



ELEMENTARY STATISTICS

in Social Research

TWELFTH EDITION

Jack Levin | James Alan Fox | David R. Forde

Elementary Statistics in Social Research

TWELFTH EDITION

Jack Levin

Northeastern University

James Alan Fox

Northeastern University

David R. Forde

University of North Florida

PEARSON

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Executive Editor: *Stephen Frail*
Editorial Assistant: *Caroline Beimford*
Executive Marketing Manager: *Wendy Albert*
Marketing Assistant: *Frank Alarcon*
Production Manager: *Fran Russello*
Art Director: *Jayne Conte*
Editorial Production and Composition Service: *Soumitra Borkakati / PreMediaGlobal*
Cover Designer: *Suzanne Behnke*
Cover Image Credit: *Aleksandar Videnovic / Shutterstock*
Printer: *Courier Companies, Inc.*
Cover Printer: *Lehigh Phoenix*

Copyright © 2014 Pearson Education, Inc. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Permissions Department, One Lake Street, Upper Saddle River, NJ 07458, or fax your request to 201-236-3290.

Many of the designations by manufacturers and seller to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

Library of Congress Cataloging-in-Publication Data

Levin, Jack

Elementary statistics in social research / Jack Levin, Northeastern University, James Alan Fox, Northeastern University, David R. Forde, University of North Texas.—12th ed.

pages cm

Includes bibliographical references and index.

ISBN-13: 978-0-205-84548-4

ISBN-10: 0-205-84548-7

1. Social sciences—Statistical methods. 2. Statistics. I. Fox, James Alan. II. Forde, David R. (David Robert), 1959– III. Title.

HA29.L388 2013

300.72'7—dc23

2013005380

10 9 8 7 6 5 4 3 2 1

PEARSON

ISBN 13: 978-0-205-84548-4

ISBN 10: 0-205-84548-7

Contents

Preface xi

1 *Why the Social Researcher Uses Statistics* 1

The Nature of Social Research 1

Why Test Hypotheses? 9

The Stages of Social Research 11

Using Series of Numbers to Do Social Research 11

The Functions of Statistics 18

Summary 24

Terms to Remember 25

Questions and Problems 25

Computer Exercises 30

Looking at the Larger Picture: A Student Survey 31

PART I • Description 33

2 *Organizing the Data* 35

Frequency Distributions of Nominal Data 35

Comparing Distributions 36

Proportions and Percentages 37

Ratios and Rates 38

Simple Frequency Distributions of Ordinal and Interval Data 41

Grouped Frequency Distributions of Interval Data 42

<i>Cumulative Distributions</i>	45
<i>Percentiles</i>	47
<i>Dealing with Decimal Data</i>	48
<i>More on Class Limits</i>	51
<i>Flexible Class Intervals</i>	53
<i>Cross-Tabulations</i>	54
<i>Graphic Presentations</i>	61
<i>Summary</i>	71
<i>Terms to Remember</i>	72
<i>Questions and Problems</i>	72
<i>Computer Exercises</i>	80

3 *Measures of Central Tendency* **81**

<i>The Mode</i>	82
<i>The Median</i>	82
<i>The Mean</i>	83
<i>Taking One Step at a Time</i>	87
BOX 3.1 • <i>Step-by-Step Illustration: Mode, Median, and Mean</i>	88
<i>Obtaining the Mode, Median, and Mean from a Simple Frequency Distribution</i>	89
<i>Comparing the Mode, Median, and Mean</i>	91
<i>Summary</i>	95
<i>Terms to Remember</i>	96
<i>Questions and Problems</i>	96
<i>Computer Exercises</i>	102

4 *Measures of Variability* **104**

<i>The Range</i>	105
<i>The Inter-Quartile Range</i>	106
<i>The Variance and the Standard Deviation</i>	108
BOX 4.1 • <i>Step-by-Step Illustration: Standard Deviation</i>	110

<i>The Raw-Score Formula for Variance and Standard Deviation</i>	111
BOX 4.2 • Step-by-Step Illustration: Variance and Standard Deviation Using Raw Scores	112
<i>Obtaining the Variance and Standard Deviation from a Simple Frequency Distribution</i>	113
BOX 4.3 • Step-by-Step Illustration: Variance and Standard Deviation of a Simple Frequency Distribution	114
<i>The Meaning of the Standard Deviation</i>	115
<i>Coefficient of Variation</i>	119
<i>Comparing Measures of Variability</i>	120
<i>Visualizing Distributions</i>	121
<i>Obtaining the Mean and Standard Deviation from a Grouped Frequency Distribution</i>	123
<i>Summary</i>	124
<i>Terms to Remember</i>	125
<i>Questions and Problems</i>	125
<i>Computer Exercises</i>	131
<i>Looking at the Larger Picture: Describing Data</i>	132

PART II • *From Description to Decision Making* 135

5 *Probability and the Normal Curve* 137

<i>Rules of Probability</i>	138
<i>Probability Distributions</i>	140
<i>The Normal Curve as a Probability Distribution</i>	146
<i>Characteristics of the Normal Curve</i>	147
<i>The Model and the Reality of the Normal Curve</i>	147
<i>The Area under the Normal Curve</i>	149
<i>Standard Scores and the Normal Curve</i>	155
<i>Finding Probability under the Normal Curve</i>	157
BOX 5.1 • Step-by-Step Illustration: Probability under the Normal Curve	158
BOX 5.2 • Step-by-Step Illustration: Finding Scores from Probability Based on the Normal Curve	163

<i>Summary</i>	164
<i>Terms to Remember</i>	164
<i>Questions and Problems</i>	165
<i>Computer Exercises</i>	172

6 *Samples and Populations* 173

<i>Errors in the Conduct of Research</i>	174
<i>Sampling Methods</i>	175
<i>Sampling Error</i>	180
<i>Sampling Distribution of Means</i>	180
<i>Standard Error of the Mean</i>	188
<i>Confidence Intervals</i>	190
BOX 6.1 • <i>Step-by-Step Illustration: 95% Confidence Interval Using z</i>	194
BOX 6.2 • <i>Step-by-Step Illustration: 99% Confidence Interval Using z</i>	195
<i>The t Distribution</i>	196
BOX 6.3 • <i>Step-by-Step Illustration: Confidence Interval Using t (small sample case)</i>	201
BOX 6.4 • <i>Step-by-Step Illustration: Confidence Interval Using t (large sample case)</i>	204
<i>Estimating Proportions</i>	206
BOX 6.5 • <i>Step-by-Step Illustration: Confidence Interval for Proportions</i>	208
<i>Summary</i>	209
<i>Terms to Remember</i>	209
<i>Questions and Problems</i>	210
<i>Computer Exercises</i>	214
<i>Looking at the Larger Picture: Generalizing from Samples to Populations</i>	215

PART III • Decision Making 217

7 *Testing Differences between Means* 219

<i>Testing the Difference between Sample and Population Means</i>	220
<i>Testing the Difference between Two Sample Means</i>	222

<i>The Null Hypothesis: No Difference between Means</i>	223
<i>The Research Hypothesis: A Difference between Means</i>	224
<i>Sampling Distribution of Differences between Means</i>	225
<i>Testing Hypotheses with the Distribution of Differences between Means</i>	228
<i>Levels of Significance</i>	232
<i>Standard Error of the Difference between Means</i>	237
<i>Testing the Difference between Means</i>	238
BOX 7.1 • <i>Step-by-Step Illustration: Test of Difference between Means</i>	239
<i>Comparing Dependent Samples</i>	244
BOX 7.2 • <i>Step-by-Step Illustration: Test of Difference between Means for Same Sample Measured Twice</i>	244
BOX 7.3 • <i>Step-by-Step Illustration: Test of Difference between Means for Matched Samples</i>	247
<i>Two-Sample Test of Proportions</i>	249
BOX 7.4 • <i>Step-by-Step Illustration: Test of Difference between Proportions</i>	250
<i>One-Tailed Tests</i>	252
BOX 7.5 • <i>Step-by-Step Illustration: One-Tailed Test of Means for Same Sample Measured Twice</i>	254
BOX 7.6 • <i>Step-by-Step Illustration: Independent Groups, One-Tailed Test</i>	256
<i>Effect Size</i>	258
BOX 7.7 • <i>Step-by-Step Illustration: Cohen's d</i>	259
<i>Requirements for Testing the Difference between Means</i>	260
<i>Summary</i>	261
<i>Terms to Remember</i>	262
<i>Questions and Problems</i>	262
<i>Computer Exercises</i>	275
8 <i>Analysis of Variance</i>	276
<hr/>	
<i>The Logic of Analysis of Variance</i>	278
<i>The Sum of Squares</i>	279
<i>Mean Square</i>	286

<i>The F Ratio</i>	287	
	BOX 8.1 • <i>Step-by-Step Illustration: Analysis of Variance</i>	290
<i>A Multiple Comparison of Means</i>	294	
	BOX 8.2 • <i>Step-by-Step Illustration: HSD for Analysis of Variance</i>	295
<i>Two-Way Analysis of Variance</i>	296	
	BOX 8.3 • <i>Step-by-Step Illustration: Two-Way Analysis of Variance</i>	303
<i>Requirements for Using the F Ratio</i>	307	
<i>Summary</i>	307	
<i>Terms to Remember</i>	308	
<i>Questions and Problems</i>	308	
<i>Computer Exercises</i>	317	
9	<i>Nonparametric Tests of Significance</i>	319
<hr/>		
<i>One-Way Chi-Square Test</i>	320	
	BOX 9.1 • <i>Step-by-Step Illustration: One-Way Chi-Square</i>	323
<i>Two-Way Chi-Square Test</i>	325	
	BOX 9.2 • <i>Step-by-Step Illustration: Two-Way Chi-Square Test of Significance</i>	331
	BOX 9.3 • <i>Step-by-Step Illustration: Comparing Several Groups</i>	335
<i>The Median Test</i>	341	
	BOX 9.4 • <i>Step-by-Step Illustration: Median Test</i>	341
<i>Mann-Whitney U Test</i>	344	
	BOX 9.5 • <i>Step-by-Step Illustration: Mann-Whitney U Test</i>	345
<i>Kruskal-Wallis Test</i>	347	
	BOX 9.6 • <i>Step-by-Step: Kruskal-Wallis Test</i>	347
<i>Summary</i>	349	
<i>Terms to Remember</i>	349	
<i>Questions and Problems</i>	349	
<i>Computer Exercises</i>	361	
<i>Looking at the Larger Picture: Testing for Differences</i>	363	

PART IV • From Decision Making to Association 365

10 Correlation 367

Strength of Correlation 367

Direction of Correlation 369

Curvilinear Correlation 370

The Correlation Coefficient 370

Pearson's Correlation Coefficient 371

BOX 10.1 • *Step-by-Step Illustration: Pearson's Correlation Coefficient* 377

The Importance of Scatter Plots 379

Partial Correlation 382

Summary 388

Terms to Remember 389

Questions and Problems 389

Computer Exercises 399

11 Regression Analysis 400

The Regression Model 401

Interpreting the Regression Line 406

Prediction Errors 408

Regression and Pearson's Correlation 413

Regression and Analysis of Variance 414

BOX 11.1 • *Step-by-Step Illustration: Regression Analysis* 416

Multiple Regression 418

Logistic Regression 428

Summary 433

Terms to Remember 434

Questions and Problems 434

Computer Exercises 442

12 *Nonparametric Measures of Correlation* 444

Spearman's Rank-Order Correlation Coefficient 444

BOX 12.1 • *Step-by-Step Illustration: Spearman's Rank-Order Correlation Coefficient* 450

Goodman's and Kruskal's Gamma 452

BOX 12.2 • *Step-by-Step Illustration: Goodman's and Kruskal's Gamma* 453

Correlation Coefficient for Nominal Data Arranged in a 2×2 Table 457

Correlation Coefficients for Nominal Data in Larger than 2×2 Tables 459

Elaboration 462

Summary 464

Terms to Remember 464

Questions and Problems 464

Computer Exercises 474

Looking at the Larger Picture: Measuring Association 475

PART V • *Applying Statistics* 477

13 *Choosing Statistical Procedures for Research Problems* 479

Research Situations 483

Research Solutions 499

APPENDIXES 509

APPENDIX A *Using SPSS and the General Social Survey* 511

APPENDIX B *StatCrunch* 532

APPENDIX C *Tables* 545

APPENDIX D *List of Formulas* 560

APPENDIX E *A Review of Some Fundamentals of Mathematics* 568

Glossary 573

Answers to Problems 578

Index 589

Preface

New to this edition:

- New discussion of Meta-Analysis in Chapter 1
- New introduction to the concept and calculation of Coefficient of Variation in Chapter 4
- New introduction to the One Sample Test of Means in Chapter 7
- New introduction to the concept and calculation of Effect Size in Chapter 7
- New discussion of the Mann-Whitney U Test and Kruskal-Wallis Test in Chapter 9
- New discussion of Elaboration in Chapter 12
- Expanded discussions of Non-Sampling Error, Standard Error, Partial Correlation, and Multicollinearity
- Updated Examples and Exercises throughout the text
- Supplementary Computer Exercises that utilize a subset of the 2010 General Social Survey

The twelfth edition of *Elementary Statistics in Social Research* provides an introduction to statistics for students in sociology, criminal justice, political science, social work and related fields. This book is not intended to be a comprehensive reference work on statistical methods. On the contrary, our first and foremost objective is to be understandable to a broad range of students, particularly those who may not have a strong background in mathematics.

Through several editions, *Elementary Statistics in Social Research* has undergone refinements and improvements in response to instructor and student feedback. For this revision, we have added several new techniques, including meta-analysis (Chapter 1), coefficient of variation (Chapter 4), one sample test of means and effect size (Chapter 7), Mann-Whitney U test and Kruskal-Wallis test (Chapter 9), logistic regression (Chapter 11), and elaboration (Chapter 12). We have expanded the discussion of several topics, including non-sampling error, standard error, partial correlation, and multicollinearity. We have updated examples and exercises throughout the text. Finally, the supplementary computer exercises in this edition utilize a subset of the 2010 General Social Survey (GSS) available for download from www.pearsonhighered.com/careers. For those instructors who do not teach with computer software, however, this feature can be easily excluded.

This edition continues to contain a number of pedagogical features. Most notably, detailed step-by-step illustrations of statistical procedures continue to be located at important points throughout the text. We have again attempted to provide clear and logical explanations for the rationale and use of statistical methods in social research. And, as in the earlier editions, we have included a number of end-of-chapter questions and problems. Students

sometimes get lost in the trees of statistics, without seeing the forest. To counter this tendency, we have ended each part of the text with a section entitled “Looking at the Larger Picture,” which carries the student through the entire research process based on hypothetical data.

Following a detailed overview in Chapter 1, the text is divided into five parts. Part One (Chapters 2 through 4) introduces the student to the most common methods for describing and comparing data. Part Two (Chapters 5 and 6) serves a transitional purpose. Beginning with a discussion of the basic concepts of probability, Part Two leads the student from the topic of the normal curve as an important descriptive device to the use of the normal curve as a basis for generalizing from samples to populations. Continuing with this decision-making focus, Part Three (Chapters 7 through 9) contains several well-known tests of significance. Part Four (Chapters 10 through 12) includes procedures for obtaining correlation coefficients and an introduction to regression analysis. Finally, Part Five consists of an important chapter (Chapter 13) in which students learn, through examples, the conditions for applying statistical procedures to various types of research problems.

The text provides students with background material for the study of statistics. An Introduction to using SPSS, StatCrunch™, a set of statistical tables, a review of basic mathematics, and a list of formulas are located in appendixes at the end of the book. Students will also find a glossary of terms and answers to the end-of-chapter problems at the end of the book.

Instructor Resources

Instructor supplements include an Instructors Manual, a solutions manual containing step-by-step calculations for the end-of-chapter problems, the General Social Survey data set, a companion statistics calculator (ABCcalc), Powerpoint slides featuring step-by-step solutions to problems, and a test generator.

To access supplementary materials online, instructors need to request an instructor access code. Go to www.pearsonhighered.com/irc, where you can register for an instructor access code. Within 48 hours of registering you will receive a confirming e-mail including an instructor access code. Once you have received your code, locate your text in the online catalog and click on the Instructor Resources button on the left side of the catalog product page. Select a supplement and a log in page will appear. Once you have logged in, you can access instructor material for all Pearson textbooks.

Acknowledgements

We are grateful to the many reviewers of this and previous editions of the text who have given us countless insightful and helpful suggestions. We also want to thank instructors who generously shared with us any errors they located in our problems and illustrations. We acknowledge the outstanding work of Chris Prener and Chris Hovey who assisted us in updating and improving problem sets, and Adam Stearn who constructed several of the book’s supplementary materials. In addition, we benefitted tremendously from the skillful editorial assistance of Jenna Savage. We also thank the National Opinion Research Council (NORC) for permission to use the 2010 General Social Survey. Finally, we note the important role of our personal computers, without “whose” assistance this book would not have been possible.

Jack Levin

James Alan Fox

David R. Forde

1

Why the Social Researcher Uses Statistics

The Nature of Social Research

Why Test Hypotheses?

The Stages of Social Research

Using Series of Numbers to Do Social Research

The Functions of Statistics

Summary

Terms to Remember

Questions and Problems

Computer Exercises

Looking at the Larger Picture: A Student Survey

A little of the social scientist can be found in all of us. Almost daily, we take educated guesses concerning the future events in our lives in order to plan for new situations or experiences. As these situations occur, we are sometimes able to confirm or support our ideas; other times, however, we are not so lucky and must face the sometimes unpleasant consequences.

Consider some familiar examples: We might invest in the stock market, vote for a political candidate who promises to solve domestic problems, play the horses, take medicine to reduce the discomfort of a cold, throw dice in a gambling casino, try to anticipate the questions on a midterm, or accept a blind date on the word of a friend.

Sometimes we win; sometimes we lose. Thus, we might make a sound investment in the stock market, but be sorry about our voting decision; win money at the craps table, but discover we have taken the wrong medicine for our illness; do well on a midterm, but have a miserable blind date; and so on. It is unfortunately true that not all of our everyday predictions will be supported by experience.

The Nature of Social Research

Similar to our everyday approach to the world, social scientists attempt to explain and predict human behavior. They also take “educated guesses” about the nature of social reality, although in a far more precise and structured manner. In the process, social scientists

examine characteristics of human behavior called *variables*—characteristics that differ or vary from one individual to another (for example, age, social class, and attitude) or from one point in time to another (for example, unemployment, crime rate, and population).

Not all human characteristics vary. It is a fact of life, for example, that the gender of the person who gave birth to you is female. Therefore, in any group of individuals, gender of mother is the *constant* “female.” A biology text would spend considerable time discussing why only females give birth and the conditions under which birth is possible, but a social scientist would consider the mother’s gender a given, one that is not worthy of study because it never varies. It could not be used to explain differences in the mental health of children because all of their mothers are females. In contrast, a mother’s age, race, and mental health are variables: In any group of individuals, they will differ from person to person and can be the key to a greater understanding of the development of the child. A researcher therefore might study differences in the mental health of children depending on the age, race, and mental health of their mothers.

In addition to specifying variables, the social researcher must also determine the *unit of observation* for the research. Usually, social scientists collect data on individual persons. For example, a researcher might conduct interviews to determine if the elderly are victimized by crime more often than younger respondents. In this case, an individual respondent is the unit to be observed by the social scientist.

However, researchers sometimes focus their research on *aggregates*—that is, on the way in which measures vary across entire collections of people. For example, a researcher might study the relationship between the average age of the population and the crime rate in various metropolitan areas. In this study, the units of observation are metropolitan areas rather than individuals.

Whether focusing on individuals or aggregates, the ideas that social scientists have concerning the nature of social reality are called *hypotheses*. These hypotheses are frequently expressed in a statement of the relationship between two or more variables: at minimum, an *independent variable* (or presumed cause) and a *dependent variable* (or presumed effect). For example, a researcher might hypothesize that socially isolated children watch more television than children who are well integrated into their peer groups, and he or she might conduct a survey in which both socially isolated and well-integrated children are asked questions regarding the time they spend watching television (social isolation would be the independent variable; TV-viewing behavior would be the dependent variable). Or a researcher might hypothesize that the one-parent family structure generates greater delinquency than the two-parent family structure and might proceed to interview samples of delinquents and nondelinquents to determine whether one or both parents were present in their family backgrounds (family structure would be the independent variable; delinquency would be the dependent variable).

Thus, not unlike their counterparts in the physical sciences, social researchers often conduct research to increase their understanding of the problems and issues in their field. Social research takes many forms and can be used to investigate a wide range of problems. Among the most useful research methods employed by social researchers for testing their hypotheses are the experiment, the survey, content analysis, participant observation, secondary analysis, and meta-analysis. For example, a researcher may conduct an experiment to determine if arresting a wife batterer will deter this behavior in the future, a sample survey

to investigate political opinions, a content analysis of values in youth magazines, a participant observation of an extremist political group, or a secondary analysis of government statistics on unemployment. Each of these research strategies is described and illustrated in this chapter.

The Experiment

Unlike everyday observation (or, for that matter, any other research approach), the *experiment* is distinguished by the degree of *control* a researcher is able to apply to the research situation. In an experiment, researchers actually manipulate one or more of the independent variables to which their subjects are exposed. The manipulation occurs when an experimenter assigns the independent variable to one group of people (called an *experimental group*), but withholds it from another group of people (called a *control group*). Ideally, all other initial differences between the experimental and control groups are eliminated by assigning subjects on a random basis to the experimental and control conditions.

For example, a researcher who hypothesizes that frustration increases aggression might assign a number of subjects to the experimental and control groups at random by flipping a coin (“heads” you’re in the experimental group; “tails” you’re in the control group), so that in all likelihood the groups do not differ initially in any major way. The researcher might then manipulate frustration (the independent variable) by asking the members of the experimental group to solve a difficult (frustrating) puzzle, whereas the members of the control group are asked to solve a much easier (nonfrustrating) version of the same puzzle. After all subjects have been given a period of time to complete their puzzle, the researcher might obtain a measure of aggression by asking them to administer “a mild electrical shock” to another subject (actually, the other subject is a confederate of the researcher who never really gets shocked, but the subjects presumably do not know this). If the willingness of subjects to administer an electrical shock is greater in the experimental group than in the control group, this difference would be attributed to the effect of the independent variable, frustration. The conclusion would be that frustration does indeed tend to increase aggressive behavior.

In 2011, Levin and Genovesi sought to investigate aspects of serial murder that contribute to its widespread fascination. In a regular classroom setting, the researchers randomly distributed to 232 male and female students various versions of a review of a motion picture featuring a serial killer’s biography. The following independent variables were experimentally manipulated: killer’s body count (half of the subjects read that 3 victims were murdered; half read that 30 victims were murdered), torture (one-third read that the killer’s torture was sexual; one-third read that the killer’s torture was physical; one-third read that the killer shot his victims without any torture), and victims’ gender (one-half read the victims were female; one-half read the victims were male). After reading a version of the review, fascination was measured, on a four-point rating scale, by asking all students to indicate “How badly do you want to see this movie when it is released to the public?” Greater fascination was indicated by a greater desire to view the film.

Results suggest that female subjects were significantly more fascinated than their male counterparts, but no significant differences were found for either body count or victims’ gender. Moreover, torture—whether sexual or physical—was significantly more

fascinating than killing alone, at least for male subjects. Findings provide only partial support for the interpretation that females are motivated more than males by a protective fascination with serial murder. The appeal of media-depicted torture may, in part, be a result of treating the sadistic and cruel punishment of victims as “forbidden fruit.” Not coincidentally, those relatively few cases of serial murder motivated by money or terror almost never receive the tremendous amount of public and media attention given to sadistic serial killers; that is, those who torture their victims.

The Survey

As we have seen, experimenters actually have a direct hand in creating the effect that they seek to achieve. By contrast, *survey* research is *retrospective*—the effects of independent variables on dependent variables are recorded *after*—and sometimes long after—they have occurred. Survey researchers typically seek to reconstruct these influences and consequences by means of verbal reports from their respondents in self-administered questionnaires, face-to-face interviews, telephone interviews, or online surveys.

Surveys lack the tight controls of experiments: Variables are not manipulated and subjects are not assigned to groups at random. As a consequence, it is much more difficult to establish cause and effect. Suppose, for instance, in a survey measuring fear of crime, that a researcher finds that respondents who had been victims of crime tend to be more fearful of walking alone in their neighborhoods than those who had not been victimized. Because the variable *victimization* was not manipulated, we cannot make the logical conclusion that victimization *causes* increased fear. An alternative explanation that the condition of their neighborhoods (poverty, for example) produces both fear among residents and crime in the streets is just as plausible.

At the same time, surveys have advantages precisely because they do not involve an experimental manipulation. As compared with experiments, survey research can investigate a much larger number of important independent variables in relation to any dependent variable. Because they are not confined to a laboratory setting in which an independent variable can be manipulated, surveys can also be more *representative*—their results can be generalized to a broader range of people.

In 2000, for example, two Stanford University researchers interested in assessing the social consequences of Internet use conducted surveys with a national sample of the adult population, including both Internet users and nonusers. Norman Nie and Lutz Erbing contacted 4,113 respondents in 2,689 households around the country and asked them to report how many hours they spend on the Internet and in various social activities.

Results obtained by Nie and Erbing consistently indicated that regular Internet users (defined as those who spend at least five hours per week on the Web) are more isolated than nonusers. More specifically, of regular Internet users, some 25% said that they spend less time with family and friends, 8% less time attending social events outside the home, and 25% less time shopping in stores. In addition, more than 25% of workers who are also regular Internet users reported that the Internet has increased the amount of time they work at the office. About the only undeniably beneficial change associated with Internet use was the finding that 14% of regular Internet users spend less time commuting in traffic!

At the August 2007 meeting of the American Sociological Association, Oregon State University's Scott Akins presented the results of a study in which he and his colleagues surveyed 6,713 adult residents of Washington State including 1,690 persons who identified themselves as Hispanic. Members of their sample were questioned about their use of illicit drugs and their ethnic identity. They were asked to indicate their marital status, educational level, socioeconomic status, and place of residence, urban versus rural.

Holding constant these other factors, Akins and his collaborators determined that illicit drug use increased among recent Hispanic immigrants as they remained longer in the United States and became more acculturated into American society. That is, to the extent that Hispanic immigrants became acculturated, they replaced their traditional cultural beliefs, language, and social patterns with those of their host society. Specifically, when asked whether they had used illicit drugs in the previous month, less than 1% of nonacculturated Hispanics indicated that they had. But 7.2% of acculturated Hispanics (not unlike 6.4% of white residents) responded in the affirmative when asked the same question about their drug use.

The Gallup polling organization employs a survey approach. On an annual basis, Gallup surveys a random sample of adult Americans regarding their views of the honesty and ethics of numerous occupations from accountants to telemarketers. During the dates November 28 to December 1, 2011, Gallup conducted telephone interviews with 1,012 adults, aged 18 and older, representing the residents of all 50 states and the District of Columbia. Gallup asked: "Please tell me how you would rate the honesty and ethical standards of people in these different fields—very high, high, average, low, or very low?"

Nurses topped the list, receiving a rating of very high or high ethical standards from some 84% of the sample. Pharmacists were in second place; 73% of the sample gave them high or very high ratings for honesty and ethics. At the other end of the spectrum, only 7% gave members of Congress and car salespeople a high or very high honesty rating, indicating their very low credibility among American adults. Even telemarketers got a slightly higher rating at 8%.

Members of Congress have long had low ratings for honesty and ethics, but never quite this low. In 2001, 22% rated their honesty as low or very low, in contrast to the results a decade later, when members of Congress received low or very low ratings from a large majority—some 64% of American adults.

Content Analysis

As an alternative to experiments and surveys, *content analysis* is a research method, whereby a researcher seeks objectively to describe the content of previously produced messages. Researchers who conduct a content analysis have no need directly to observe behavior or to question a sample of respondents. Instead, they typically study the content of books, magazines, newspapers, films, radio broadcasts, photographs, cartoons, letters, verbal dyadic interaction, political propaganda, or music.

In 2001, for example, James Alan Fox, Jack Levin, and Jason Mazaik performed a content analysis of celebrities depicted in *People* magazine cover stories. The researchers sought to determine how the celebrities chosen to be featured by the most popular

celebrity magazine in the United States (circulation: 3,552,287 per issue) had changed over almost three decades. Using appropriate coding sheets, each of the more than 1,300 covers of issues of *People* from its inception in 1974 through 1998 was scrutinized for various characteristics of the celebrity and the overall tone of the cover presentation.

Results obtained by Fox, Levin, and Mazaik indicated that the basis for *People* celebrities appearing in a cover story has, over the decades, become dramatically more negative. In 1974, during its first year of publication, less than 3% of all celebrities were featured for negative reasons, such as drug or alcohol dependence, child abuse, or the commission of a violent crime. Instead, most celebrities were on the cover because they had accomplished a positive goal—either by overcoming a personal problem or by accomplishing a career objective. By 1988 and continuing as a pattern through the 1990s, however, there was a major reversal in tone, so that almost half of all cover stories focused not on celebrities' positive accomplishments, but on their untoward characteristics. Along with musicians, athletes, and political figures, *People* for the first time also featured murderers and rapists.

In the past, various forms of communication aimed at children—children's books, video games, educational software, films, television shows, commercials, music videos, comic strips, and magazines—have been shown to be dominated by males. More specifically, such media tend to under-represent female characters who are frequently portrayed in stereotypical ways. In 2009, Black, Marola, Littman, Chrisler, and Neace published a study examining gender images found on product packaging—specifically, on the cereal boxes sitting on the kitchen table each morning as children eat their breakfast. The researchers content analyzed a sample of 217 cereal boxes—every box in the cereal aisle of a major supermarket in the northeastern United States. Each of the 1,568 characters appearing on the sampled boxes was coded for gender (female or male), species (person, animal, object), age (child/adolescent, adult), and activity level (passive, active).

Results of the content analysis were similar to those obtained in studies of gender representations in other forms of communication. On cereal boxes, male characters outnumbered female characters by more than two to one. Moreover, males were more likely than females to be depicted as authority figures (that is, adults rather than children). Overall, female characters were under-represented and portrayed as having lower status than their male counterparts.

Participant Observation

Another widely used research method is participant observation, whereby a researcher actively participates in the daily life of the people being studied, either openly in the role of a social researcher or covertly in an “undercover” role, observing events, interactions, and conversations as they happen, and questioning people over some period of time.

In 2005, sociologist Janese Free explored the influence of alternative schooling on at-risk youth. As a participant observer, Free spent approximately 750 hours over the course of a year at an alternative school in New England, volunteering as a classroom aide. The alternative school was a middle school for at-risk students who had been removed from conventional programs for reasons such as drug use, violence, truancy, or failing out of school.

Free was a semi-covert participant observer, meaning the teachers knew she was there to collect data for her research but the students did not. In her role as a “classroom aide,” she participated in the students’ daily activities at school (including classroom time, recess, gym, lunch duty), and in her role as a “researcher,” she observed and recorded the students’ behaviors and interactions in each of these settings. In an attempt to gather the most accurate data possible, Free recorded her observations in private (bathrooms, teacher’s lounge, and her car) on small pieces of paper carried in her pocket so the students would not be aware they were being observed. These field notes were transcribed into documents and then analyzed by coding the data and looking for common themes.

One of the participant observer’s guiding research questions was “How does attending an alternative school influence its students’ development, both academically and behaviorally?” Free identified six major influential factors (or “themes”) influencing students’ development, namely (1) alternative educational practices; (2) dedication of teachers and staff; (3) school culture and environment; (4) student home lives; (5) student behaviors, violence, and arrests; and (6) school disciplinary responses. Of these major influences, Free argued that alternative education practices and the dedication of teachers and staff were “enhancers” to student development, whereas the school culture and environment; student home lives; student behaviors, violence, and arrests; and school disciplinary responses were “inhibitors” to student development and progress.

Secondary Analysis

On occasion, it is possible for a social researcher not to gather his or her own data but to take advantage of data sets previously collected or assembled by others. Often referred to as *archival data*, such information comes from government, private agencies, and even colleges and universities. The social researcher is therefore not the primary or first one to analyze the data; thus, whatever he or she does to examine the data is called *secondary analysis*. This approach has an obvious advantage over firsthand data collection: It is relatively quick and easy but still exploits data that may have been gathered in a scientifically sophisticated manner. On the other hand, the researcher is limited to what is available, and has no say as to how variables are defined and measured.

The April 2008 issue of the *American Sociological Review* contains an article in which the effect of age on happiness is researched. The author, Yang Yang, a University of Chicago sociologist, conducted a secondary analysis of data from the National Opinion Research Center’s General Social Survey (GSS), which reports the results of face-to-face interviews with a representative sample of as many as 3,000 respondents collected every year between 1972 and 2004. In order to measure their degree of happiness, the GSS asked respondents: “Taken all together, how would you say things are these days—would you say that you are very happy, pretty happy, or not too happy?”

Yang found that his respondents became happier with advancing age, supporting the notion that people mature and develop in positive ways as they grow older. Moreover, happiness was greater during periods of economic prosperity.

Generations also differed in terms of degree of happiness, with baby boomers—those born between 1946 and 1964—being the least happy of any generational group. Yang’s results indicated also that among 18-year-old respondents, white women are the

TABLE 1.1 *Government Websites Containing Social Science Data*

Website/Agency	Types of Data	URL
FEDSTATS	Links to data and reports from over 70 federal agencies	www.fedstats.gov
Bureau of the Census	Population, families, business, income, housing, voting	www.census.gov
Bureau of Justice Statistics (BJS)	Crime offenders, victims, justice system	www.bjs.gov
Bureau of Labor Statistics (BLS)	Employment, unemployment, prices, wages	www.bls.gov
Bureau of Transportation Statistics (BTS)	Travel, aviation, boating, trucking, roads, highways	www.bts.gov
National Center for Health Statistics (NCHS)	Births, illness, injury, deaths, health care, nursing homes	www.cdc.gov/nchs/
National Center for Education Statistics (NCES)	Elementary, secondary, higher education	www.nces.ed.gov

happiest overall, followed by white men, black women, and black men. However, these race and gender differences almost disappear as respondents mature into old age. Apparently, happiness continues to increase even into the eighth decade of life.

In this text, we occasionally make use of archival data sources. In Chapter 2, for example, we present and analyze birth rates, homicide rates, unemployment figures, and income data drawn from various government agencies. Table 1.1 provides a list of useful government websites from which various kinds of data related to health, housing, population, crime, education, transportation, and the economy can be found and downloaded.

Meta-Analysis

An especially important use of the existing literature for conducting research can be found in the approach known as *meta-analysis*. Rather than collect new data or depend on a few unrepresentative studies to draw a conclusion, an investigator combines the results obtained in a number of previous studies that have addressed the same hypothesis and subjects all of them collectively to a single statistical test. In conducting a meta-analysis, the researcher may, with the aid of a well-supported finding, be able to end a debate in the literature or confirm the validity of an important idea. The meta-analysis process essentially consists of a literature search in which the conclusions reached by previous investigators are ignored. Instead, the researcher re-examines the data collected previously and then calculates an estimate of what is called the *effect size*—a measure of the extent to which a relationship exists in the population.

Social interdependence, where students work together cooperatively to accomplish their learning objectives, is one of the most fertile areas of research in education. When students work interdependently, they are able to succeed academically only if other students

in their group also succeed. In contrast to competitive or individualistic approaches, social interdependence has become a widely used instructional procedure from preschool through graduate school levels. It is therefore important to determine the effectiveness of social interdependence with respect to student achievement.

In 2000, Johnson, Johnson, and Stanne examined the education literature since 1970 and located 158 studies in which the comparative effect of a particular kind of social interdependence (that is, students working cooperatively) versus control methods (that is, students working alone in competition with classmates) was tested for its impact on academic achievement. Combining all of the data collected in these 158 studies, eight different forms of cooperative learning were determined by the researchers to have produced significantly higher student achievement than working alone. Based on the results of this meta-analysis, teachers should feel comfortable using a cooperative learning method in their classrooms. There is substantial evidence that social interdependence is effective.

Why Test Hypotheses?

Social science is often referred to, quite unfairly, as the study of the obvious. However, it is desirable, if not necessary, to test hypotheses about the nature of social reality, even those that seem logical and self-evident. Our everyday commonsense observations are generally based on narrow, often biased preconceptions and personal experiences. These can lead us to accept without criticism invalid assumptions about the characteristics of social phenomena and behavior.

To demonstrate how we can be so easily misled by our preconceptions and stereotypes, consider what we “know” about mass murderers—those individuals who simultaneously kill at least four victims. According to popular thinking (and media portrayals), mass murderers are typically insane individuals who go berserk or run amok, expressing their anger in a spontaneous and impulsive outpouring of aggression. Moreover, they are usually regarded as total strangers to their victims, who are unlucky enough to be in the wrong place at the wrong time—at a shopping mall, on a commuter train, or in a fast-food restaurant.

The foregoing conception of mass murderers may seem clear-cut and obvious. Yet, compiling detailed information from FBI reports about 697 mass killers over the period from 1976 to 2009, Fox and Levin found instead that mass murderers are rarely insane and spontaneous—they know exactly what they are doing and are not driven to kill by voices of demons. Random shootings in a public place are the exceptions; most mass murders occur within families or among acquaintances. Typically, mass murderers target spouses and all of their children, or bosses and their co-workers. Far from being impulsive, most mass killers are methodical and selective. They usually plan their attacks and are quite selective as to the victims they choose to kill. In an office massacre, for example, a mass killer might choose to murder only those co-workers and supervisors whom the murderer blames for losing an important promotion or getting fired.

Until recently, even criminologists all but ignored mass killings, perhaps believing that mass murder was merely a special case of homicide (albeit, by definition, yielding a larger body count), explainable by the same theories applied to single-victim incidents and