University, School of Criminal Justice

Fawn T. Ngo
University of South Florida Sarasota/Manatee

Thomas Petee
Auburn University

Dan Powers
University of Texas

Andre Rosay
University of Alaska

Stephen M. Schnebly
Arizona State University

Christina DeJong Schwitzer
Michigan State University

Hanna Scott
University of Ontario Institute of Technology

Lenonore Simon
Temple University

Christopher J. Sullivan
University of Cincinnati

Gary Sweeten
Arizona State University

Christine Tartaro
The Richard Stockton College of New Jersey

Dale W. Willits
Washington State University

There also have been others who have used the book and have made corrections and invaluable suggestions; they include Marc Riedel, Jacques Press, Douglas Thomson, Bonnie Lewis, Alexis Durham, and especially Keith Marcus, who went through a previous edition with a microscope. (Keith, we're sorry we could not incorporate all of your suggestions into this edition!) In addition, our hats are off to two PhD students—Lin Liu (now an assistant professor at Florida International University) and Matthew Kijowski—for their meticulous proofreading under pressure to catch any errors in our statistical calculations. There were more than a few!

Of course, we continue to be indebted to the many students we have had an opportunity to teach and mentor over the years at both the undergraduate and graduate levels. In many respects, this book could not have been written without these ongoing and reciprocal teaching and learning experiences. To all of our students, past and future: You inspire us to become better teachers!

Ronet is indebted to an amazing circle of friends who endured graduate school together and continue to hold retreats one weekend of the year (29 years and counting!) for guidance, support, therapy, chocolate, and laughter: Dianne Carmody, Gerry King, Peggy Plass, and Barbara Wauchope. You are the most amazing women in the world, and I am so blessed to have you in my life. To Alex Alvarez, Michelle Meloy, and Lori Williams, my other kindred spirits, for their support and guidance. To my father, Ron, for his steadfast critical eye in all matters of life. And heart-filling gratitude and remembrance to my husband, Raymond Paternoster, to whom this book is dedicated. His amazing spirit and wisdom are with me always even though his body is not bound to this planet. Like always, his guidance continues to make me a better person. And to our son, John Bachman-Paternoster, and our dog, Leo, who bring me joy and laughter beyond my wildest dreams!

Teddy is lucky to have an incredible cast of supporters including family, friends, mentors, and compatriots from the trenches of graduate school. I am particularly grateful to my mother, Karen Wilson, for her steadfast support through all times including when I diverted from aerospace engineering to pursue my passions in criminology, criminal justice, and statistics. To the "goons" in the DC metro area for teaching me the need for balance in life among work, family, friends, and self. To my fiancé for her encouragement, support, love, and occasional smiles at my terrible jokes. And most importantly, to Raymond Paternoster, to whom this book is dedicated, for teaching me how to be a professor, a good man, and a scholar.

ABOUT THE AUTHORS

Ronet D. Bachman, PhD, is a professor in the Department of Sociology and Criminal Justice at the University of Delaware. She is coauthor of *The Practice of Research in Criminology and Criminal Justice* 6th ed. (with Russell Schutt), coauthor of *Violence: The Enduring Problem* and *Murder American Style* (with Alexander Alvarez), coeditor of *Explaining Crime and Criminology: Essays in Contemporary Criminal Theory* (with Raymond Paternoster), and author of *Death and Violence on the Reservation: Homicide, Suicide and Family Violence in Contemporary American Indian Communities*, as well as author/coauthor of numerous articles that examine the long-term patterns of offending and the etiology of violence, with a particular emphasis on women, the elderly, and minority populations. Her longest ongoing research project is a longitudinal study related to delinquency and crime prevention, but it has only one subject—her son, John Bachman-Paternoster.

Raymond Paternoster, PhD, to whom this book is dedicated, was a professor in the Department of Criminology and Criminal Justice at the University of Maryland. He received his BA in sociology at the University of Delaware, where he was introduced to criminology by Frank Scarpitti and obtained his PhD at Florida State University under the careful and dedicated tutelage of Gordon Waldo and Ted Chiricos. He is coauthor of *The Death Penalty: America's Experience with Capital Punishment.* In addition to his interest in statistics, he also pursued questions related to offender decision making and rational choice theory, desistance from crime, and capital punishment. He was as devoted to teaching and mentorship as he was to his scholarship.

Theodore (Teddy) H. Wilson, PhD, is an assistant professor in the School of Criminal Justice at the University at Albany, SUNY. He received his BA, MA, and PhD in criminology and criminal justice at the University of Maryland under the careful instruction and mentoring of David Maimon, Thomas Loughran, and Raymond Paternoster. Joining his passion for statistics, he studies offender decision making, courts and sentencing, and courtroom actor decision making. His latest funded project is directed at exploring and explaining racial disparities in punishment for illegal gun possession in New York.

THE IMPORTANCE OF STATISTICS IN THE CRIMINOLOGICAL SCIENCES OR WHY DO I HAVE TO LEARN THIS STUFF?

You gain strength, courage, and confidence by every experience in which you really stop to look fear in the face.

—Eleanor Roosevelt

Do not worry about your difficulties in Mathematics. I can assure you mine are still greater.

—Albert Einstein

Learning Objectives

1. Describe the role statistical analyses play in criminological and criminal justice research.
2. Identify the difference between a sample and a population.
3. Explain the purpose of probability sampling techniques.
4. Define the different types of probability and nonprobability samples.
5. State the difference between descriptive and inferential statistics.
6. Specify the different types of validity in research.

INTRODUCTION

Most of you reading this book are probably taking a course in statistics because it is required to graduate, not because you were seeking a little adventure and thought it would be fun. Nor are you taking the course because there is something missing in your life and, thus, you think the study of statistics is necessary to make you intellectually "well rounded." At least this has been our experience when teaching statistics courses. Everyone who has taught a statistics course has probably heard the litany of sorrows expressed by their students at the beginning of the course—the "wailing and gnashing of teeth." "Oh, I have been putting this off for so long—I dreaded having to take this." "I have a mental block when it comes to math—I haven't had any math courses since high school." "Why do I have to learn this, I'm never going to use it?"

There are those fortunate few for whom math comes easy, but the rest of us experience apprehension and anxiety when approaching our first statistics course. Psychologists, however, are quick to tell us that what we most often fear is not real—it is merely our mind imagining the worst possible scenario. FEAR has been described as an

acronym for False Expectations Appearing Real. In fact, long ago it was Aristotle who said, "Fear is pain arising from anticipation." But then, this may not comfort you either because it is not Aristotle who is taking the course—it's you!

Although it is impossible for us to allay all of the fear and apprehension you may be experiencing right now, it may help to know that virtually everyone can and will make it through this course, even those of you who have trouble counting change. This is not a guarantee, and we are not saying it will be easy, but it can be done with hard work. We have found that persistence and tenacity can overcome even the most extreme mathematical handicaps. Those of you who are particularly rusty with your math, and those of you who just want a quick confidence builder, should refer to Appendix A at the back of this book. Appendix A reviews some basic math lessons. Our book also includes practice problems and, more important, the answers to those problems. After teaching this course for over two decades, we have found that every student who puts forth effort and time can pass the course! Our chapters are designed to provide step-by-step instructions for calculating the statistics with real criminal justice data and case studies so you will not only learn about statistics but also a little about research going on in our discipline.

Another incentive to learning this material is that careers related to criminology and criminal justice are increasingly becoming data driven. Understanding how to organize data and interpret statistics will be a tremendous asset to you, no matter what direction you plan to take in your career. Virtually every job application, as well as applications to graduate school and law school, now asks you about your data analysis skills. We now exist in a world where programs to organize and manipulate data are everywhere. The quotes from former students at the beginning of each chapter are testimony to this.

We hope that after this course you will be able to understand and manipulate statistics for yourself and that you will be a knowledgeable consumer of the statistical material you are confronted with daily. In addition to the mathematical skills required to compute statistics, we also hope to leave you with an understanding of what different statistical tests or operations can and cannot do, and what they do and do not tell us about a given problem. The foundations for the statistics presented in this book are derived from complicated mathematical theory. You will be glad to know that it is *not* the purpose of this book to provide you with the proofs necessary to substantiate this body of theory.

In this book, we provide you with two basic types of knowledge: (1) knowledge about the basic mathematical foundations of each statistic, as well as the ability to manipulate and conduct statistical analysis for your own research, and (2) an ability to interpret the results of statistical analysis and to apply these results to the real world. We want you, then, to have the skills to both calculate and comprehend social statistics. These two purposes are not mutually exclusive but related. We think that the ability to carry out the mathematical manipulations of a formula and come up with a statistical result is almost worthless unless you can interpret this result and give it meaning. Therefore, information about the mechanics of conducting statistical tests and information about interpreting the results of these tests will be emphasized equally throughout this text.

Learning about statistics for perhaps the first time does not mean that you will always have to calculate your statistics by hand with the assistance of only a calculator. Most, if not all, researchers do their statistical analyses with a computer and software programs. Many useful and "user-friendly" statistical software programs are available, including SPSS, SAS, Stata, R, Excel, and Minitab. Because learning to conduct statistical analyses with a computer is such an essential task to master, we provide a discussion of the computer software program SPSS on the student website for this book, along with data sets that can be downloaded. We have also included SPSS data analysis exercises at the end of each chapter; however, you can easily use these exercises for virtually any other statistical software program including the spreadsheet program Excel.

You may be wondering why you have to learn statistics and how to calculate them by hand if you can avoid all of this by using a computer. First, we believe it is important for you to understand exactly what it is the computer is doing when it is calculating statistics.

Without this knowledge, you may get results, but you will have no understanding of the logic behind the computer's output and little comprehension of how those results were obtained. This is not a good way to learn statistics; in fact, it is not really learning statistics at all. Without a firm foundation in the basics of statistics, you will have no real knowledge of what to request of your computer or how to recognize an incorrect result. Computer programs provide results even if those results are wrong. Despite its talent, the computer is actually fairly stupid; it has no ability to determine whether what it is told to do is correct—it will do pretty much anything it is asked, and it will calculate and spit out virtually anything you want it to, correct or not. The adage "garbage in, garbage out" is appropriate here. Moreover, a computer program won't interpret the results. That is your responsibility!

SETTING THE STAGE FOR STATISTICAL INQUIRY

Before we become more familiar with statistics in the upcoming chapters, we first want to set the stage for statistical inquiry. The data we use in criminology are derived from many different sources: from official government agency data such as the Federal Bureau of Investigation's (FBI) Uniform Crime Reports; from social surveys conducted by the government (the Bureau of Justice Statistics' National Crime Victimization Survey), ourselves, or other researchers; from experiments; from direct observation, as either a participant observer or an unobtrusive observer; or from a content analysis of existing images (historical or contemporary), such as newspaper articles or films. As you can see, the research methods we employ are very diverse.

Criminological researchers often conduct "secondary data analysis" (Riedel, 2012), which, simply put, means reanalyzing data that already exist. These data usually come from one of two places: Either they are official data collected by local, state, and federal agencies (e.g., rates of crime reported to police, information on incarcerated offenders from state correctional authorities, or adjudication data from the courts), or they are data collected from surveys sponsored by government agencies or conducted by other researchers. Virtually all of these data collected by government agencies and a great deal of survey data collected by independent researchers are made available to the public through the Inter-University Consortium for Political and Social Research (ICPSR), which is located at the University of Michigan.

The ICPSR maintains and provides access to a vast archive of criminological data for research and instruction, and it offers training in quantitative methods to facilitate effective data use. For example, data available online at ICPSR include the Supplementary Homicide Reports (SHR) provided by the U.S. Department of Justice, which contain information for each homicide from police reports, including such details as the relationship between victims and offenders, use of weapons, and other characteristics of victims and offenders; survey data from the National Crime Victimization Survey (NCVS), which interviews a sample of U.S. household residents to determine their experiences with both property and violent crime, regardless of whether the crimes were reported to police or anyone else; survey data from samples of jail and prison inmates; and survey data from the National Opinion Survey of Crime and Justice, which asked adults for their opinion about a wide range of criminal justice issues. These are just a few examples of the immense archive of data made available at the ICPSR. Take a look at what is available by going on the website: www.icpsr.umich.edu.

THE ROLE OF STATISTICAL METHODS IN CRIMINOLOGY AND CRIMINAL JUSTICE

Over the past few decades, statistics and numerical summaries of phenomena such as crime rates have increasingly been used to document how "well" or "poorly" a society is doing. For example, cities and states are described as relatively safe or unsafe depending on their respective

levels of violent crime, and age groups are frequently monitored and compared with previous generations to determine their relative levels of deviancy based on criteria such as their drug and alcohol use. Other criminal justice-related phenomena like the number of people under correctional supervision are also monitored closely.

Research and statistics are important in our discipline because they enable us to monitor phenomena over time and across geographic locations, and they allow us to determine relationships between phenomena. Of course, we make conclusions about the relationships between phenomena every day, but these conclusions are most often based on biased perceptions and selective personal experiences.

Science: A set of logical, systematic, documented methods for investigating nature and natural processes; the knowledge produced by these investigations.

In criminological research, we rely on scientific methods, including statistics, to help us perform these tasks. **Science** relies on logical and systematic methods to answer questions, and it does so in a way that allows others to inspect and evaluate its methods. In the realm of criminological research, these methods are not so unusual. They involve asking questions, observing behavior, and counting people, all of which we often do in our everyday lives. The difference is that researchers develop, refine, apply, and report their understanding of the social world more systematically.

CASE STUDY

Youth Violence

The population of the United States all too frequently mourns the deaths of young innocent lives taken in school shootings. The deadliest elementary school shooting to date took place on December 14, 2012, when a 20-year-old man named Adam Lanza walked into an elementary school in Newtown, Connecticut, armed with several semiautomatic weapons and killed 20 children and 6 adults. On April 16, 2007, Cho Seung-Hui perpetrated the deadliest college mass shooting by killing 32 students, faculty, and staff and left over 30 others injured on the campus of Virginia Tech in Blacksburg, Virginia. Cho was armed with two semiautomatic handguns that he had legally purchased and a vest filled with ammunition. As police were closing in on the scene, he killed himself. The deadliest high school shooting occurred on February 4, 2018, in Parkland, Florida, at Marjory Stoneman Douglas High School. Nikolas Cruz, a former student at the school, killed 17 and injured another 17 before fleeing. Cruz is currently awaiting trial in Florida.

None of these mass murderers were typical terrorists, and each of these incidents caused a media frenzy. Headlines such as "The School Violence Crisis" and "School Crime Epidemic" were plastered across national newspapers and weekly news journals. Unfortunately, the media play a large role in how we perceive both problems and solutions. What are your perceptions of violence committed by youth, and how did you acquire them? What do you believe are the causes of youth violence? Many (frequently conflicting) factors have been blamed for youth violence in American society, including the easy availability of guns, the lack of guns in classrooms for protection, the use of weapons in movies and television, the moral decay of our nation, poor parenting, unaware teachers, school and class size, racial prejudice, teenage alienation, the Internet and the World Wide Web, anti-Semitism, violent video games, rap and rock music, and the list goes on.

Of course, youth violence is not a new phenomenon in the United States. It has always been a popular topic of social science research and the popular press. Predictably, whenever a phenomenon is perceived as an epidemic, numerous explanations emerge to explain it. Unfortunately, most of these explanations are based on the media and popular culture, not on empirical research. Unlike the anecdotal information floating around in the mass media, social scientists interested in this phenomenon have amassed a substantial body of findings

that have refined knowledge about the factors related to the problem of gun violence, and some of this knowledge is being used to shape social policy. Research that relies on statistical analysis generally falls into three categories of purposes for social scientific research: Descriptive, Explanatory, and Evaluation.

Descriptive Research

Defining and describing social phenomena of interest is a part of almost any research investigation, but **descriptive research** is the primary focus of many studies of youth crime and violence. Some of the central questions used in descriptive studies are as follows: "How many people are victims of youth violence?" "How many youth are offenders?" "What are the most common crimes committed by youthful offenders?" and "How many youth are arrested and incarcerated each year for crime?"

Descriptive research: Research in which phenomena are defined and described.

Police reports: Data used to measure crime based on incidents that become known to police departments.

Uniform Crime Reports (UCR): Official reports about crime incidents that are reported to police departments across the United States and then voluntarily reported to the Federal Bureau of Investigation (FBI), which compiles them for statistics purposes.

National Incident-Based Reporting System (NIBRS): Official reports about crime incidents that are reported to police departments across the United States and then voluntarily reported to the Federal Bureau of Investigation (FBI), which compiles them for statistics purposes. This system is slowly replacing the older UCR program.

Surveys: Research method used to measure the prevalence of behavior, attitudes, or any other phenomenon by asking a sample of people to fill out a questionnaire either in person, through the mail or Internet, or on the telephone.

CASE STUDY

How Prevalent Is Youth Violence?

Police reports: One of the most enduring sources of information on lethal violence in the United States is the FBI's SHR. Data measuring the prevalence of nonlethal forms of violence such as robbery and assaults are a bit more complicated. How do we know how many young people assault victims each year? People who report their victimizations to police represent one avenue for these calculations. The FBI compiles these numbers in its **Uniform Crime Reports (UCR)** system, which is slowly being replaced by the **National Incident-Based Reporting System (NIBRS)**. Both of these data sources rely on state, county, and city law enforcement agencies across the United States to participate voluntarily in the reporting program. Can you imagine why relying on these data sources may be problematic for estimating prevalence rates of violent victimizations? If victimizations are never reported to police, they are not counted. This is especially problematic for victimizations between people who know each other and other offenses like rape in which only a fraction of incidents are ever reported to police.

Surveys: Many, if not most, social scientists believe the best way to determine the magnitude of violent victimization is through random sample surveys. This basically means randomly selecting individuals in the population of interest and asking them about their victimization experiences via a mailed or Internet, telephone, or in-person questionnaire. The only ongoing survey to do this on an annual basis is the NCVS, which is sponsored by the U.S. Department of Justice's Bureau of Justice Statistics. Among other questions, the NCVS asks questions like, "Has anyone attacked or threatened you with a weapon, for instance, a gun or knife; by something thrown, such as a rock or bottle, include any grabbing, punching, or choking?" Estimates indicate that youth aged 12 to 24 years all have the highest rates of violent victimization, which have been declining steadily since the highs witnessed in the early 1990s, despite the recent increases observed in homicide rates for this age group in some locations.

Another large research survey that estimates the magnitude of youth violence (along with other risk-taking behavior such as taking drugs and smoking) is called the Youth Risk Behavior Survey (YRBS), which has been conducted every two years in the United States since 1990. Respondents to this survey are a national sample of approximately 16,000 high-school students in grades 9 through 12. To measure the extent of youth violence, students are asked a number of questions, including the following: "During the past 12 months, how many times were you in a physical fight?" "During the past 12 months, how many times were you in a physical

fight in which you were injured and had to be seen by a doctor or nurse?" "During the past 12 months, how many times were you in a physical fight on school property?" and "During the past 12 months, how many times did someone threaten or injure you with a gun, knife, or club on school property?"

Of course, another way to measure violence would be to ask respondents about their offending behaviors. Some surveys do this, including the Rochester Youth Development Study (RYDS). The RYDS sample consists of 1,000 students who were in the seventh and eighth grades in the Rochester, New York, public schools during the spring semester of the 1988 school year. This project has interviewed the original respondents at 12 different times including the last interview that took place in 1997 when respondents were in their early 20s (Thornberry, Krohn, Lizotte, & Bushway, 2008). As you can imagine, respondents are typically more reluctant to reveal offending behavior compared with their victimization experiences. However, these surveys have been a useful tool for examining the factors related to violent offending and other delinquency. We should also point out that although this discussion has been specific to violence, the measures we have discussed in this section, along with their strengths and weaknesses, apply to measuring all crime in general.

Explanatory Research

Explanatory research: Research that seeks to identify causes and/or effects of social phenomena.

Dependent variable: Variable that is expected to change or vary depending on the variation in the independent variable.

Independent variable: Variable that is expected to cause or lead to variation or change in the dependent variable.

Many people consider explanation to be the premier goal of any science. **Explanatory research** seeks to identify the causes and effects of social phenomena, to predict how one phenomenon will change or vary in response to variation in some other phenomenon. Researchers adopted explanation as a goal when they began to ask such questions as "Are kids who participate in after school activities less likely to engage in delinquency?" and "Does the unemployment rate influence the frequency of youth crime?" In explanatory research, studies are often interested in explaining a **dependent variable** by using one or more independent variables. In research, the dependent variable is expected to vary or change depending on variation or change in the **independent variable**. One of our students came up with this way to describe it, "The dependent variable is dependent on the level or change in the independent variable, whereas, the independent variable is not dependent, but independent!" In this causal type of explanation, the independent variable is the cause and the dependent variable the effect.

CASE STUDY

What Factors Are Related to Youth Delinquency and Violence?

When we move from description to explanation, we want to understand the direct relationship between two or more things. Does x (the independent variable) explain y (the dependent variable) or if x happens, is y also likely to occur? What are some of the factors related to youth violence? Nathalie Fontaine and her colleagues Fontaine, Brendgen, Vitaro, and Tremblay (2016) were interested in how several factors including parental supervision and attachment to school affected the probability of adolescents engaging in violent behavior. They used a longitudinal data set collected in Montreal, Canada, which followed boys from kindergarten until they were 17 years old. By following this sample of boys over time, the researchers could determine that parental supervision and attachments in school came before the violent offending, which is extremely important when attempting to determine factors that predict violence.

Theory: Logically interrelated set of propositions about empirical reality that can be tested.

Testing hypotheses generated from theory is often a goal of explanatory research. A **theory** is a logically interrelated set of propositions about empirical reality. Examples of criminological

theories include social learning theory, general strain theory, social disorganization theory, and routine activities theory. A **hypothesis** is simply a tentative statement about empirical reality, involving a relationship between two or more variables.

Hypothesis: Tentative statement about empirical reality, involving the relationship between two or more variables.

Social control theory contends that conformity to the rules of society is produced and maintained by the ties individuals have to different things including family, friends, schools, jobs, and so on. These ties or bonds serve to increase the costs of delinquency and crime, and as such, serve to control individual offending. Fontaine and her colleagues (2016) measured two types of social control: parental attachment and school engagement (independent variables) that they hypothesized would be related to self-reported violent offending (dependent variable). Results indicated that boys who had greater parental supervision and school engagement were less likely to engage in violent delinquency compared with their less supervised and engaged counterparts. In fact, while boys who had been aggressive as children were more likely to be violent as adolescents, the relationship between childhood and adolescent violence was virtually eliminated for those boys who had high levels of parental supervision and school engagement.

Evaluation Research

Evaluation research seeks to determine the effects of a social program or other types of intervention. It is a type of explanatory research because it deals with cause and effect. However, evaluation research differs from other forms of explanatory research because evaluation research considers the implementation and effects of social policies and programs. These issues may not be relevant in other types of explanatory research.

Evaluation research: Research about social programs or interventions.

Evaluation research is a type of explanatory research, but instead of testing theory, it is most often used to determine whether an implemented program or policy had the intended outcome. To reduce violence and create a safer atmosphere at schools across the country, literally thousands of schools have adopted some form of violence prevention training. These programs generally provide cognitive-behavioral and social skills training on various topics using a variety of methods. Such programs are commonly referred to as conflict resolution and peer mediation training. Many of these prevention programs are designed to improve interpersonal problem-solving skills among children and adolescents by training children in cognitive processing, such as identifying interpersonal problems and generating nonaggressive solutions. Despite the millions of dollars being paid by school districts for such violence and bullying prevention programs, you may be surprised to learn that very few of these programs have been evaluated scientifically. That is, we know very little about whether they actually prevent bullying and violence.

CASE STUDY

How Effective Are School Bullying and Violence Prevention Programs?

Randomized control trials (RCTs), otherwise known as **true experimental designs**, are the gold standard in science to determine whether there is true causality between and independent and dependent variable. In the case of violence prevention programs, the program would be the independent variable that is assumed to affect bullying or violent behavior. RCTs randomly assign individuals to either receive the treatment or participate in the program, which is called the experimental group. Those who are not randomly assigned to the experimental group do not receive the new program and are called the control group. Random assignment to the two

Randomized control trial (RCT) or true experimental design: When two groups are randomly assigned with one group receiving the treatment or program (experimental group) while the other group (control group) does not. After the program or treatment, a post-test determines whether there is a change in the experimental group.

groups is the key because in this way, researchers can assume that the two groups are expected to be equivalent except for one group receiving a new treatment or program.

There are several anti-bullying programs being marketed both in the United States and in other countries, including the *Olweus Bullying Prevention Program*, *Steps to Respect*, *Restorative Whole-School Approach*, to name a few. Many of these programs include elements intended to increase staff supervision to prevent bullying and to increase the emotional intelligence of students to increase empathy and the ability to resolve conflicts without violence. Both of these mechanisms are hypothesized to decrease rates of bullying in schools. But do they work?

Gaffney, Ttofi, and Farrington (2019) reviewed studies that have evaluated these prevention programs since 2009. They performed a meta-analysis, which is actually a complicated statistical analysis that determines the average effect of similar programs on an outcome, which in this case was bullying behavior. They used the Centers for Disease Control and Prevention definition of bullying that includes three elements: (1) an intention to harm; (2) repetitive in nature; and (3) a clear power imbalance between perpetrator and victim. Without getting too deep into the statistical nitty-gritty, the researchers determined that the average effect of the bullying prevention programs served to decrease bullying perpetration by around 19%.

POPULATIONS AND SAMPLES

The words "population" and "sample" should already have some meaning to you. When you think of a population, you probably think of the population of some locality such as the United States, or the city or state in which you reside, or the university or college you attend. As with most social science research, samples in criminology consist of samples at different units of analysis including countries, states, cities, neighborhoods, prisons, schools, individuals, etc. Since it is too difficult, too costly, and sometimes impossible to get information on the entire population of interest, we must often solicit the information of interest from samples. **Samples** are simply subsets of a larger **population**.

Sample: Subset of the population that a researcher must often use to make generalizations about the larger population.

Population: Larger set of cases or aggregate number of people that a researcher is actually interested in or wishes to know something about.

Most official statistics collected by the U.S. government are derived from information obtained from samples, not from the entire population (the U.S. Census taken every 10 years is an exception). For example, the NCVS is a survey used to obtain information on the incidence and characteristics of criminal victimization in the United States based on a sample of the U.S. population. Every year, the NCVS interviews more than 100,000 individuals aged 12 years or older to solicit information on their experiences with victimization that were both reported and unreported to the police. Essentially, professional interviewers ask persons who are selected into the sample if they were the victim of a crime in the past six months, regardless of whether this victimization was reported to police.

You may be thinking right now, "Well, what if I am only interested in a small population?" Good question! Let's say we were interested in finding out about job-related stress experienced by law enforcement officers in your state. Although it would be easier to contact every individual in this population compared with every U.S. citizen, it would still be extremely difficult and costly to obtain information from every law enforcement officer, even within one state. In fact, in almost all instances, we have to settle for a sample derived from the population of interest rather than study the full population. For this reason, the "population" usually remains an unknown entity whose characteristics we can only estimate. The **generalizability** of a study is the extent to which it can be used to inform us about persons, places, or events that were *not* studied.

Generalizability: Extent to which information from a sample can be used to inform us about persons, places, or events that were not studied in the entire population from which the sample was taken.

We usually make a generalization about the characteristics of a population by using information we have from a sample; that is, we make inferences from our sample data to the

population. Because the purpose of sampling is to make these generalizations, we must be very meticulous when selecting our sample. The primary goal of sampling is to make sure that the sample we select is actually representative of the population we are estimating and want to generalize to. Think about this for a minute. What is representative? Generally, if the characteristics of a sample (e.g., age, race/ethnicity, and gender) look similar to the characteristics of the population, the sample is said to be representative. For example, if you were interested in estimating the proportion of the population that favors the death penalty, then to be representative, your sample should contain about 50% men and 50% women because that is the makeup of the U.S. population. It also should contain about 85% whites and 15% nonwhites because that is the makeup of the U.S. population. If your sample included a disproportionately high number of males or nonwhites, it would be unrepresentative. If, on the other hand, your target population was individuals older than 65 years of age, your sample should have a somewhat different gender distribution. To reflect the gender distribution of all individuals in the United States older than 65, a sample would have to contain approximately 60% women and 40% men since this is the gender distribution of all individuals older than age 65 in the United States as defined by the Census Bureau.

In sum, the primary question of interest in sample generalizability is as follows: *Can findings from a sample be generalized to the population from which the sample was drawn?* Sample generalizability depends on sample quality, which is determined by the amount of **sampling error** present in your sample. Sampling error can generally be defined as the difference between the sample estimate and the population value that you are estimating. The larger the sampling error, the less representative the sample and, as a result, the less generalizable the findings are to the population.

Sampling error: The difference between a sample estimate (called a sample statistic) and the population value it is estimating (called a population parameter).

With a few special exceptions, a good sample should be representative of the larger population from which it was drawn. A representative sample looks like the population from which it was selected in all respects that are relevant to a particular study. In an unrepresentative sample, some characteristics are overrepresented and/or some characteristics may be underrepresented. Various procedures can be used to obtain a sample; these range from the simple to the complex as we will see next.

HOW DO WE OBTAIN A SAMPLE?

From the previous discussion, it should be apparent that accuracy is one of the primary problems we face when generalizing information obtained from a sample to a population. How accurately does our sample reflect the true population? This question is inherent in any inquiry because with any sample we represent only a part—and sometimes a small part—of the entire population. The goal in obtaining or selecting a sample, then, is to select it in a way that increases the chances of this sample being representative of the entire population.

One of the most important distinctions made about samples is whether they are based on a probability or nonprobability sampling method. Sampling methods that allow us to know in advance how likely it is that any element of a population will be selected for the sample are **probability sampling methods**. Sampling methods that do not let us know the likelihood in advance are **nonprobability sampling methods**.

Probability sampling methods: These methods rely on random selection or chance and allow us to know in advance how likely it is that any element of a population is selected for the sample.

Nonprobability sampling methods: These methods are not based on random selection and do not allow us to know in advance the likelihood of any element of a population being selected for the sample.

The fundamental aspect of probability sampling is **random selection**. When a sample is randomly selected from the population, this means every element of the population (e.g., individual, school, or city) has a known, equal, and independent chance of being selected for the sample. All probability sampling methods rely on a random selection procedure.

Random selection: The fundamental aspect of probability sampling. The essential characteristic of random selection is that every element of the population has a known and independent chance of being selected for the sample.

Probability sampling techniques not only serve to minimize any potential bias we may have when selecting a sample, but also they allow us to gain access to probability theory in our data analysis, which you will learn more about later in this text. This body of mathematical theory allows us to estimate more accurately the degree of error we have when generalizing

results obtained from known sample statistics to unknown population parameters. But don't worry about probability theory now. For now, let's examine some of the most common types of probability samples used in research.

Flipping a coin and rolling a set of dice are the typical examples used to characterize random selection. When you flip a coin, you have the same chance of obtaining a head as you do of obtaining a tail: one out of two. Similarly, when rolling a die, you have the same probability of rolling a 2 as you do of rolling a 6: one out of six. In criminology, researchers generally use random numbers tables, such as Table B.1 in Appendix B, or other computer-generated random selection programs to select a sample. Because they are based on random selection, probability sampling methods have no systematic bias; nothing but chance determines which elements are included in the sample. As a result, our sample also is more likely to be representative of the entire population. When the goal is to generalize your findings to a larger population, it is this characteristic that makes probability samples more desirable than nonprobability samples. Using probability sampling techniques serves to avoid any potential bias we might introduce if we selected a sample ourselves.

PROBABILITY SAMPLING TECHNIQUES

Simple Random Samples

Simple random sample: Method of sampling in which every sample element is selected only on the basis of chance through a random process.

Perhaps the most common type of probability sample to use when we want to generalize information obtained from the sample to a larger population is called a **simple random sample**. Simple random sampling requires a procedure that generates numbers or identifies cases of the population for selection strictly on the basis of chance. The key aspect of a simple random sample is random selection. As we stated earlier, random selection ensures that every element in the population has a known, equal, and independent chance of being selected for the sample. If an element of the population is selected into the sample, true simple random sampling is done by replacing that element back into the population so that, once again, there is an equal and independent chance of every element being selected. This is called sampling with replacement. However, if your sample represents a very small percentage of a large population (say, less than 4%), sampling with and without replacement generally produce equivalent results.

Random digit dialing (RDD): Random dialing by a machine of numbers within designated phone prefixes, which creates a random sample for phone surveys.

Organizations that conduct large telephone surveys often draw random samples with an automated procedure called **random digit dialing (RDD)**. In this process, a computer dials random numbers within the phone prefixes corresponding to the area in which the survey is to be conducted. Random digit dialing is particularly useful when a sampling frame is not available. The researcher simply replaces any inappropriate numbers, such as those numbers that are no longer in service or numbers for businesses, with the next randomly generated phone number. Many surveys rely on this method and use both numbers for land lines and cell phones (Bachman & Schutt, 2017). For example, National Intimate Partner Violence and Sexual Victimization Surveys sponsored by The Centers for Disease Control and Prevention selects a random sample of adult males and females residing in the United States by using the RDD sampling technique.

Systematic Random Samples

Systematic random sampling: Method of sampling in which sample elements are selected from a list or from sequential files, with every kth element being selected after the first element is selected randomly within the first interval.

Simple random sampling is easy to do if your population is organized in a list, such as from a phone book, registered voters list, court docket, or membership list. We can make the process of simple random selection discussed earlier a little less time-consuming by systematically sampling the cases. In **systematic random sampling**, we select the first element into the sample randomly, but instead of continuing with this random selection, we *systematically* choose the rest of the sample. The general rule for systematic random sampling is to begin with a single element (any number selected randomly within the first interval, say the 10th) in the population

and then proceed to select the sample by choosing every *k*th element thereafter (say, every 12th element after the 10th). The first element is the only element that is truly selected at random. The starting element can be selected from a random numbers table or by some other random method. Systematic random sampling eliminates the process of deriving a new random number for every element selected, thus, saving time.

For systematic sampling procedures to approximate a simple random sample, the population list must be truly random, not ordered. For example, we could not have a list of convicted felons ordered by offense type, age, or some other characteristic. If the list is ordered in any way, this will add bias to the sampling process, and the resulting sample is not likely to be representative of the population. In virtually all other situations, systematic random sampling yields what is essentially a simple random sample.

Multistage Cluster Samples

There are often times when we do not have the luxury of a population list but still want to collect a random sample. Suppose, for example, we wanted to obtain a sample from the entire U.S. population. Would there be a list of the entire population available? Well, there are telephone books that list residents of various locales who have telephones; there are lists of residents who have registered to vote, lists of those who hold driver's licenses, lists of those who pay taxes, and so on. However, all these lists are incomplete (some people do not list their phone numbers or do not have telephones; some people do not register to vote or drive cars). Using these incomplete lists would introduce bias into our sample.

In such cases, the sampling procedures become a little more complex. We usually end up working toward the sample we want through successive approximations: by first extracting a sample from lists of groups or clusters that are available and then sampling the elements of interest from within these selected clusters. A cluster is a naturally occurring, mixed aggregate of elements of the population, with each element appearing in one and only one cluster. Schools could serve as clusters for sampling students, prisons could serve as clusters for sampling incarcerated offenders, neighborhoods could serve as clusters for sampling city residents, and so on. Sampling procedures of this nature are typically called **multistage cluster samples**.

Multistage cluster sampling: Sampling in which elements are selected in two or more stages, with the first stage being the random selection of naturally occurring clusters and the last stage being the random selection of multilevel elements within clusters.

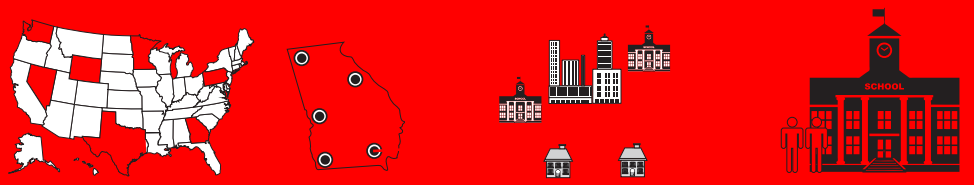
Drawing a cluster sample is at least a two-stage procedure. First, the researcher draws a random sample of clusters (e.g., blocks, prisons, and counties). Next, the researcher draws a random sample of elements within each selected cluster. Because only a fraction of the total clusters from the population are involved, obtaining a list of elements within each of the selected clusters is usually much easier.

Many large surveys sponsored by the federal government use multistage cluster samples. The U.S. Justice Department's NCVS is an excellent example of a multistage cluster sample. Because the target population of the NCVS is the entire U.S. population, the first stage of sampling requires selecting a first-order sample of counties and large metropolitan areas called primary sampling units (PSUs). From these PSUs, another stage of sampling involves the selection of geographic districts within each of the PSUs that have been counted by the 2000 census. And finally, a probability sample of residential dwelling units are selected from these geographic districts. These dwelling units, or addresses, represent the final stage of the multistage sampling. Or in a cluster sample of students, a researcher could contact the schools selected in the first stage and make arrangements with the registrars to obtain lists of students at each school. Figure 1.1 displays the multiple stages of a cluster sample like this.

Weighted or Stratified Samples

In some cases, the types of probability samples described earlier do not actually serve our purposes. Sometimes, we may want to make sure that certain segments of the population of interest are represented within our sample, and we do not want to leave this to chance. Say, for example,

Figure 1.1 Example of Cluster Sampling



Stage 1: Randomly select states

Stage 2: Randomly select cities, towns, and counties within those states

Stage 3: Randomly select schools within those cities and towns

Stage 4: Randomly select students within each school

that we are interested in incidents of personal larceny involving contact, such as purse snatching. We know from the NCVS that Americans older than 65 years of age are as vulnerable to this type of crime as those who are younger than 65. We may be interested in whether there are differences in the victimization circumstances (e.g., place or time of occurrence and number of offenders) between two groups of persons: those younger than 65 and those older than 65. To investigate this, we want to conduct a sample survey with the entire U.S. population. A simple random sample of the population, however, may not result in a sufficient number of individuals older than 65 to use for comparison purposes because individuals older than 65 make up a relatively small proportion of the entire population (approximately 12%).

Stratified or weighted sampling: Method of sampling in which sample elements are selected separately from population strata or are weighted differently for selection in advance by the researcher.

One way to achieve this goal would be to weight the elements in our population disproportionately. These samples are referred to as **stratified or weighted samples**. Instead of having an equal chance of being selected, as in the case of random samples, individuals would have a known but unequal chance of being selected. That is, some elements would have a greater probability of being selected into the sample than others. This would be necessary in our study of purse snatching because those older than 65 represent only about 12% of the total U.S. population. Because we want to investigate differences between the victimizations of those younger than and older than 65, we want to have more than this 12% proportion represented in our sample. To do this, we would disproportionately weight our sample selection procedures to give persons older than 65 a better chance of being selected. It is important to note that if we were going to make generalizations from a weighted sample to the population, then adjustments to our statistics would be necessary to take this sample weighting into account. This is a somewhat complicated procedure that is usually accomplished through the aid of computer technology.

NONPROBABILITY SAMPLING TECHNIQUES

As you can imagine, obtaining a probability sample such as those described in the previous section can be a very laborious, and sometimes costly, task. Many researchers do not have the resources, in either time or money, to obtain a probability sample. Instead, many rely on nonprobability sampling procedures. Unlike the samples we have already discussed, when samples are collected using nonprobability sampling techniques, elements within the target population do *not* have a known, equal, and independent probability of being selected. Because the chance of one element being selected versus another element remains unknown, we cannot be certain that the selected sample actually represents our target population. Since we are

generally interested in making inferences to a larger population, this uncertainty can represent a significant problem.

Why, then, would we want to use nonprobability sampling techniques? Well, they are useful for several purposes, including those situations in which we do not have a population list. Moreover, nonprobability-sampling techniques are often the only way to obtain samples from particular populations or for certain types of research questions, especially those about hidden or deviant subcultures. At other times when we are just exploring issues we may not need the precision (and added costs and labor) of a probability sample. We will briefly discuss three types of nonprobability samples in this section: availability, quota, and purposive or judgement samples.

Availability Samples

The first type of sampling technique we will discuss is one that is perhaps too frequently used and is based solely on the availability of respondents. This type of sample is appropriately termed an **availability sample**. The media often pass availability samples off as probability samples. Popular magazines and Internet sites periodically survey their readers by asking them to fill out questionnaires, and those individuals inclined to respond make up the availability sample for the survey. Follow-up articles then appear in the magazine or on the site displaying the results under such titles as "What You Think About the Death Penalty for Teenagers." Even if the number of people who responded is large, however, these respondents only make up a tiny fraction of the entire readership and are probably unlike other readers who did not have the interest or time to participate. In sum, these samples are not representative of the total population—or even of the total population of all readers.

Availability sampling: Sampling in which elements are selected on the basis of convenience.

You have probably even been an element in one of these samples. Have you ever been asked to complete a questionnaire in class, say as a course requirement for a psychology class? University researchers frequently conduct surveys by passing out questionnaires in their large lecture classes. Usually, the sample obtained from this method consists of those students who voluntarily agree to participate or those who receive course credit for doing so. This voluntary participation injects yet another source of bias into the sample. It is not surprising that this type of sample is so popular; it is one of the easiest and least expensive sampling techniques available. But it may produce the least representative and least generalizable type of samples.

Quota Samples

Quota sampling is intended to overcome availability sampling's biggest downfall: the likelihood that the sample will just consist of who or what is available, without any concern for its similarity to the population of interest. The distinguishing feature of a quota sample is that quotas are set to ensure that the sample represents certain characteristics in proportion to their prevalence in the population.

Quota sampling: Nonprobability sampling method in which elements are selected to ensure that the sample represents certain characteristics in proportion to their prevalence in the population or to oversampled segments of the population.

Quota samples are similar to stratified probability samples, but they are generally less rigorous and precise in their selection procedures. Quota sampling simply involves designating the population into proportions of some group that you want to be represented in your sample. Similar to stratified samples, in some cases, these proportions may actually represent the true proportions observed in the population. At other times, these quotas may represent predetermined proportions of subsets of people you deliberately want to oversample.

The problem is that even when we know that a quota sample is representative of the particular characteristics for which quotas have been set, we have no way of knowing if the sample is representative in terms of any other characteristics. Realistically, researchers can set quotas for only a small fraction of the characteristics relevant to a study, so a quota sample is really not much better than an availability sample (although following careful, consistent procedures for selecting cases within the quota limits always helps).

Purposive or Judgment Samples

Purposive or judgment sampling: Nonprobability sampling method in which elements are selected for a purpose usually because of their unique position.

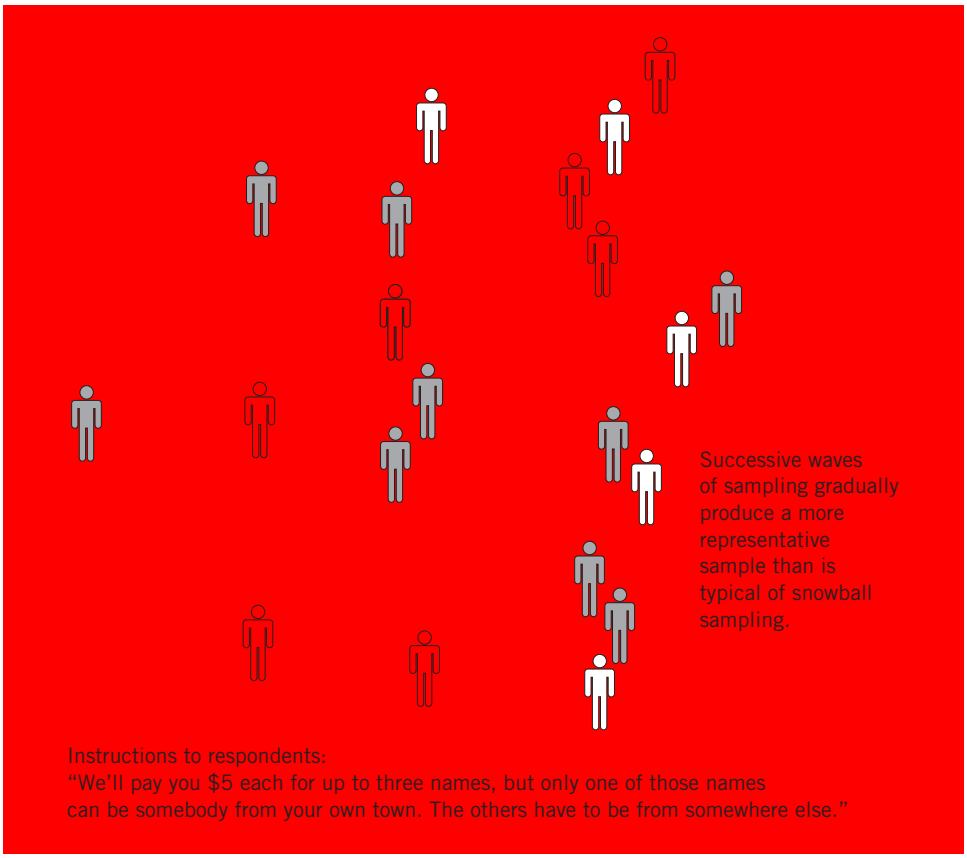
Another type of nonprobability sample that is often used in the field of criminology is called a **purposive** or **judgment sample**. In general, this type of sample is selected based on the purpose of the researcher's study and on his or her judgment of the population. It is often referred to as judgment sampling because the researcher uses her or his own judgment about whom to select into the sample, rather than drawing sample elements randomly. Although this type of sample does not provide the luxury of generalizability, it can provide a wealth of information not otherwise attainable from a typical random sample.

Many noted studies in the field of criminology have been carried out by using a purposive or judgment sample. For example, in the classic book *The Booster and the Snitch: Department Store Shoplifting*, Mary Cameron (1964) tracked a sample of individuals who had been caught shoplifting by department store employees.

Snowball sample: Type of purposive sample that identities one member of a population and then asks him or her to identify others in the population. The sample size increases as a snowball would rolling down a slope.

Another variation of a purposive sample is called a **snowball sample**. By using this technique, you identify one member of the population and speak to him or her, then ask that person to identify others in the population and speak to them, then ask them to identify others, and so on. The sample size increases with time as a snowball would rolling down a slope. This technique is useful for hard-to-reach or hard-to-identify interconnected populations where at least some members of the population know each other, such as drug dealers, prostitutes, practicing criminals, gang leaders, and informal organizational leaders. Figure 1.2 displays the process of snowball sampling.

Figure 1.2 Example of Snowball Sampling



To investigate the co-occurrence of intimate partner violence (IPV) and animal abuse, Fitzgerald and her colleagues Barrett, Stevenson, and Cheung (2019) obtained a purposive obtaining a random sample of the population would not be the ideal way to answer this research question. Of the women they surveyed, 55 women reported owning a pet while with their abusive partner. Of these women, 49 reported that their abusers also maltreated their pets. While the most common form of pet abuse included threats to get rid of the pet, or scaring or intimidating the pet, over 50% of the women reported that their pets were hit or had objects thrown at them, 20% reported that their pets were injured, and about 15% reported that their pets had been killed by their abusers. Based on these results, Fitzgerald et al. (2019) concluded that abuser's maltreatment of their partner's pets was driven by a "desire to cause them emotional harm and/or to enact power and control over them." (p. 1823)

We believe it is fundamental to identify the types of samples that are used in research before beginning a course in statistics. All inferential statistics we will examine in this text assume that the data being examined were obtained from a probability sample. What are inferential statistics, you ask? Good question. We will answer this next.

DESCRIPTIVE AND INFERENTIAL STATISTICS

Traditionally, the discipline of statistics has been divided into descriptive and inferential statistics. In large part, this distinction relies on whether one is interested in simply describing some phenomenon or in "inferring" characteristics of some phenomenon from a sample to the entire population. See? An understanding of sampling issues is already necessary.

Descriptive statistics can be used to describe characteristics or some phenomenon from either a sample or a population. The key point here is that you are using the statistics for "description" only. For example, if we wanted to describe the number of parking tickets given out by university police or the amount of revenues these parking tickets generated, we could use various statistics, including simple counts or averages.

Descriptive statistics: Statistics used to describe the distribution of a sample or population.

If, however, we wanted to generalize this information to university police departments across the country, we would need to move into the realm of **inferential statistics**. Inferential statistics are mathematical tools for estimating how likely it is that a statistical result based on data from a random sample is representative of the population from which the sample was selected. If our interest is in making inferences, a **sample statistic** is really only an estimate of the population statistic, called a **population parameter**, which we want to estimate. Because this sample statistic is only an estimate of the population parameter, there will always be some amount of error present. Inferential statistics are the tools used for calculating the magnitude of this sampling error. As we noted earlier, the larger the sampling error, the less accurate the sample statistic will be as an estimate of the population parameter. Of course, before we can use inferential statistics, we must be able to assume that our sample is actually representative of the population. And to do this, we must obtain our sample using appropriate probability sampling techniques. We hope the larger picture is beginning to come into focus!

Inferential statistics: Mathematical tools for estimating how likely it is that a statistical result based on data from a random sample is representative of the population from which the sample was selected.

Sample statistic: Statistic (i.e., mean, proportion, etc.) obtained from a sample of the population.

Population parameter: Statistic (i.e., mean, proportion, etc.) obtained from a population. Since we rarely have entire population data, we typically estimate population parameters using sample statistics.

VALIDITY IN CRIMINOLOGICAL RESEARCH

Before we conclude this introductory chapter, it is important to cover two more concepts. In criminological research, we seek to develop an accurate understanding of empirical reality by conducting research that leads to valid knowledge about the world. But when is knowledge valid? In general, we have reached the goal of validity when our statements or conclusions about empirical reality are correct. If you look out your window and observe that it is raining, this is probably a valid observation. However, if you read in the newspaper that the majority of Americans favor the death penalty for adolescents who commit murder, this conclusion should

be held up to stronger scrutiny because it is probably based on an interpretation of a social survey. There are two types of validity that we will examine here: measurement validity and causal validity.

Measurement Validity

Measurement validity: When we have actually measured what we intended to measure.

In general, we can consider **measurement validity** the first concern in establishing the validity of research results because if we haven't measured what we think we have measured, our conclusions may be completely false. To see how important measurement validity is, let's go back to the descriptive research question we addressed earlier: "How prevalent is youth violence and delinquency in the United States?"

Data on the extent of juvenile delinquency come from two primary sources: official data and surveys. Official data are based on the aggregate records of juvenile offenders and offenses processed by agencies of the criminal justice system: police, courts, and corrections. As noted earlier, one primary source of official statistics on juvenile delinquency is the UCR or the newer NIBRS produced by the FBI. However, the validity of these official statistics for measuring the extent of juvenile delinquency is a subject of heated debate among criminologists. Although some researchers believe official reports are a valid measure of serious delinquency, others contend that these data say more about the behavior of the police than about delinquency. These criminologists think the police are predisposed against certain groups of people or certain types of crimes.

Unquestionably, official reports underestimate the actual amount of delinquency. Obviously, not all acts of delinquency become known to the police. Sometimes delinquent acts are committed and not observed; other times they are observed and not reported, and if the official data include arrests, then even crimes that are observed and reported frequently do not result in anyone being arrested. In addition, there is evidence that UCR data often reflect the political climate and police policies as much as they do criminal activity. Take the U.S. "War on Drugs," which heated up in the 1980s. During this time, arrest rates for drug offenses soared, giving the illusion that drug use was increasing at an epidemic pace. However, self-report surveys that asked citizens directly about their drug use behavior during this same time period found that the use of most illicit drugs was actually declining (Regoli & Hewitt, 1994). In your opinion, then, which measure of drug use, the UCR or self-report surveys, was more valid? Before we answer this question, let's continue our delinquency example.

Despite the limitations of official statistics for measuring delinquency, these data were relied on by criminologists and used as a valid measure of delinquency for many decades. As a result, delinquency and other violent offending were thought to involve primarily minority populations and/or disadvantaged youth. In 1947, however, James Wallerstein and Clement Wyle surveyed a sample of 700 juveniles and found that 91% admitted to having committed at least one offense that was punishable by one or more years in prison and 99% admitted to at least one offense for which they could have been arrested had they been caught. In 1958, James Short and F. Ivan Nye reported the results from the first large-scale self-report study involving juveniles from a variety of locations. In their research, Short and Nye concluded that delinquency was widespread throughout the adolescent population and that youth from high-income families were just as likely to engage in delinquency as youth from low-income families. Contemporary studies using self-report data from the National Youth Survey (NYS) indicate that the actual amount of delinquency is much greater than that reported by the UCR and that, unlike these official data where nonwhites are overrepresented, self-report data indicate that white juveniles report almost exactly the same number of delinquencies as non-whites, but fewer of them are arrested (Elliott & Ageton, 1980).

This is just one example that highlights the importance of measurement validity, but it should convince you that we must be very careful in designing our measures and in subsequently evaluating how well they have performed. We cannot just assume that the measures we

use are measuring what we believe them to measure. Remember this as we use real data and case studies from the criminology and criminal justice literature throughout this book.

Reliability

There are several types of reliability, but we are only going to concentrate on the basic concept here. **Reliability** means that a measure procedure yields consistent scores as long as the phenomenon being measured is not changing. For example, if we gave students a survey about alcohol consumption with the same questions, the measure would be reliable if the same students gave approximately the same answers six months later, assuming their drinking patterns had not changed much. Reliability is a prerequisite for measurement validity; we cannot really measure a phenomenon if the measure we are using gives inconsistent results. Figure 1.3 illuminates the difference between reliability and measurement validity.

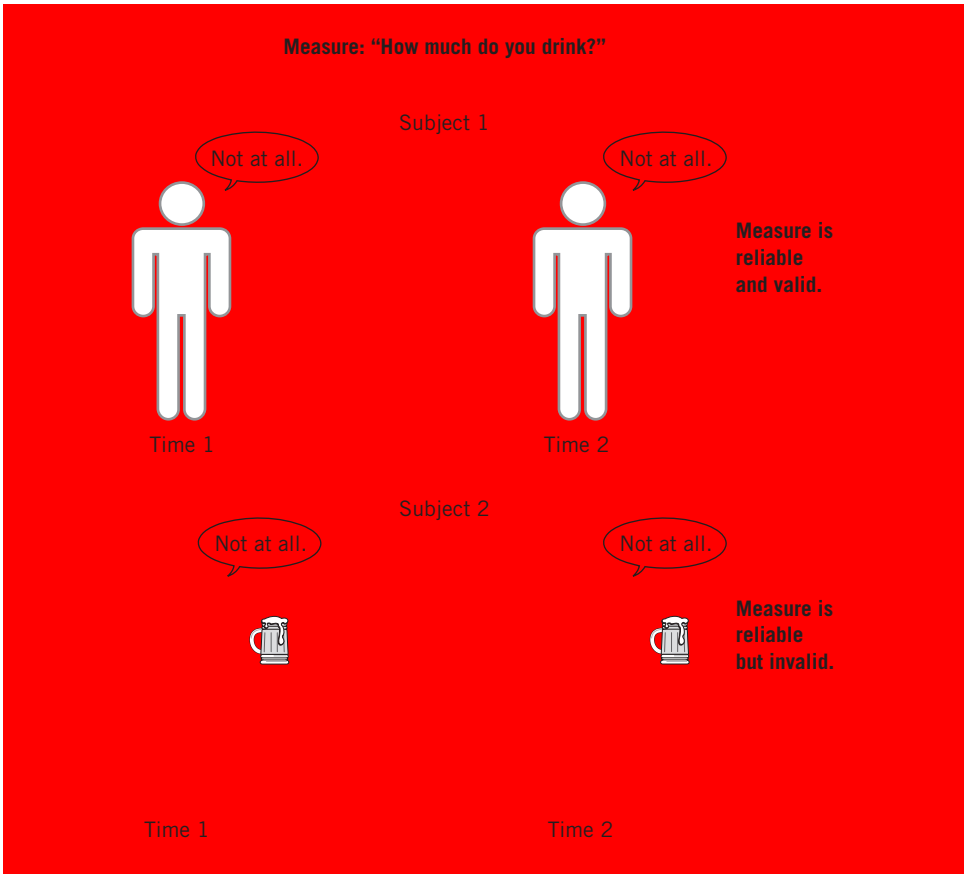
Reliability: Measure that is reliable when it yields consistent scores or observations of a given phenomenon on different occasions. Reliability is a prerequisite for measurement validity.

Causal Validity

Causal validity, also known as **interval validity**, is another issue of validity we are concerned with and has to do with the truthfulness of an assertion that an independent variable did, in

Causal validity (internal validity): When we can assume that our independent variable did cause the dependent variable.

Figure 1.3 Difference Between Reliability and Measurement Validity: Drinking Behavior



fact, cause the dependent variable, or that *X* caused *Y*. Let's go back to the issue of violence prevention programs in schools. Imagine that we are searching for ways to reduce violence in high schools. We start by searching for what seems to be particularly effective violence prevention programs in area schools. We find a program at a local high school—let's call it Plainville Academy—that a lot of people have talked about, and we decide to compare rates of violence reported to the guidance counselor's office in that school with those in another school, Cool School, that does not offer the violence prevention program. We find that students in the school with the special program have lower rates of reported violence, and we decide that the program caused the lower rates. Are you confident about the causal validity of our conclusion? Probably not. Perhaps the school with the special program had lower rates of reported violence even before the special program began. Maybe kids who go to Cool School are at a greater risk of violence because of where it is located.

This is the sort of problem that randomized experiments, like those reviewed by Gaffney and her colleagues (2019) are designed to resolve. Randomly assigning students to either receive a bullying prevention program or not made it very unlikely that students who were more aggressive would be disproportionately represented in either group. In addition, causal conclusions can be mistaken because of some factor that was not recognized during planning for the study, even in randomized experiments. Statistical control of other factors thought also to explain or predict the phenomenon of interest is essential in determining causal validity. The final two multiple regression chapters in this book highlight the ways research uses statistical methods to control for many independent variables thought to affect a dependent variable.

Our goal in this introductory chapter is to underscore the nature of the importance of statistics in criminology and criminal justice along with several fundamental aspects of the research process. We have set the stage for us to begin our exploration into the realm of statistics. Can't wait!

We have seen that, unlike observations we make in everyday life, criminological research relies on scientific methods. Statistical methods play a role in three types of research we conduct in our field: descriptive research, explanatory research, and evaluation research. The goal of all research is validity—for our statements or conclusions about empirical reality to be correct. Measurement validity exists when we have actually measured what we think we have measured. Causal or internal validity exists when the assertion that an independent variable causes a dependent variable, or that *X* causes *Y*, is correct. Generalizability, also known as external validity, exists when we can assume that results obtained from a sample can be generalized to the population.

Because it is almost never possible to obtain information on every individual or element in the population of interest, our investigations usually rely on data taken from samples of the population. Furthermore, because virtually all of the statistics we will examine in this text are based on assumptions about the origins of our data, we have provided a discussion of the most common types of samples used in our field of study. Samples generally fall within two categories: those derived from probability sampling techniques and those derived from nonprobability sampling techniques. The fundamental element in probability sampling is random selection. When a sample is randomly selected from the population, it means that every element (e.g., individual) has a known and independent chance of being selected for the sample.

We examined four types of probability samples: the simple random sample, the systematic random sample, the multistage cluster sample, and the weighted sample. In addition, we discussed three types of nonprobability samples: quota samples, purposive or judgment samples, and availability samples. We concluded the chapter with a brief discussion of descriptive and inferential statistics and highlighted the importance of measurement and causal validity.

Obtain a list of students from the statistics or research methods course in which you are currently using this book. Using this list and the random numbers table in Appendix B, select a simple random sample of 15 students. What are the steps you performed in doing this? Comment on how well this sample represents the entire sophomore class. Now draw a systematic random sample from the same list. Are there any differences?

How can you approximate a simple random sample when you do not have a list of the population?

Discuss the importance of probability sampling techniques.

How does random selection ensure that we are obtaining the most representative sample possible?

If we wanted to make sure that certain segments of the population were represented and/or overrepresented within our sample, what are two types of sampling techniques we could use?

What is the danger in using nonprobability samples in research?

In what types of situations would nonprobability samples be the most appropriate?

What is reliability? Why is this important?

What is measurement validity? Why is this important?

In their book, *Armed Robbers in Action: Stickups and Street Culture*, Richard Wright and Scott Decker (1997) conducted in-depth interviews with armed robbers to understand their mindset and decision making when planning, executing, and leaving an armed robbery. The sample was obtained through connections with current respondents where they were paid to get others to participate for interviews. What type of sampling is this?

Data for Exercise

Dataset	Description
2013YRBS.sav	The 2013YRBS, short for Youth Risk Behavior Survey, is a national study of high school students. It focuses on gauging various behaviors and experiences of the adolescent population, including substance use and some victimization.

SPSS introduction: SPSS, short for "Statistical Package for the Social Sciences," is a professional statistical analysis program that is used by universities, hospitals, and businesses. The exercises at the end of each chapter are intended to get you comfortable with the basics of SPSS. The first thing we've got to do is open some data:

First, go to the website for this textbook (**edge.sagepub.com/bachmansccj5e**) and download all the data sets somewhere you can access easily.

Opening a data set in SPSS: After double clicking on the SPSS icon you'll see a spreadsheet in the background and a welcome screen pop up, asking what you want to do. Select "Open another file . . ." and "Open." (or simply double click "Open another file . . ." This will cause a browser to open, at which point you must simply go to the folder that contains the data set of interest; in this case, that is the 2013 YRBS.sav file. SPSS also allows you to import data from other packages by clicking "File," "Import Data" and then clicking the file the data you wish to import is in.

Alternately, if the welcome screen does not pop up, you can always select file->open->data from the menu bar to access your data set.

Navigating SPSS:

SPSS uses two main screens through which you can view your data set. The buttons to switch between "views" are on the bottom left of the SPSS window. By default you open up to the **Data View** screen. However, let's begin with **Variable View**:

Variable View in SPSS:

Variable view in SPSS lets you look at information on each variable in the data set. Each numbered row corresponds to a different measure from the survey. Some of the information in the columns is not of interest to us for this book; instead, focus on the following:

Name: The name of the variable. Double clicking allows you to edit this field.

Label: A summary of what the variable tells us or how the question was asked. Double clicking allows you to edit this field.

Values: This allows us to put labels on numeric values. For instance, we could tell SPSS that responses with a 1 should be labeled as "White." You can specify your own labels or view them by clicking the cell for a given row and clicking the ". . ." field.

Missing: This tells us values that SPSS will treat as missing, excluding them from analyses. Many surveys code cases as –9 or a similar value rather than leaving them blank so it is important to make sure these are treated as missing. You can specify missing values by clicking in the cell and pressing the ". . ." box.

Variable View Exercises:

Identify the variable name, label, and value labels for the following variables:

Row 2

Row 4

Row 23

Row 45

Searching for variables: If you know the variable name, you can search for it to make the process faster. Do this by clicking on the column "Name" and press ctrl+F on your keyboard, enter the name of the variable, and click "Find Next"; or selecting edit->find. There are also tabs for "Go to case" and "Go to variable" that can be used for finding specific cases and variables, respectively. Find the labels and values for the following variables:

qn43

qhallucDrug

qnowt

How many variables are in this data set?

Data View:

On the bottom left of the screen next to "Variable View" is "Data View" In Data View each column is a different variable. Each row corresponds to a respondent; it contains a person's specific responses to all the survey questions.

You'll notice lots of "." marks; these are "system missing" responses. We just don't have data for that person for whatever reason! SPSS ignores these automatically.

Data for Exercise

Dataset	Description
2013YRBS.xls	The 2013YRBS, short for Youth Risk Behavior Survey, is a national study of high-school students. It focuses on gauging various behaviors and experiences of the adolescent population, including substance use and some victimization.

Excel introduction: Excel is not short for anything, but is rumored to have been named Excel because it can do anything. Excel is also a popular program used in universities, hospitals, and businesses just like SPSS and Stata. Excel is unique in that it can do everything covered in this textbook from basic analyses to complicated regression. Ironically, the easier output is typically harder to obtain than the more complicated outputs in Excel.

Go to (**edge.sagepub.com/bachmansccj5e**) to download the data.

Opening a data set in Excel: Excel data files contain a ".xls" or ".xlsx" at the end of its name. You can simply double click any ".xls" or ".xlsx" file for it to be opened in Excel; or in programs such as Stata or SPSS where the data are already loaded for example, click "File," "Export" and then click "Data to Excel Spreadsheet."

Navigating Excel:

Excel uses one main screen that contains the list of the variables that are currently in Excel to be used on the first row with letters representing columns. You can easily add "Sheets" by clicking the plus on the bottom. This is useful to copy and paste certain variables into a new sheet for easier analysis.

Data Editor in Excel:

Everything for Excel is contained on the main screen where you can click on any cell and make any changes.

Name: The name of the variable is in the first row and the column represents the values of the variable. You can change the variable name simply by entering the new name. Note that the rows represent individual cases.

Label: There are no explicit ways to label variables in Excel unless you add it to the variable name. This can be done by adding a parenthesis after the name and including relevant information.

Missing: Missing values can take the form of dots (".") or simply blank spaces.

Data Editor Exercises:

Searching for variables: If you know the variable name, you can search for it to make the process faster. Do this by clicking "Ctrl+F" and entering the variable name. Find the values for the following variables:

qn43

qhallucDrug

qnowt

How many variables are in this data set?

Data Editor Exercises:

What was respondent 1's (i.e., row 1) response to question q13?

What was respondent 71's race according to the variable race7?

How many respondents do we have in this data set?

Data for Exercise

Dataset	Description
2013YRBS.dta	The 2013YRBS, short for Youth Risk Behavior Survey, is a national study of high-school students. It focuses on gauging various behaviors and experiences of the adolescent population, including substance use and some victimization.

While SPSS is focused upon a point and click method, Stata primarily utilizes command lines but does allow for point and click.

Similarly to SPSS, go to (**edge.sagepub.com/bachmansccj5e**) to download the data.

Opening a data set in Stata: Stata data files contain a ".dta" at the end of the file name. You can simply double click any ".dta" file for it to be opened in Stata, or in Stata, click "File," "Open" and then click the dataset you wish to use. If this is not the case, however, Stata is very flexible in incorporating data from other software packages. To import an SPSS dataset that ends in ".sav", click "File" in Stata, "Import" and then "SPSS data (*.sav)." Note also all of the other programs you can import into Stata.

Navigating Stata:

Stata uses one main screen that contains the command line history, output window, command line, and variable list of the variables that are currently in Stata to be used.

Data Editor in Stata:

At the top of Stata, there is a task bar with two "Data Editor" tabs: one is for browsing and the other is for editing. It is safer to stay in the browsing tab so as to not make unwanted changes to the data; you can easily transfer over to the editor if need be.

Name: The name of the variable is in the top shaded row. If you click on a variable, a box titled "Properties" appears and describes the variable. In edit mode, you can change the variable name simply by entering the new name. Note that the rows represent individual cases.

Label: Under the "Properties" tab, describes the variable. Similar to "Name," you can click on this in editor mode and change the values.

Missing: In the viewer, all of the dots (".") represents data that are missing.

Data Editor Exercises:

Searching for variables: If you know the variable name, you can search for it to make the process faster. Do this by entering the name of the variable in the line titled "Filter variables here" below the "Variables" tab. Find the labels and values for the following variables:

qn43

qhallucDrug

qnowt

How many variables are in this data set?

Data Editor Exercises:

What was respondent 1's (i.e., row 1) response to question q13?

What was respondent 71's race according to the variable race7?

How many respondents do we have in this data set?

Access the eFlashcards, data sets, and software output for SPSS, Excel, and Stata at

.

PART ONE

UNIVARIATE ANALYSIS

Describing Variable Distributions

LEVELS OF MEASUREMENT AND AGGREGATION

Learning Objectives

1. Summarize the role of variables in research.
2. Identify the four levels of measurement variables can have.
3. Describe the difference between variables that identify qualities compared with variables that identify quantities.
4. Explain the differences among raw frequencies, proportions, percentages, and rates.
5. Define the units of analysis in any particular data set.

Science cannot progress without reliable and accurate measurement of what it is you are trying to study. The key is measurement, simple as that.

—Robert D. Hare

When you can measure what you are speaking about, and express it in numbers, you know something about it.

—The Lord Kelvin

INTRODUCTION

In Chapter 1, we examined various sampling techniques that can be used for selecting a sample from a given population. Once we have selected our sample, we can begin the process of collecting information. The information we gather is usually referred to as "data" and in its entirety is called a "data set." In this chapter, we will take a closer look at the types of variables that can make up a data set.

This may be the first time you have been formally exposed to statistics, but we are sure each of you has some idea what a variable is even though you may not call it that. A **variable** is any element to which different values can be attributed. Although society is beginning to acknowledge the fluidity of gender, it is still a variable that is typically measured using two values, male and female. Race/ethnicity is a variable with many values, such as American Indian, African American, Asian, Latinx, Caucasian, and mixed. Age is another variable that can take on different values, such as 2, 16, 55 years, and so on.

Variable: Characteristic or property that can vary or take on different values or attributes.

As we noted in the last chapter, in explanatory research, we are interested in explaining a dependent variable by using one or more independent variables. In research, the dependent variable is expected to vary or change depending on variation or change in the independent variable. In this causal type of explanation, the independent variable is the cause and the dependent variable the effect or outcome. The entire set of values a variable takes on is called a **frequency distribution** or an **empirical distribution**. In a given data set, a frequency distribution is a distribution (a list) of outcomes or values for a variable. It is referred to as an empirical distribution because it is a distribution of empirical (real and observed) data, and it is called a frequency distribution because it tells us how frequent each value or outcome is in the entire data set. For example, suppose we conducted a survey from a sample of 100 persons in your class at your university. In one question we asked for respondent's age. Suppose this "age" variable ranged from 18 to 42. There might be 15 people who were 18 years of age, 30 people who were 19 years of age, 17 people who were 20 years of age, only 1 person who was 42 years of age, and so on. An empirical, or frequency, distribution would tell you not only what the different ages were but also how many people in the sample were represented by each age in the entire distribution.

Frequency or empirical distribution: Distribution of values that make up a variable distribution.

In contrast, a characteristic of your sample that does not vary in a data set is called a **constant**. Unlike a variable, whose values vary or are different, a constant has only one value. For example, if you have a sample of inmates from a male correctional institution, the value for "gender" would be considered a constant—"male." Since all elements of the sample would be male, respondent's gender would not vary in that data set. Similarly, if you selected a sample of 20-year-olds from the sophomore class at a state university, age would be a constant rather than a variable in that sample because all members of the sample would be the same age (20 years).

Constant: Characteristic or property that does not vary but takes on only one value.

Notice that a given characteristic, such as respondent's gender or age, is not always a variable or a constant. Under different conditions, it may be one or the other. For example, in a sample of male prisoners, gender is a constant, but age is a variable because the male inmates are likely to be different ages. In the sample of 20-year-old sophomore students from a university, age is a constant, and respondent's gender is a variable because some persons in the sample would be male and some would be female.

We can classify variables in many different ways and make several distinctions among them. First, there are differing levels of measurement that can be associated with variables. The next section of the chapter examines these measurement differences, beginning with the classification of variables as either continuous or categorical variables. We then examine the four measurement classifications within these broad categories: nominal, ordinal, interval, and ratio measurement. The second section of the chapter addresses the difference between independent and dependent variables and the different ways of reporting the features of variables. In the final section, you will learn how to identify the units of analysis in a research design so that you can state conclusions about the relationships between your variables in the appropriate units. As you will soon see, understanding this information is extremely important as every statistical application we use typically depends on a variable's levels of measurement. And when we interpret statistical results, we can only generalize that result to the units of analysis that were observed in the sample.

LEVELS OF MEASUREMENT

Recall that data generally come from one of three places: They are gathered by us personally, gathered by another researcher, or gathered by a government agency. Doing research on a previously collected data set is often referred to as "secondary data analysis" because the data already existed and had been analyzed before. No matter how they were collected, however, data sets are by definition simply a collection of many variables. For illustrative purposes, imagine that we were interested in the relationship between levels of student drinking and drug use and student demographic characteristics such as gender, age, religion, and year in college (freshman, sophomore, junior, senior). Table 2.1 displays the small data set we might have obtained had we investigated this issue by collecting surveys from 20 college students (a random sample, of course).

Table 2.1 Example of the Format of a Data Set From a Survey of 20 College Students

ID Number	Gender	Age	College Year	*GPA*	Average Month		Religion
					# Drinks	# Times Used Marijuana	
1	Female	19	Sophomore	2.3	45	22	Catholic
2	Male	22	Senior	3.1	30	10	Other
3	Female	22	Senior	3.8	0	0	Protestant
4	Female	18	Freshman	2.9	35	5	Jewish
5	Male	20	Junior	2.5	20	20	Catholic
6	Female	23	Senior	3.0	10	0	Catholic
7	Male	18	Freshman	1.9	45	25	Not religious
8	Female	19	Sophomore	2.8	28	3	Protestant
9	Male	28	Junior	3.3	9	0	Protestant
10	Female	21	Junior	2.7	0	0	Muslim
11	Female	18	Freshman	3.1	19	2	Jewish
12	Male	19	Sophomore	2.5	25	20	Catholic
13	Female	21	Senior	3.5	2	0	Other
14	Male	21	Junior	1.8	19	33	Protestant
15	Female	42	Sophomore	3.9	10	0	Protestant
16	Female	19	Sophomore	2.3	45	0	Catholic
17	Male	21	Junior	2.8	29	10	Not religious
18	Male	25	Sophomore	3.1	14	0	Other
19	Female	21	Junior	3.5	5	0	Catholic
20	Female	17	Freshman	3.5	28	0	Jewish

Qualitative or categorical variables: Values that refer to qualities or categories. They tell us what kind, what group, or what type a value is referring to.

To measure the extent to which each student used alcohol and marijuana, let's say we asked them these questions: "How many drinks do you consume in an average month? By 'drinks' we mean a beer, a mixed drink, or a glass of wine." "How many times during an average month do you use marijuana?" Each of the other variables in the table relates to other information about each student in the sample. Everything listed in this table, including the respondent's identification number, is a variable. All of these variables combined represent our data set. The first thing you may notice about these variables is that some are represented by categories and some are represented by actual numbers. Gender, for example, is divided into two categories, female and male. This type of variable is often referred to as a **qualitative** or **categorical variable**, implying that the values represent qualities or categories only. The values of this variable have no numeric or quantitative meaning. Other examples in the data set of qualitative variables include college year and religion.

The rest of the variables in our data set, however, have values that do represent numeric values that can be quantified—hence the name **quantitative or continuous variables**. The values of quantitative variables can be compared in a numerically meaningful way. Respondent's identification number, age, grade point average, number of drinks, and number of times drugs were used are all quantitative variables. We can compare the values of these variables in a numerically meaningful way. For example, from Table 2.1, we can see that respondent 1 has a lower grade point average than respondent 19. We can also see that respondents 7 and 16 have the highest levels of alcohol consumption in the sample.

Quantitative or continuous variables: Values that refer to quantities or different measurements. They tell us how much or how many.

In Table 2.1, it is relatively easy to identify which variables are qualitative and which are quantitative simply because the qualitative variables are represented by **alphanumeric data** (by letters rather than by numbers). Data that are represented by numbers are called **numeric data**. A good way to remember the distinction between these two types of data is to note that alphanumeric data consist of letters of the alphabet, whereas numeric data consist of numbers.

Alphanumeric data: Values of a variable that are represented by letters rather than by numbers.

Numeric data: Values of a variable that represent numerical qualities.

It is certainly possible to include alphanumeric data in a data set, as we have done in Table 2.1, but when stored in a computer, as most data are, alphanumeric data take up a great deal of space, and alphanumeric data are difficult to statistically analyze. For this reason, these data are usually converted to or represented by numeric values. For example, females may arbitrarily be identified with the number 1, rather than with the word "female," and males with the number 2. Assigning numbers to the categorical values of qualitative variables is called "coding" the data. Of course, which numbers get assigned to qualitative variables (for example, 1 for females and 2 for males) is arbitrary because the numeric code (number) assigned has no real quantitative meaning. Males could be given either a 1 or a 2, or a 0, with females coded either a 2 or a 1; it makes no difference. The numbers would still only be representing qualities or categories.

Table 2.2 redisplays the data in Table 2.1 numerically as they would normally be stored in a computer data set. Because values of each variable are represented by numbers, it is a little more difficult to distinguish the qualitative variables from the quantitative variables. You have to ask yourself what each of the values really means. For example, for the variable gender, what does the "1" really represent? It represents the code for a female student and is therefore not numerically meaningful. Similarly, the number "1" coded for the religion variable represents those students who said they were Catholic, and the code "3" represents those students who said they were Jewish. There is nothing inherently meaningful about the numbers 1 and 3. They simply represent categories for the religion variable and we changed the letters of the alphabet to numbers. For the variable age, what does the number 19 represent? This is actually a meaningful value—it tells us that this respondent was 19 years of age, and it is therefore a quantitative variable.

Table 2.2 Example of the Data Presented in Table 2.1 as They Would Be Stored in a Computer Data File

ID Number	Gender	Age	College Year	*GPA*	Average Month		Religion
					# Drinks	# Times Drugs Used	
1	1	19	2	2.3	45	22	1
2	2	22	4	3.1	30	10	6
3	1	22	4	3.8	0	0	2
4	1	18	1	2.9	35	5	3
5	2	20	3	2.5	20	20	1

(Continued)

Table 2.2 (Continued)

ID Number	Gender	Age	College Year	*GPA*	Average Month		Religion
					# Drinks	# Times Drugs Used	
6	1	23	4	3.0	10	0	1
7	2	18	1	1.9	45	25	5
8	1	19	2	2.8	28	3	2
9	2	28	3	3.3	9	0	2
10	1	21	3	2.7	0	0	4
11	1	18	1	3.1	19	2	3
12	2	19	2	2.5	25	20	1
13	1	21	4	3.5	2	0	6
14	2	21	3	1.8	19	33	2
15	1	42	2	3.9	10	0	2
16	1	19	2	2.3	45	0	1
17	2	21	3	2.8	29	10	5
18	2	25	2	3.1	14	0	6
19	1	21	3	3.5	5	0	1
20	1	17	1	3.5	28	0	3

Level of measurement: Mathematical nature of the values for a variable.

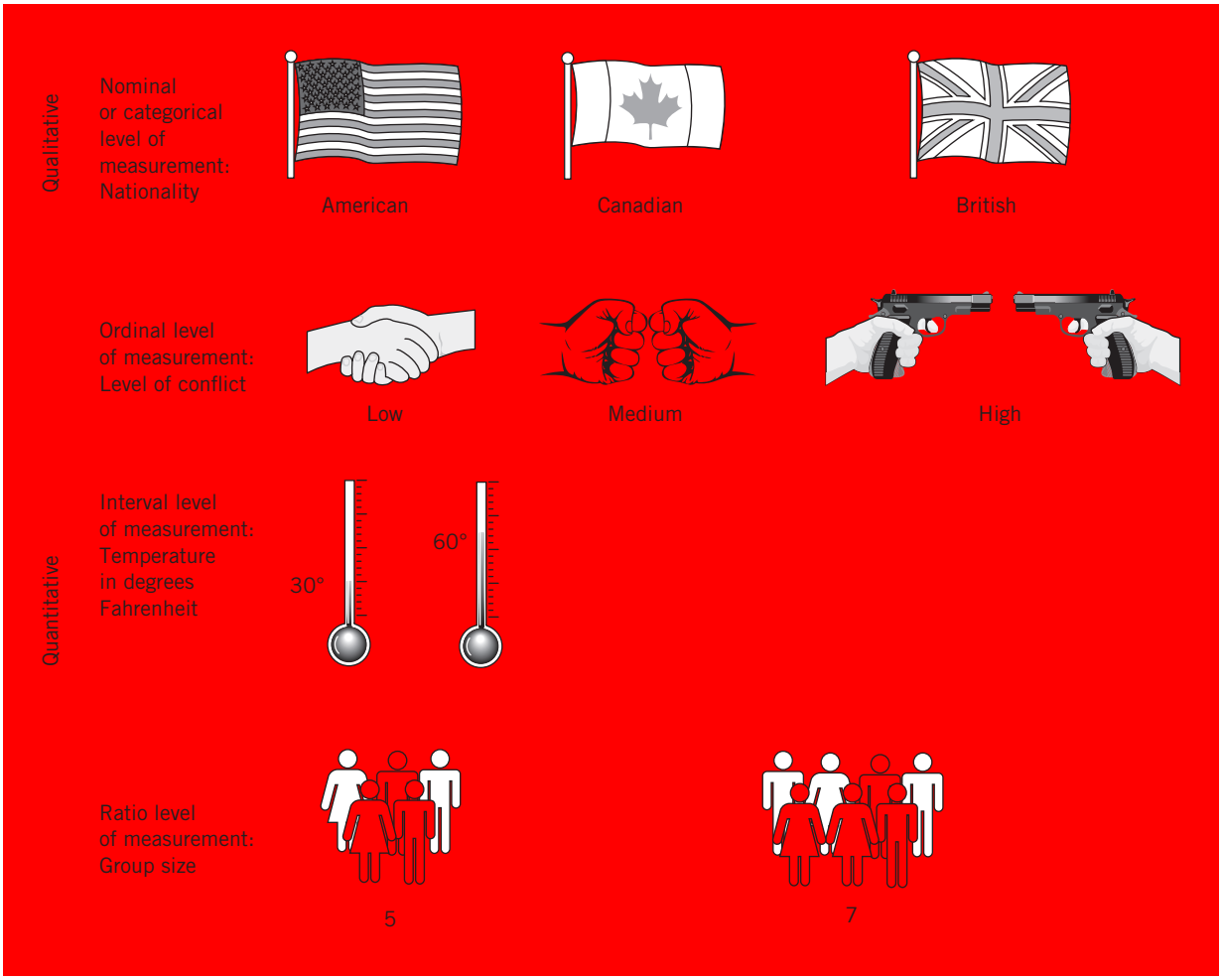
In addition to distinguishing between qualitative and quantitative, we can differentiate among variables in terms of what is called their **level of measurement**. The four levels of measurement are (1) nominal, (2) ordinal, (3) interval, and (4) ratio. Figure 2.1 depicts the difference among these four levels of measurement.

Nominal Level of Measurement

Nominal-level variables: Values that represent categories or qualities of a case only.

Variables measured at the nominal level are exclusively qualitative in nature. The values of **nominal-level variables** convey classification or categorization information *only*. Therefore, the only thing we can say about two or more nominal-level values of a variable is that they are different. We cannot say that one value reflects more or less of the variable than the other. The most common types of nominal-level variables are gender (male and female), religion (Protestant, Catholic, Jewish, Muslim, etc.), and political party (Democrat, Republican, Independent, etc.). The values of these variables are distinct from one another and can give us only descriptive information about the type or label attached to a value. Notice we can say that males are different from females but not that they have more "gender." We can say that Protestants have a different religion than Catholics or Jews, but again, not that they have more "religion." The only distinction we can make with nominal-level variables is that their values are different.

Figure 2.1 Levels of Measurement



Because they represent distinctions only of kind (one is merely different from the other), the categories of a nominal-level variable are not related to one another in any meaningful numeric way. This is true even if the alphanumeric values are converted or coded into numbers. For example, in Table 2.2, the values assigned to the variables gender and religion are given numeric values. Remember, however, that these numbers were simply assigned for convenience and have no numeric meaning. The fact that Catholics are assigned the code of 1 and Protestants are assigned the code of 2 does not mean that Protestants have twice as much religion as Catholics or that the Protestant religion is "more than" the Catholic religion. The only thing that the codes of 1 and 2 mean is that they refer to different religions. Because we cannot make distinctions of "less than" or "more than" with them, then, nominal-level variables do not allow us to rank-order the values of a given variable. In other words, nominal-level measurement does not have the property of order. It merely reflects the fact that some values are different from others. Consequently, mathematical operations cannot be performed with nominal-level data. With our religion variable, for example, we cannot subtract a 2 (Protestant) from a 3 (Jewish) to get a 1 (Catholic). Do you see how meaningless mathematical operations are with variables measured at the nominal level?

Ordinal Level of Measurement

Ordinal-level variables: Values that not only represent categories but also have a logical order.

The values of **ordinal-level variables** not only are categorical in nature, but the categories also have some type of relationship to each other. This relationship is one of order or transitivity. That is, categories on an ordinal variable can be rank-ordered from high (more of the variable) to low (less of the variable) even though they still cannot be exactly quantified. As a result, although we can know whether a value is more or less than another value, we do not know exactly *how much* more or less. The properties of ordinal-level measurement are clearer with an example.

Let's say that on a survey, we have measured income in such a way that respondents simply checked the income category that best reflected their annual income. The categories the survey provided are as follows:

1. Less than $20,000
2. $20,001 to $40,000
3. $40,001 to $60,000
4. $60,001 to $80,000
5. more than $80,001

Now suppose that one of our respondents (respondent 1) checked the first category and that another respondent (respondent 2) checked the third category. We don't know the exact annual income of each respondent, but we do know that the second respondent makes more than the first. Thus, in addition to knowing that our respondents have different annual incomes (nominal level), we also know that one income is more than the other. In reality, respondent 1 may make anywhere between no money and $20,000, but we can never know. Had we measured income in terms of actual dollars earned per year, we would be able to make more precise mathematical distinctions between respondents' annual incomes. Suppose we had a third person (respondent 3) who checked the response of more than $80,000. The property of transitivity says that if respondent 1 makes less than respondent 2, and if respondent 2 makes less than respondent 3, then respondent 1 also makes less than respondent 3. The rank order is thus:

1. Less than $20,000 respondent 1
2. $20,001 to $40,000
3. $40,001 to $60,000 respondent 2
4. $60,001 to $80,000
5. more than $80,001 respondent 3

Other examples of ordinal-level variables include the "Likert-type" response questions found on surveys that solicit an individual's attitudes or perceptions. You are probably familiar with this type of survey question. A typical one follows: "Please respond to the following statement by circling the appropriate number: '1' Strongly Agree, '2' Agree, '3' Disagree, '4' Strongly Disagree." The answers to these questions represent the ordinal level of measurement. Often these categories are displayed like this:

1	2	3	4
Strongly Agree	Agree	Disagree	Strongly Disagree

Response categories that rank-order attitudes in this way are often called *Likert* responses after Rensis Likert, who is believed to have developed them back in the 1930s. There are other ways to measure judgements using a Likert-type response. For example, the aggression questionnaires (AQs) in the literature are designed to measure an individual's propensity to feel anger and hostility (Buss & Warren, 2000). It consists of 34 items, such as "Given enough provocation, I may hit another person," "When people annoy me, I may tell them what I think of them," and "I have trouble controlling my temper." Individuals taking the AQ are asked to respond to the statements using a five-point Likert-type scale from "not at all like me," which is coded 0, to "completely like me," coded 5. Tracey Skilling and Geoff Sorge (2014) used the AQ to assess the validity of two other scales, one intended to measure criminal attitudes and another intended to measure antisocial attitudes. Using a sample of delinquent offenders in Canada, they found that all three measures were significantly related to each other, indicating that they were each measuring antisocial attitudes.

Interval Level of Measurement

In addition to enabling us to rank-order values, **interval-level variables** allow us to quantify the numeric relationship among them. To be classified as an interval-level variable, the difference between values along the measurement scale must be the same at every two points. For example, the difference in temperature on the Fahrenheit scale between 40 degrees and 41 degrees is the same 1 degree difference as the difference between 89 degrees and 90 degrees. This 1 degree difference is the same difference that exists between all the other values on the Fahrenheit scale.

Interval-level variable: In addition to an inherent rank order, a value's numeric relationship to other values is known. There is an equal and constant distance between adjacent values. Therefore, the values can be added and subtracted.

Another characteristic of interval-level measurement is that the zero point is arbitrary. An arbitrary zero means that, although a value of zero is possible, zero does not mean the absence of the phenomenon. A meaningless zero is an arbitrary zero. For example, a temperature on the Fahrenheit scale of 0 degrees does not mean that there is no temperature outside; it simply means that it is cold! Zero degrees on the Fahrenheit scale is arbitrary. These characteristics allow scores on an interval scale to be added and subtracted, but meaningful multiplication and division cannot be performed. This level of measurement is represented in Figure 2.1 by the difference between two Fahrenheit temperatures. Although 60 degrees is 30 degrees hotter than 30 degrees, 60 in this case is not twice as hot as 30. Why not? Because heat does not begin at 0 degrees on the Fahrenheit scale.

Social scientists often treat indices (see the AQ earlier) that were created by combining responses to a series of variables measured at the ordinal level as interval-level measures. Another example of an index like this could be created with responses to the Core Institute's (2020) questions about friends' disapproval of substance use (see Table 2.3). The survey has 13 questions on the topic, each of which has the same three response choices. If Do Not Disapprove is valued at 1, Disapprove is valued at 2, and Strongly Disapprove is valued at 3, then the summed index of disapproval would range from 12 to 36. The average could then be treated as a fixed unit of measurement. So a score of 20 could be treated as if it were 4 more units than a score of 16 and so on.

Ratio Level of Measurement

Ratio-level variables have all the qualities of interval-level variables, and the numeric difference between values is based on a natural, or true-zero, point. A true-zero point means that a score of zero indicates that the phenomenon is absent. For example, if people were asked how many hours they worked last month and they replied "zero hours," it would mean that there was a complete absence of work—they were unemployed that month. Ratio measurement allows meaningful use of multiplication and division, as well as addition and subtraction. We can therefore divide one number by another to form a ratio—hence the name of this level of measurement. Suppose we were conducting a survey of the victimization experiences of residents in

Ratio-level variables: Variables that we assume can be added and subtracted as well as multiplied and divided and that have true-zero points.

Table 2.3 Ordinal-Level Variables Can Be Added to Create an Index With Interval-Level Properties: Core Alcohol and Drug Survey

How Do You Think Your Close Friends Feel (or Would Feel) About You . . . (Mark One for Each Line)	Do Not Disapprove	Disapprove	Strongly Disapprove
Trying marijuana once or twice			
Smoking marijuana occasionally			
Smoking marijuana regularly			
Trying cocaine once or twice			
Taking cocaine regularly			
Trying LSD once or twice			
Taking LSD regularly			
Trying amphetamines once or twice			
Taking amphetamines regularly			
Taking one or two drinks of an alcoholic beverage (beer, wine, liquor) nearly every day			
Taking four or five drinks nearly every day			
Having five or more drinks in one sitting			
Taking steroids for bodybuilding or improved athletic performance			

Source: Adapted from *Core Alcohol and Drug Survey: Long Form* 2020 from the Core Institute.

rural areas and asked them to provide their annual income in dollars. This variable would be an example of the ratio-level of measurement because it has both a true-zero point and equal and known distances between adjacent values. For example, a value of no income, "zero dollars," has inherent meaning to all of us, and the difference between $10 and $11 is the same as that between $55,200 and $55,201.

There are a few variables in Table 2.2 that are measured at the ratio level. One is the number of drinks respondents had in an average month. Notice that there were a few respondents who had "0" drinks—this is an absolute zero! And a college student who drinks an average of 20 drinks a month has 10 more drinks than someone who has 10 drinks a month and 10 fewer drinks than someone who has an average of 30 drinks a month. We have not shown you how to calculate the mean yet, but imagine we calculate the average number of drinks a senior in college has from this table and find that it is 10.5 drinks. We then calculate the average number of drinks a first-year student has as 31.75 drinks. Because this is a ratio-level variable with an absolute zero, we could now take the ratio of drinks consumed by a first-year student compared with a senior to be (31.75 / 10.5 = 3.02) and say that first-year students consume about 3 times as much alcohol as seniors! Does this seem accurate to you? Because we can do this, the level of measurement is called "ratio."