

NINTH  
EDITION

# Psychological **TESTING**

PRINCIPLES, APPLICATIONS, AND ISSUES

Robert M. Kaplan | Dennis P. Saccuzzo

# Psychological Testing

PRINCIPLES, APPLICATIONS, & ISSUES



**ROBERT M. KAPLAN**

*Stanford University*

**DENNIS P. SACCUZZO**

*San Diego State University*



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

***Psychological Testing: Principles,  
Applications, and Issues, Ninth Edition***  
**Robert M. Kaplan and Dennis P. Saccuzzo**

Product Director: Star M. Burruto

Product Manager: Carly McJunkin

Content Developer: Tangelique  
Williams-Grayer

Product Assistant: Katie Chen

Marketing Manager: James Findlay

Art and Cover Direction, Production  
Management, and Composition:  
Lumina Datamatics, Inc.

Manufacturing Planner: Karen Hunt

Cover Image: Christian Offenberg/iStock/  
Thinkstock

Unless otherwise noted all items

© Cengage Learning®

© 2018, 2013 Cengage Learning

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced or distributed in any form or by any means, except as permitted by U.S. copyright law, without the prior written permission of the copyright owner.

For product information and technology assistance, contact us at  
**Cengage Learning Customer & Sales Support, 1-800-354-9706.**

For permission to use material from this text or product,  
submit all requests online at [www.cengage.com/permissions](http://www.cengage.com/permissions).

Further permissions questions can be e-mailed to  
[permissionrequest@cengage.com](mailto:permissionrequest@cengage.com).

Library of Congress Control Number: 2016948139

Student Edition:

ISBN: 978-1-337-09813-7

**Cengage Learning**

20 Channel Center Street  
Boston, MA 02210  
USA

Cengage Learning is a leading provider of customized learning solutions with employees residing in nearly 40 different countries and sales in more than 125 countries around the world. Find your local representative at [www.cengage.com](http://www.cengage.com).

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

To learn more about Cengage Learning Solutions, visit [www.cengage.com](http://www.cengage.com).

Purchase any of our products at your local college store or at our preferred online store [www.cengagebrain.com](http://www.cengagebrain.com).

Printed in the United States of America  
Print Number: 02 Print Year: 2017

# Brief Contents

## **PART I** PRINCIPLES

- 1 Introduction 1
- 2 Norms and Basic Statistics for Testing 23
- 3 Correlation and Regression 63
- 4 Reliability 99
- 5 Validity 133
- 6 Writing and Evaluating Test Items 159
- 7 Test Administration 187

## **PART II** APPLICATIONS

- 8 Interviewing Techniques 203
- 9 Theories of Intelligence and the Binet Scales 225
- 10 The Wechsler Intelligence Scales: WAIS-IV, WISC-V, and WPPSI-IV 247
- 11 Tests for Infants, Disabilities, and Special Populations 267
- 12 Standardized Tests in Education, Civil Service, and the Military 299
- 13 Applications in Clinical and Counseling Settings 329
- 14 Projective Personality Tests 371
- 15 Computers and Basic Psychological Science in Testing 401
- 16 Testing in Counseling Psychology 425
- 17 Testing in Health Psychology and Health Care 443
- 18 Testing in Industrial and Business Settings 483

## **PART III** ISSUES

- 19 Test Bias 513
- 20 Testing and the Law 547
- 21 The Future of Psychological Testing 587



# Contents

## **PART I** PRINCIPLES

---

### **1 Introduction 1**

- Basic Concepts 6
  - What a Test Is 6
  - Types of Tests 7
- Overview of the Book 9
  - Principles of Psychological Testing 10
  - Applications of Psychological Testing 10
  - Issues of Psychological Testing 11
- Historical Perspective 11
  - Early Antecedents 11
  - Charles Darwin and Individual Differences 12
  - Experimental Psychology and Psychophysical Measurement 13
  - The Evolution of Intelligence and Standardized Achievement Tests 14
  - Personality Tests: 1920–1940 16
  - The Emergence of New Approaches to Personality Testing 18
  - The Period of Rapid Changes in the Status of Testing 20
  - The Current Environment 21
- Summary 21

### **2 Norms and Basic Statistics for Testing 23**

- Why We Need Statistics 24
- Scales of Measurement 25
  - Properties of Scales 25
  - Types of Scales 27
  - Permissible Operations 28
- Frequency Distributions 29
- Percentile Ranks 32
- Percentiles 36
- Describing Distributions 37
  - Mean 37
  - Standard Deviation 38

Z Score	40
Standard Normal Distribution	41
McCall's <i>T</i>	48
Quartiles and Deciles	50
Norms	51
Age-Related Norms	54
Tracking	54
Criterion-Referenced Tests	58
Summary	61

### **3 Correlation and Regression 63**

The Scatter Diagram	64
Correlation	66
Regression	67
The Regression Line	67
The Best-Fitting Line	68
Testing the Statistical Significance of a Correlation Coefficient	74
How to Interpret a Regression Plot	76
Other Correlation Coefficients	80
Terms and Issues in the Use of Correlation	82
Residual	82
Standard Error of Estimate	83
Coefficient of Determination	83
Coefficient of Alienation	83
Shrinkage	84
Cross Validation	84
The Correlation-Causation Problem	84
Third Variable Explanation	85
Restricted Range	86
Multivariate Analysis (Optional)	87
General Approach	87
An Example Using Multiple Regression	87
Discriminant Analysis	88
Factor Analysis	89
Summary	92
APPENDIX 3.1: Calculation of a Regression Equation and a Correlation Coefficient	93
Calculation of a Regression Equation (Data from Table 3.5)	94

### **4 Reliability 99**

History and Theory of Reliability	100
Conceptualization of Error	100

- Spearman's Early Studies 101
- Basics of Test Score Theory 101
- The Domain Sampling Model 103
- Item Response Theory 104
- Models of Reliability 106
  - Sources of Error 106
  - Time Sampling: The Test–Retest Method 107
  - Item Sampling: Parallel Forms Method 108
  - Split-Half Method 109
  - $KR_{20}$  Formula 111
  - Coefficient Alpha 112
  - Reliability of a Difference Score 113
- Reliability in Behavioral Observation Studies 115
- Connecting Sources of Error with Reliability Assessment Method 119
- Using Reliability Information 122
  - Standard Errors of Measurement and the Rubber Yardstick 122
  - How Reliable Is Reliable? 123
  - What to Do about Low Reliability 124
- Summary 129
- APPENDIX 4.1: Using Coefficient Alpha to Estimate Split-Half Reliability When the Variances for the Two Halves of the Test Are Unequal 130
- APPENDIX 4.2: The Calculation of Reliability Using  $KR_{20}$  130

## 5 Validity 133

- Defining Validity 135
- Aspects of Validity 136
  - Face Validity 136
  - Content-Related Evidence for Validity 136
  - Criterion-Related Evidence for Validity 138
  - Construct-Related Evidence for Validity 149
  - Relationship between Reliability and Validity 155
- Summary 157

## 6 Writing and Evaluating Test Items 159

- Item Writing 160
  - Item Formats 161
  - Guessing 163
  - Other Possibilities 170
- Item Analysis 173
  - Item Difficulty 173

Discriminability	175
Pictures of Item Characteristics	177
Linking Uncommon Measures	182
Items for Criterion-Referenced Tests	184
Limitations of Item Analysis	185
Summary	186

## **7 Test Administration 187**

Why We Changed Our Minds	188
The Examiner and the Subject	188
The Relationship Between Examiner and Test Taker	188
The Race of the Tester	189
Stereotype Threat	190
How Stereotype Threat Does Damage	191
Remedies for Stereotype Threat	192
Language of Test Taker	193
Training of Test Administrators	193
Expectancy Effects	193
Effects of Reinforcing Responses	196
Computer-Assisted Test Administration	197
Mode of Administration	199
Subject Variables	201
Summary	201

# **PART II APPLICATIONS**

---

## **8 Interviewing Techniques 203**

The Interview as a Test	207
Reciprocal Nature of Interviewing	208
Principles of Effective Interviewing	208
The Proper Attitudes	209
Responses to Avoid	209
Effective Responses	211
Responses to Keep the Interaction Flowing	212
Measuring Understanding	215
Mental Status Examination	217
Developing Interviewing Skills	218
Sources of Error in the Interview	219
Interview Validity	219
Interview Reliability	222
Summary	223



**9 Theories of Intelligence and the Binet Scales 225**

- The Problem of Defining Intelligence 226
- Binet's Principles of Test Construction 228
  - Principle 1: Age Differentiation 228
  - Principle 2: General Mental Ability 229
- Spearman's Model of General Mental Ability 229
  - Implications of General Mental Intelligence (*g*) 230
  - The *gf-gc* Theory of Intelligence 230
- The Early Binet Scales 231
  - The 1905 Binet-Simon Scale 231
  - The 1908 Scale 232
- Terman's Stanford-Binet Intelligence Scale 234
  - The 1916 Stanford-Binet Intelligence Scale 234
  - The Intelligence Quotient (IQ) 234
  - The 1937 Scale 235
  - The 1960 Stanford-Binet Revision and Deviation IQ (SB-LM) 237
- The Modern Binet Scale 238
  - Model for the Fourth and Fifth Editions of the Binet Scale 238
  - Characteristics of the 1986 Revision 240
  - Characteristics of the 2003 Fifth Edition 242
  - Psychometric Properties of the 2003 Fifth Edition 243
  - Median Validity 244
- Summary 244

**10 The Wechsler Intelligence Scales: WAIS-IV, WISC-V, and WPPSI-IV 247**

- The Wechsler Intelligence Scales 249
  - Point and Performance Scale Concepts 249
- From the Wechsler-Bellevue Intelligence Scale to the WAIS-IV 251
- Scales, Subtests, and Indexes 251
  - A Closer Look at Subtests 252
  - From Raw Scores to Scaled and Index Scale Scores 257
  - Index Scores 258
  - FSIQs 258
- Interpretive Features of the Wechsler Tests 259
  - Index Comparisons 259
  - Pattern Analysis 260
  - Hypothetical Case Studies 260
- Psychometric Properties of the Wechsler Adult Scale 262
  - Standardization 262
  - Reliability 263
  - Validity 263

Evaluation of the Wechsler Adult Scales	263
Downward Extensions of the WAIS-IV: The WISC-V and the WPPSI-IV	264
The WISC-V	264
The WPPSI-IV	265
Summary	266

## **11 Tests for Infants, Disabilities, and Special Populations 267**

Alternative Individual Ability Tests Compared With the Binet and Wechsler Scales	268
Alternatives Compared With One Another	270
Early Tests	272
Infant Scales	272
Major Tests for Young Children	279
General Individual Ability Tests for Handicapped and Special Populations	284
Testing Learning Disabilities	287
Visiographic Tests	292
Creativity: Torrance Tests of Creative Thinking (TTCT)	294
Individual Achievement Tests: Wide Range Achievement Test-4 (WRAT-4)	296
Summary	297

## **12 Standardized Tests in Education, Civil Service, and the Military 299**

Comparison of Group and Individual Ability Tests	301
Advantages of Individual Tests	302
Advantages of Group Tests	302
Overview of Group Tests	303
Characteristics of Group Tests	303
Selecting Group Tests	303
Using Group Tests	304
Group Tests in the Schools: Kindergarten Through 12th Grade	305
Achievement Tests Versus Aptitude Tests	305
Group Achievement Tests	305
Group Tests of Mental Abilities (Intelligence)	308
College Entrance Tests	312
The New (2016) SAT	312
The American College Test	313
Graduate and Professional School Entrance Tests	314
Graduate Record Examination Aptitude Test	314

- Miller Analogies Test 318
- The Law School Admission Test 319
- Nonverbal Group Ability Tests 321
  - Raven Progressive Matrices 322
  - Goodenough-Harris Drawing Test (G-HDT) 324
  - The Culture Fair Intelligence Test 325
  - Standardized Tests Used in the U.S. Civil Service System 326
  - Standardized Tests in the U.S. Military: The Armed Services Vocational Aptitude Battery 326
- Summary 327

## 13 Applications in Clinical and Counseling Settings 329

- Strategies of Structured Personality Test Construction 331
  - Deductive Strategies 332
  - Empirical Strategies 332
  - Criteria Used in Selecting Tests for Discussion 334
- The Logical-Content Strategy 334
  - Woodworth Personal Data Sheet 334
  - Early Multidimensional Logical-Content Scales 335
  - Mooney Problem Checklist 335
  - Criticisms of the Logical-Content Approach 336
- The Criterion-Group Strategy 336
  - Minnesota Multiphasic Personality Inventory 336
  - California Psychological Inventory (CPI)–Third Edition 347
- The Factor Analytic Strategy 349
  - Guilford’s Pioneer Efforts 349
  - Cattell’s Contribution 350
  - Problems With the Factor Analytic Strategy 352
- The Theoretical Strategy 353
  - Edwards Personal Preference Schedule (EPPS) 353
  - Personality Research Form, Third Edition (PRF-III) and Jackson Personality Inventory Revised (JPI-R) 355
  - Self-Concept 357
- Combination Strategies 358
  - Positive Personality Measurement and the NEO Personality Inventory–Three (NEO-PI-3) 358
  - The NEO Personality Inventory–Three (NEO PI-R™) 359
- Frequently Used Measures of Positive Personality Traits 362
  - Rosenberg Self-Esteem Scale 362
  - General Self-Efficacy Scale (GSE) 363
  - Ego Resiliency Scale Revised 363
  - Dispositional Resilience Scale (DRS) 363
  - Hope Scale 364

Life Orientation Test-Revised (LOT-R) 364  
 Satisfaction with Life Scale (SWLS) 365  
 Positive and Negative Affect Schedule (PANAS) 365  
 Coping Intervention for Stressful Situations (CISS) 366  
 Core Self-Evaluations 366

Future of Positive Personality Research 367

Summary 368

## 14 Projective Personality Tests 371

The Projective Hypothesis 373

The Rorschach Inkblot Test 374

Historical Antecedents 374

Stimuli, Administration, and Interpretation 375

Psychometric Properties 380

An Alternative Inkblot Test: The Holtzman 389

The Thematic Apperception Test 390

Stimuli, Administration, and Interpretation 391

Psychometric Properties 394

Alternative Apperception Procedures 395

Nonpictorial Projective Procedures 395

Word Association Test 396

Sentence Completion Tasks 396

Figure Drawing Tests 398

Summary 398

## 15 Computers and Basic Psychological Science in Testing 401

Cognitive-Behavioral Assessment Procedures Versus the Medical Model of Assessment 403

The Rationale for Cognitive-Behavioral Assessment 403

Early Procedures Based on Operant Conditioning 405

Self-Report Techniques 407

The Dysfunctional Attitude Scale 411

Irrational Beliefs Test 411

Irrational Beliefs Inventory (IBI) 412

Cognitive Functional Analysis 412

Psychophysiological Procedures 414

Physiological Variables With Treatment Implications 414

Evaluation of Psychophysiological Techniques 415

Computers and Psychological Testing 416

Computer-Assisted Interview 416

Computer-Administered Tests 417

Computer Diagnosis, Scoring, and Reporting of Results 418

- Internet Usage for Psychological Testing 419
- The Computerization of Cognitive-Behavioral Assessment 420
- Tests Possible Only by Computer 420
- Computer-Adaptive Testing 421
- Summary 423

## 16 Testing in Counseling Psychology 425

- Measuring Interests 426
  - The Strong Vocational Interest Blank 427
  - The Evolution of the Strong Measures 428
  - The Campbell Interest and Skill Survey 429
  - The Reemergence of the Strong Interest Inventory 430
  - The Kuder Occupational Interest Survey 432
  - The Career Assessment Inventory 436
  - The Self-Directed Search 436
  - Eliminating Gender Bias in Interest Measurement 437
  - Aptitudes and Interests 439
- Measuring Personal Characteristics for Job Placement 439
  - Are There Stable Personality Traits? 440
  - Other Uses of Interest Matching Methods: The Case of Internet Dating 440
- Summary 441

## 17 Testing in Health Psychology and Health Care 443

- Neuropsychological Assessment 444
  - Clinical Neuropsychology 444
  - Developmental Neuropsychology 449
  - Adult Neuropsychology 453
  - California Verbal Learning Test (CVLT) 459
  - Automated Neuropsychological Testing 462
  - Anxiety and Stress Assessment 463
  - Stress and Anxiety 464
  - The State-Trait Anxiety Inventory 464
  - Measures of Coping 466
  - Ecological Momentary Assessment 466
  - Depression 467
  - NIH Toolbox 470
- Quality-of-Life Assessment 472
  - What Is Health-Related Quality of Life? 472
  - Common Methods for Measuring Quality of Life 473
  - mHealth and New Mobile Technologies 476
  - The 2015 Medical College Admissions Test (MCAT) 477
- Summary 482

## 18 Testing in Industrial and Business Settings 483

- Personnel Psychology—The Selection of Employees 484
  - Employment Interview 484
- Base Rates and Hit Rates 486
  - Taylor-Russell Tables 489
  - Utility Theory and Decision Analysis 493
  - Value-Added Employee Assessments 495
  - Incremental Validity 499
- Personnel Psychology From the Employee's Perspective:  
Fitting People to Jobs 501
  - The Myers-Briggs Type Indicator 501
  - Tests for Use in Industry: Wonderlic Personnel Test (WPT) 502
- Measuring Characteristics of the Work Setting 503
  - Classifying Environments 503
- Job Analysis 505
- Measuring the Person–Situation Interaction 508
- Summary 511

## PART III ISSUES

---

### 19 Test Bias 513

- Why Is Test Bias Controversial? 514
- The Traditional Defense of Testing 520
  - Content-Related Evidence for Validity 521
  - Criterion-Related Sources of Bias 524
- Other Approaches to Testing Minority Group Members 529
  - Ignorance Versus Stupidity 529
- Suggestions for Solutions 531
  - Ethical Concerns and the Definition of Test Bias 531
  - Thinking Differently: Finding New Interpretations of Data 535
  - Developing Different Criteria 535
  - When Tests Harm 536
  - Does It Matter? More Testing and Less Testing 537
  - Changing the Social Environment 540
- Summary 544

### 20 Testing and the Law 547

- Laws Governing the Use of Tests 549
  - Federal Authorities 549
  - Specific Laws 553



- Federal Initiatives in Education 555
  - The Common Core 556
- Major Lawsuits That Have Affected Psychological Testing 558
  - Early Desegregation Cases 558
    - Stell v. Savannah-Chatham County Board of Education* 559
    - Hobson v. Hansen* 560
    - Diana v. State Board of Education* 561
    - Larry P. v. Wilson Riles* 561
    - Parents in Action on Special Education v. Hannon* 563
    - Crawford et al. v. Honig et al.* 564
    - Marchall v. Georgia* 568
    - Debra P. v. Turlington* 568
    - Regents of the University of California v. Bakke* 571
    - Golden Rule Insurance Company et al. v. Washburn et al.* 571
    - Adarand Constructors, Inc. v. Peña, Secretary of Transportation, et al.* 572
  - Affirmative Action in Higher Education 572
    - Grutter v. Bollinger* and *Gratz v. Bollinger* 573
    - Parents v. Seattle* 575
    - Meredith v. Jefferson County Board of Education* 576
    - Fisher v. University of Texas* 576
  - Personnel Cases 577
  - Cases Relevant to the Americans With Disabilities Act (ADA) 583
    - A Critical Look at Lawsuits 584
- Summary 585

## 21 The Future of Psychological Testing 587

- Issues Shaping the Field of Testing 588
  - Professional Issues 588
  - Moral Issues 591
  - Social Issues 594
- Current Trends 596
  - The Proliferation of New Tests 596
  - Higher Standards, Improved Technology, and Increasing Objectivity 597
  - Greater Public Awareness and Influence 598
  - The Computerization of Tests 599
  - Testing on the Internet 599
- Future Trends 599
  - Future Prospects for Testing Are as Promising as Ever Before 600
  - Controversy, Disagreement, and Change Will Continue 600
  - The Integration of Cognitive Science and Computer Science Will Lead to Several Innovations in Testing 601
- Summary 601

APPENDIX 1	Areas of a Standard Normal Distribution	603
APPENDIX 2	Critical Values of $r$ for $\alpha = .05$ and $\alpha = .01$ (Two-Tailed Test)	606
APPENDIX 3	Critical Values of $t$	607
APPENDIX 4	Code of Fair Testing Practices in Education	609
	GLOSSARY	614
	REFERENCES	618
	NAME INDEX	683
	SUBJECT INDEX	700

## List of Sample Test Profiles

---

- FIGURE 9.7 Cover page of Stanford-Binet Intelligence Scale 239
- FIGURE 12.1 Example of a score report for the Stanford Achievement Test 307
- FIGURE 12.2 A sample student profile from the ACT 313
- FIGURE 12.3 GRE verbal ability sample items 315
- FIGURE 12.4 GRE quantitative ability sample items 317
- FIGURE 12.5 MAT sample items 319
- FIGURE 13.2 An MMPI profile sheet 337
- FIGURE 13.3 An MMPI-2 profile sheet 344
- FIGURE 13.4 Jackson Personality Inventory profile sheet 356
- FIGURE 13.5 NEO Personality Inventory profile sheet 360
- TABLE 14.1 Summary of Rorschach scoring 381
- FOCUSED EXAMPLE 14.2 The danger of basing Rorschach interpretations on insufficient evidence 386–387
- Sentence completion tasks 396
- FIGURE 17.5 Profile of a patient tested with the Luria-Nebraska battery 459
- FIGURE 18.2 Sample questions from the Wonderlic 503
- FIGURE 19.8 Sample SOMPA profile 525
- TABLE 20.1 Examples of items from a minimum competence test 569

# Preface

Psychology is a broad, exciting field. Psychologists work in settings ranging from schools and clinics to basic research laboratories, pharmaceutical firms, and private international companies. Despite this diversity, all psychologists have at least two things in common: They all study behavior, and they all depend to some extent on its measurement. This book concerns a particular type of measurement, psychological tests, which measure characteristics pertaining to all aspects of behavior in human beings.

*Psychological Testing* is the result of a long-standing partnership between the authors. As active participants in the development and use of psychological tests, we became disheartened because far too many undergraduate college students view psychological testing courses as boring and unrelated to their goals or career interests. In contrast, we see psychological testing as an exciting field. It has a solid place in the history of psychology, yet it is constantly in flux because of challenges, new developments, and controversies. A book on testing should encourage, not dampen, a student's interest. Thus, we provide an overview of the many facets of psychological tests and measurement principles in a style that will appeal to the contemporary college student.

To understand the applications and issues in psychological testing, the student must learn some basic principles, which requires some knowledge of introductory statistics. Therefore, some reviewing and a careful reading of Part I will pave the way for an understanding of the applications of tests discussed in Part II. Part III examines the issues now shaping the future of testing. Such issues include test anxiety, test bias, and the interface between testing and the law. The future of applied psychology may depend on the ability of psychologists to face these challenging issues.

Throughout the book, we present a series of focused discussions and focused examples. These sections illustrate the material in the book through examples or provide a more detailed discussion of a particular issue. We also use box features called "Psychological Testing in Everyday Life" to demonstrate material such as statistical calculations.

## Increased Emphasis on Application

Students today often favor informal discussions and personally relevant examples. Consequently, we decided to use models from various fields and to write in an informal style. However, because testing is a serious and complicated field in

which major disagreements exist even among scholars and experts, we have treated the controversial aspects of testing with more formal discussion and detailed referencing.

The first edition of *Psychological Testing: Principles, Applications, and Issues* was published in 1982. The world has changed in many ways in the 35 years since the text was first introduced. For example, personal computers were new in 1982. Most students and professors had never heard of the Internet, nobody communicated by e-mail, and the inventor of Facebook had not yet been born. Nobody had even imagined smart portable phones. The first edition of *Psychological Testing* was produced on typewriters, before word processors were commonly used. At the time, few professors or students had access to private computers. The early editions of the book offered instruction for preparing the submission of statistical analyses to main-frame computers. There were far fewer applications of psychological testing than there are today. On the other hand, principles of psychological testing have remained relatively constant. Thus, newer editions have included improvements and refinements in the Principles chapters. The later chapters on Applications and Issues have evolved considerably.

Not only has the field of psychological testing changed, but so have the lives of the authors. One of us (RMK) spent most of his career as a professor in a school of medicine, eventually moved to a school of public health, then to the federal government, and back again to a school of medicine. The other (DPS) completed law school and works extensively with attorneys and the U.S. legal system on many of the applied issues discussed in this book. While maintaining our central identities as psychologists, we have also had the opportunity to explore cutting-edge practice in medicine, public health, government regulation, education, and law. The ninth edition goes further than any previous edition in spelling out the applications of psychological testing in a wide variety of applied fields.

In developing this edition, we have organized topics around the application areas. Chapter 11 considers psychological testing in education and special education. Chapter 12 looks at the use of standardized tests in education, civil service, and the military. Chapters 13 and 14 consider the use of psychological tests in clinical and counseling settings.

The age of computers has completely revolutionized psychological testing. We deal with some of these issues in the Principles chapters by discussing computer-adaptive testing and item response theory. In Chapter 15, we discuss applications of psychological science in the computer age. Chapter 16 discusses the use of psychological testing in the field of counseling psychology and focuses primarily on interest inventories. Chapter 17 explores the rapidly developing fields of psychological assessment in health psychology, medicine, and health care. Chapter 18 reviews psychological testing in industry and business settings. Several of these chapters discuss the role of new electronic technologies, such as cell phones and sensors, in the acquisition of information about human behavior.

Over the last 35 years psychological testing has faced important challenges related to fairness and to social justice. Chapter 19 takes a careful look at these controversies and attempts to spell out some of the differing perspectives in these debates. Chapter 20 focuses on legal challenges to testing practices. Ethical issues relevant to psychological tests are considered in Chapter 21.

Following a trend in our recent editions, the final chapters on issues in psychological testing have been extensively updated to reflect new developments in social justice, law, and ethics.

## Organization of the Ninth Edition: A Note to Professors for Planning

---

Producing nine editions of *Psychological Testing* over the course of more than 35 years has been challenging and rewarding. We are honored that hundreds of professors have adopted our text, and that it is now used in hundreds of colleges and universities all over the world. However, some professors have suggested that we reorganize the book to facilitate their approach to the class. To accommodate the large variety of approaches, we have tried to keep the chapters independent enough for professors to teach them in whatever order they choose. For example, one approach to the course is to go systematically through the chapter sequence.

Professors who wish to emphasize psychometric issues, however, might assign Chapters 1 through 7, followed by Chapters 19 and 20. Then, they might return to certain chapters from the Applications section. On campuses that require a strong statistics course as a prerequisite, Chapters 2 and 3 might be dropped. Professors who emphasize applications might assign Chapters 1 through 5 and then proceed directly to Part II, with some professors assigning only some of its chapters. Although Chapters 9 through 13 are most likely to be used in a basic course, we have found sufficient interest in Chapters 14 through 18 to retain them. Chapters 17 and 18 represent newer areas into which psychological testing is expanding. Finally, Chapters 19 and 20 were written so that they could be assigned either at the end of the course or near the beginning. For example, some professors prefer to assign Chapters 19 and 20 after Chapter 5.

## MindTap for Kaplan and Saccuzzo's *Psychological Testing*

---

MindTap is a personalized teaching experience with relevant assignments that guide students to analyze, apply, and improve thinking, allowing instructors to measure skills and outcomes with ease.

- ▶ **Guide Students:** A unique learning path of relevant readings, media, and activities that moves students up the learning taxonomy from basic knowledge and comprehension to analysis and application.
- ▶ **Personalized Teaching:** Becomes yours with a Learning Path that is built with key student objectives. Control what students see and when they see it. Use it as-is or match to your syllabus exactly—hide, rearrange, add and create your own content.
- ▶ **Promote Better Outcomes:** Empower instructors and motivate students with analytics and reports that provide a snapshot of class progress, time in course, engagement and completion rates.



## Supplements Beyond Compare

---

### Cognero

Cengage Learning Testing Powered by Cognero is a flexible, online system that allows you to:

- ▶ author, edit, and manage test bank content from multiple Cengage Learning solutions
- ▶ create multiple test versions in an instant
- ▶ deliver tests from your LMS, your classroom or wherever you want.

### Instructor's Resource Manual and Test Bank

The Instructor's Resource Manual (IRM) was written by Katherine Nicolai of Rockhurst University the Test Bank by TBD. The IRM includes suggestions for:

- ▶ designing your course,
- ▶ using psychological tests in your course,
- ▶ using student data to teach measurement,
- ▶ using class time,
- ▶ demonstrations, activities, and activity-based lectures.

The IRM also provides a description of integrative assignments found on the instructor's companion Web site and *unique* mock projectives and much more.

The test bank contains more than 800 multiple-choice questions in addition to many "thought" essay questions.

## Acknowledgments

---

We are highly indebted to the many reviewers and professors who provided feedback that helped shape this textbook. Special thanks go to reviewers of all editions of the text: Glen M. Adams, *Harding University*, John Dale Alden III, *Lipscomb University*, Steven Anolik, *St. Francis College*; Michael DeDonno, *Barry University*, John C. Hotz, *St. Cloud State University*, Jacqueline Massa, *Kean University*, Katherine Noll, *University of Illinois at Chicago*; Janet Panter, *Rhodes College*; and Joneis Frandele Thomas, *Howard University*; Virginia Allen, *Idaho State University*, David Bush, *Utah State University*; Ira Bernstein, *University of Texas, Arlington*; Jeff Conte, *San Diego State University*, Imogen Hall, *University of Windsor*, Maureen Hannah, *Siena College*; Ronald McLaughlin, *Juniata College*; Michael Mills, *Loyola Marymount University*, Philip Moberg, *University of Akron*; M. J. Monnot, *Central Michigan University*, Jennifer Neemann, *University of Baltimore*; Karen Obremski Brandon, *University of South Florida*; Frederick Oswald, *Michigan State University*, S. Mark Pancer, *Wilfrid Laurier University*, Christopher Ralston, *Iowa State University*, Sharon Rostosky, *University of Kentucky*, Stefan Schulenberg, *University of Mississippi*; Theresa Sparks, *Clayton State University*; Chockalingam Viswesvaran, *Florida International University*, Mark Wagner, *Wagner College*; and Nancy Zook *SUNY Purchase*.

The nine editions of this book have been developed under seven different Cengage editors. The earlier editions benefited from the patient and inspired supervision of Todd Lueders, C. Deborah Laughton, Phil Curson, Marianne Taflinger, and Jim Brace-Thompson, and Tim Matray. We are most appreciative of the support we have received from current content developer, Tangelique Williams-Grayer. She has been patient, helpful, and very well organized in directing the development of the current edition. Each of our editors has come to the task with a different personality and a different set of insights. We learned immensely from each of them and the ninth edition represents a collection of what we have gained from advice and consultations over many years. We want to give particular thanks to Kate Nicolai for preparing the student workbook for past editions, and the ninth edition online Instructor's Manual. And, we also thank the editorial and production teams, including Jennifer Ziegler, content production manager; Katie Chen, product assistant; and Sharib Asrar of Lumina Datamatics.

The ninth edition was completed while one of us (RMK) was a fellow at the Center for Advanced Studies in the Behavioral Sciences at Stanford University. The Center graciously provided office space, library services, and collegial support that greatly facilitated the timely revision of the manuscript.

*Robert M. Kaplan*  
*Dennis P. Saccuzzo*  
*September 2016*

# About the Authors



Robert M. Kaplan

**ROBERT M. KAPLAN** has served as Chief Science Officer at the US Agency for Health Care Research and Quality (AHRQ) and Associate Director of the National Institutes of Health, where he led the behavioral and social sciences programs. He is also a Distinguished Emeritus Professor of Health Services and Medicine at UCLA, where he led the UCLA/RAND AHRQ health services training program and the UCLA/RAND CDC Prevention Research Center. He was Chair of the Department of Health Services from 2004 to 2009. From 1997 to 2004 he was Professor and Chair of the Department of Family and Preventive Medicine, at the University of California, San Diego. He is a past President of several organizations, including the American Psychological Association Division of Health Psychology, Section J of the American Association for the Advancement of Science (Pacific), the International Society for Quality of Life Research, the Society for Behavioral Medicine, and the Academy of Behavioral Medicine Research. Kaplan is a former Editor-in-Chief of *Health Psychology* and of the *Annals of Behavioral Medicine*. His 20 books and over 500 articles or chapters have been cited nearly 30,000 times and the ISI includes him in the listing of the most cited authors in his field (defined as above the 99.5th percentile). Kaplan is an elected member of the National Academy of Medicine (formerly the Institute of Medicine). Dr. Kaplan is currently Regenstein Distinguished Fellow at Purdue University and Adjunct Professor of Medicine at Stanford University, where he works with Stanford's Clinical Excellence Research Center (CERC).



Robert M. Kaplan

**DENNIS P. SACCUZZO** is a professor emeritus at San Diego State University, president and co-founder of Applications of Psychology to Law, Inc., an educational corporation devoted to applying cutting-edge psychological concepts to the law, and a founding partner of Saccuzzo Johnson & Poplin, LLP, a law firm from which he uses his knowledge of testing and his legal background to fight for the rights of special education students and other vulnerable groups of individuals. He has been a scholar and practitioner of psychological testing for over 40 years. He has authored numerous peer-reviewed publications and professional presentations in the field. Dr. Saccuzzo's research has been supported by the National Science Foundation, the National Institutes of Mental Health, the National Institutes of Health, the U.S. Department of Education,

the Scottish Rite Foundation, and the U.S. armed services. He is also a California-licensed psychologist and a California-licensed attorney. He is board certified in clinical psychology by the American Board of Professional Psychology (ABPP). In addition, he is a diplomate of the American Board of Assessment Psychology (ABAP). He is a fellow of the American Psychological Association and American Psychological Society for outstanding and unusual contributions to the field of psychology. Dr. Saccuzzo is the author or co-author of over 300 peer-reviewed papers and publications, including 12 textbooks and over 20 law manuals.



# Introduction

## LEARNING OBJECTIVES

*When you have completed this chapter, you should be able to:*

- ▶ Define the basic terms pertaining to psychological and educational tests
- ▶ Distinguish between an individual test and a group test
- ▶ Define the terms *achievement*, *aptitude*, and *intelligence* and identify a concept that can encompass all three terms
- ▶ Distinguish between ability tests and personality tests
- ▶ Define the term *structured personality test*
- ▶ Explain how structured personality tests differ from projective personality tests
- ▶ Explain what a normative or standardization sample is and why such a sample is important
- ▶ Identify the major developments in the history of psychological testing
- ▶ Explain the relevance of psychological tests in contemporary society



You are sitting at a table. You have just been fingerprinted and have shown a picture ID. You look around and see over 200 nervous people. A test proctor with a stopwatch passes out booklets. You are warned not to open the booklet until told to do so; you face possible disciplinary action if you disobey. This is not a nightmare or some futuristic fantasy—this is real.

Finally, after what seems like an eternity, you are told to open your booklet to page 3 and begin working. Your mouth is dry; your palms are soaking wet. You open to page 3. You have 10 minutes to solve a five-part problem based on the following information.<sup>1</sup>

A car drives into the center ring of a circus and exactly eight clowns—Q, R, S, T, V, W, Y, and Z—get out of the car, one clown at a time. The order in which the clowns get out of the car is consistent with the following conditions:

- V gets out at some time before both Y and Q.
- Q gets out at some time after Z.
- T gets out at some time before V but at some time after R.
- S gets out at some time after V.
- R gets out at some time before W.

*Question 1.* If Q is the fifth clown to get out of the car, then each of the following could be true *except*:

- Z is the first clown to get out of the car.
- T is the second clown to get out of the car.
- V is the third clown to get out of the car.
- W is the fourth clown to get out of the car.
- Y is the sixth clown to get out of the car.
- Not quite sure how to proceed, you look at the next question.

*Question 2.* If R is the second clown to get out of the car, which of the following must be true?

- S gets out of the car at some time before T does.
- T gets out of the car at some time before W does.
- W gets out of the car at some time before V does.
- Y gets out of the car at some time before Q does.
- Z gets out of the car at some time before W does.

Your heart beats a little faster and your mind starts to freeze up. You glance at your watch and notice that 2 minutes have elapsed and you still don't have your bearings. The person sitting next to you looks a bit faint. Welcome to the world of competitive, "high stakes," standardized psychological tests. The questions you just faced were actual problems from a past version of the LSAT—the Law School Admission Test. Whether or not a student is admitted into law school in the United States is almost entirely determined by that person's score on the LSAT and undergraduate college grade point average. Thus, one's future can depend to a tremendous extent on a single score from a single test given in a tension-packed morning or afternoon. Despite

<sup>1</sup> Used by permission from the Law School Admission Test, October 2002. Answer to Question 1 is D; answer to Question 2 is E.

efforts to improve tests like the LSAT to increase diversity (Kirkland & Hansen, 2011; Pashley, Thornton, & Duffy, 2005), standardized tests tend to disadvantage women, test takers whose parents have lower incomes and levels of education, and ethnic minorities (Atkinson & Geiser, 2009).

Partly because of diversity concerns, growing numbers of 4-year colleges are not relying on the SAT test (Berger 2012; Espenshade & Chung, 2010). In 2011, the website of the National Center for Fair and Open Testing named hundreds of 4-year colleges that do not use the SAT test to admit substantial numbers of freshmen (Fair Test, 2011), and updates the list to keep its website current, <http://www.fairtest.org/university/optional>. As a result, there continues to be changes to the SAT to make it more responsive to modern realities (Wainer, 2014). Similar problems have appeared on the GRE—the Graduate Record Exam, a test that plays a major role in determining who gets to study at the graduate level in the United States. (Later in this book, we discuss how to prepare for such tests and what their significance, or predictive validity, is.) ETS, creator of the GRE General Test, recently revised the test in several significant ways. The revised GRE General Test was introduced on August 1, 2011 (<http://www.ets.org/gre/institutions/about/general>), and is now even being used to evaluate European students (Schwager, Hülshager, Lang, & Bridgeman, 2015)

Today, some careers do ride on a single test. Perhaps you have already taken the GRE or LSAT. Or perhaps you have not graduated yet but are thinking about applying for an advanced degree or professional program and will soon be facing the GRE, LSAT, or MCAT (Medical College Admission Test). Clearly, it will help you have a basic understanding of the multitude of psychological tests people are asked to take throughout their lives.

From birth, tests have a major influence on our lives. When the pediatrician strokes the palms of our hands and the soles of our feet, he or she is performing a test. When we enter school, tests decide whether we pass or fail classes. Testing may determine if we need special education. In the United States, Europe, and many other industrialized countries, competence tests determine if students will graduate from high school (Lamb, 2011; Reardon, Nicole, Allison, & Michal, 2010). More tests determine which college we may attend. And, of course, we still face more tests once we are in college.

After graduation, those who choose to avoid tests such as the GRE may need to take tests to determine where they will work. In the modern world, a large part of everyone's life and success depends on test results. Indeed, tests even have worldwide significance.

For example, 15-year-old children in 32 nations were given problems such as the following from the Organization for Economic Co-operation and Development (OECD) and the Programme for International Student Assessment (PISA) (Schleicher & Tamassia, 2000):

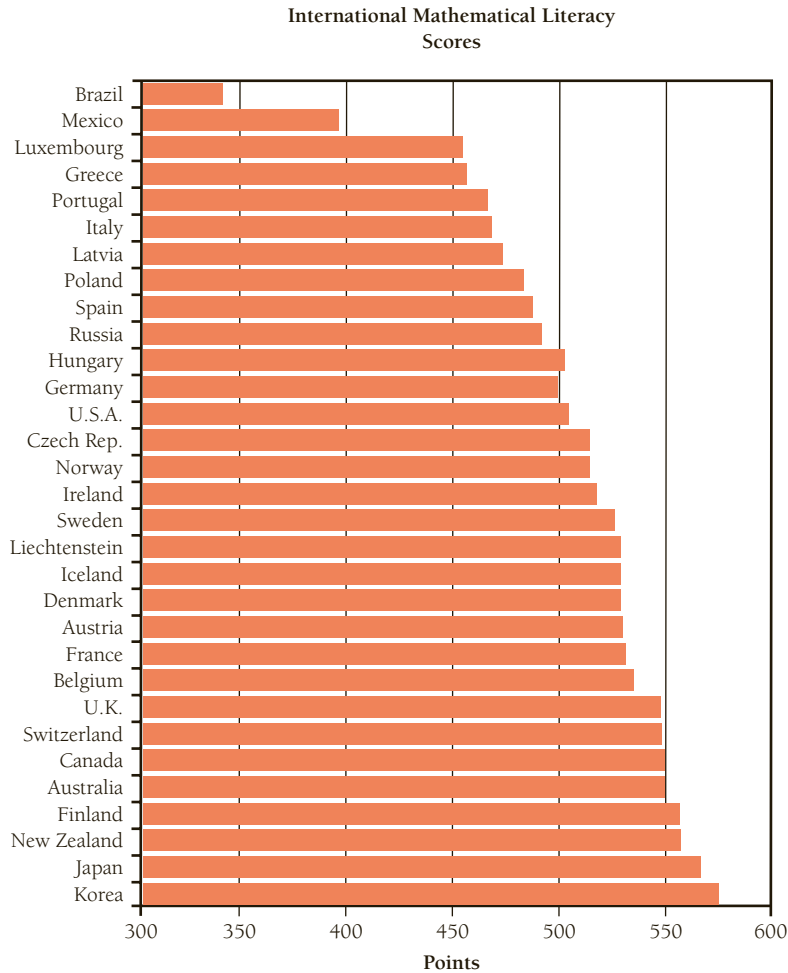
A result of global warming is that ice of some glaciers is melting.

Twelve years after the ice disappears, tiny plants, called lichen, start to grow on the rocks. Each lichen grows approximately in the shape of a circle.

The relationship between the diameter of the circles and the age of the lichen can be approximated with the formula:  $d = 7.0 \times \text{the square root of } (t - 12)$  for any/less

**FIGURE 1.1**

Approximate average scores of 15-year-old students on the OECD mathematical literacy test.



Statistics used by permission of the OECD and PISA. Figure courtesy of W. J. Koen

than or equal to 12, where  $d$  represents the diameter of the lichen in millimeters, and  $t$  represents the number of years after the ice has disappeared.

Calculate the diameter of the lichen 16 years after the ice disappeared. The complete and correct answer is:

$$d = 7.0 \times \text{the square root of } (16 - 12 \text{ mm})$$

$$d = 7.0 \times \text{the square root of } 4 \text{ mm}$$

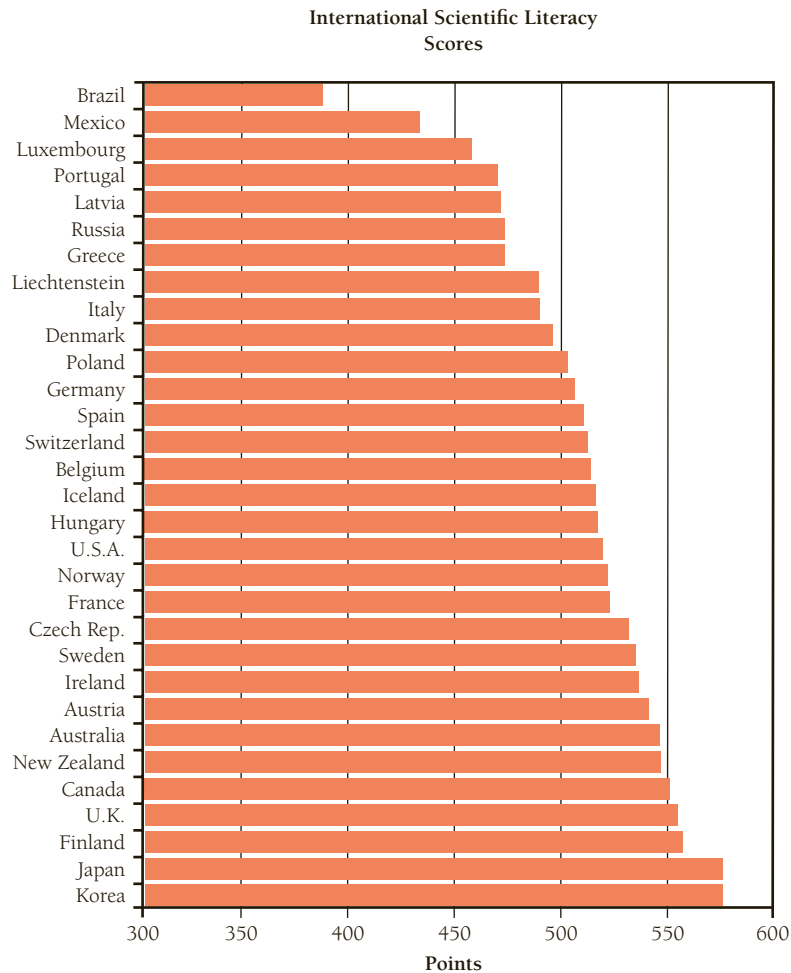
$$d = 14 \text{ mm}$$

Eighteen countries ranked above the United States in the percentage of 15-year-olds who had mastered such concepts (see Figure 1.1).

The results were similar for an OECD science literacy test (see Figure 1.2), which had questions such as the following:

A bus is moving along a straight stretch of road. The bus driver, named Ray, has a cup of water resting in a holder on the dashboard. Suddenly Ray has to slam on the brakes.

**FIGURE 1.2**  
Approximate average scores of 15-year-old students on the OECD scientific literacy test.



Statistics used by permission of the OECD and PISA. Figure courtesy of W. J. Koen

What is most likely to happen to the water in the cup immediately after Ray slams on the brakes?

- A. The water will stay horizontal.
- B. The water will spill over side 1.
- C. The water will spill over side 2.
- D. The water will spill but you cannot tell if it will spill over side 1 or side 2.

The correct answer is C.

How useful are tests such as these? Do they measure anything meaningful? How accurate are they? Such questions concern not only every U.S. citizen but also all members of the highly competitive international community. To answer them, you must understand the principles of psychological testing that you are about to learn.

To answer questions about tests, you must understand the concepts presented in this book, such as reliability, validity, item analysis, and test construction. A full

understanding of these concepts will require careful study and knowledge of basic statistics, but your efforts will be richly rewarded. When you finish this book, you will be a better consumer of tests.

## Basic Concepts

You are probably already familiar with some of the elementary concepts of psychological testing. For the sake of clarity, however, we shall begin with definitions of the most basic terms so that you will know how they are used in this textbook.

### What a Test Is

Everyone has had experience with tests. A **test** is a measurement device or technique used to quantify behavior or aid in the understanding and prediction of behavior. A spelling test, for example, measures how well someone spells or the extent to which someone has learned to spell a specific list of words. At some time during the next few weeks, your instructor will likely want to measure how well you have learned the material in this book. To accomplish this, your instructor may give you a test.

As you well know, the test your instructor gives may not measure your full understanding of the material. This is because a test measures only a sample of behavior, and error is always associated with a sampling process. Test scores are not perfect measures of a behavior or characteristic, but they do add significantly to the prediction process, as you will see.

An **item** is a specific stimulus to which a person responds overtly; this response can be scored or evaluated (e.g., classified, graded on a scale, or counted). Because psychological and educational tests are made up of items, the data they produce are explicit and hence subject to scientific inquiry.

In simple terms, items are the specific questions or problems that make up a test. The problems presented at the beginning of this chapter are examples of test items. The overt response would be to fill in or blacken one of the spaces:



A **psychological test** or educational test is a set of items that are designed to measure characteristics of human beings that pertain to behavior. There are many types of behavior. *Overt* behavior is an individual's observable activity. Some psychological tests attempt to measure the extent to which someone might engage in or "emit" a particular overt behavior. Other tests measure how much a person has previously engaged in some overt behavior. Behavior can also be *covert*—that is, it takes place within an individual and cannot be directly observed. For example, your feelings and thoughts are types of covert behavior. Some tests attempt to measure such behavior. Psychological and educational tests thus measure past or current

behavior. Some also attempt to predict future behavior, such as success in college or in an advanced degree program.

Scores on tests may be related to traits, which are enduring characteristics or tendencies to respond in a certain manner. “Determination,” sometimes seen as “stubbornness,” is an example of a trait; “shyness” is another. Test scores may also be related to the state, or the specific condition or status, of an individual. A determined individual after many setbacks may, for instance, be in a weakened state and therefore be less inclined than usual to manifest determination. Tests measure many types of behavior.

What does it mean when someone gets 75 items correct on a 100-item test? One thing it means, of course, is that 75% of the items were answered correctly. In many situations, however, knowing the percentage of correct items a person obtained can be misleading. Consider two extreme examples. In one case, out of 100 students who took the exam, 99 had 90% correct or higher, and 1 had 75% correct. In another case, 99 of the 100 students had scores of 25% or lower, while 1 had 75% correct. The meaning of the scores can change dramatically, depending on how a well-defined sample of individuals scores on a test. In the first case, a score of 75% is poor because it is in the bottom of the distribution; in the second case, 75% is actually a top score. To deal with such problems of interpretation, psychologists make use of **scales**, which relate raw scores on test items to some defined theoretical or empirical distribution. Later in the book you will learn about such distributions.

## Types of Tests

Just as there are many types of behavior, so there are many types of tests. Those that can be given to only one person at a time are known as **individual tests** (see Figure 1.3). The examiner or **test administrator** (the person giving the test) gives

**FIGURE 1.3**  
An individual test administration.



Steve Debenport/Getty Images



the test to only one person at a time, the same way that psychotherapists see only one person at a time. A **group test**, by contrast, can be administered to more than one person at a time by a single examiner, such as when an instructor gives everyone in the class a test at the same time.

One can also categorize tests according to the type of behavior they measure. Ability tests contain items that can be scored in terms of speed, accuracy, or both. On an ability test, the faster or the more accurate your responses, the better your scores on a particular characteristic. The more algebra problems you can correctly solve in a given amount of time, the higher you score in ability to solve such problems.

Historically, experts have distinguished among achievement, aptitude, and intelligence as different types of ability. **Achievement** refers to previous learning. A test that measures or evaluates how many words you can spell correctly is called a *spelling achievement test*. **Aptitude**, by contrast, refers to the potential for learning or acquiring a specific skill. A spelling aptitude test measures how many words you might be able to spell given a certain amount of training, education, and experience. Your musical aptitude refers in part to how well you might be able to learn to play a musical instrument given a certain number of lessons. Traditionally distinguished from achievement and aptitude, **intelligence** refers to a person's general potential to solve problems, adapt to changing circumstances, think abstractly, and profit from experience. When we say a person is "smart," we are usually referring to intelligence. When a father scolds his daughter because she has not done as well in school as she can, he most likely believes that she has not used her intelligence (general potential) to achieve (acquire new knowledge).

The distinctions among achievement, aptitude, and intelligence are not always so cut-and-dried because all three are highly interrelated. Attempts to separate prior learning from potential for learning, for example, have not succeeded. In view of the considerable overlap of achievement, aptitude, and intelligence tests, all three concepts are encompassed by the term **human ability**.

There is a clear-cut distinction between ability tests and personality tests. Whereas ability tests are related to capacity or potential, **personality tests** are related to the overt and covert dispositions of the individual—for example, the tendency of a person to show a particular behavior or response in a given situation. Remaining isolated from others, for instance, does not require any special skill or ability, but some people typically prefer or tend to remain thus isolated. Personality tests measure typical behavior.

There are several types of personality tests. In Chapter 13, you will learn about structured, or objective, personality tests. **Structured personality tests** provide a statement, usually of the "self-report" variety, and require the subject to choose between two or more alternative responses such as "True" or "False" (see Figure 1.4).

In contrast to structured personality tests, projective personality tests are unstructured. In a **projective personality test**, either the stimulus (test materials) or the required response—or both—are ambiguous. For example, in the highly controversial Rorschach test, the stimulus is an inkblot. Furthermore, rather than being asked to choose among alternative responses, as in structured personality tests, the individual is asked to provide a spontaneous response. The inkblot is presented

**FIGURE 1.4**  
Self-report test items.

	True	False
1. I like heavy metal music.	<input type="checkbox"/>	<input type="checkbox"/>
2. I believe that honesty is the best policy.	<input type="checkbox"/>	<input type="checkbox"/>
3. I am in good health.	<input type="checkbox"/>	<input type="checkbox"/>
4. I am easily fatigued.	<input type="checkbox"/>	<input type="checkbox"/>
5. I sleep well at night.	<input type="checkbox"/>	<input type="checkbox"/>

**TABLE 1.1** Types of Tests

- I. **Ability tests:** Measure skills in terms of speed, accuracy, or both.
  - A. **Achievement:** Measures previous learning.
  - B. **Aptitude:** Measures potential for acquiring a specific skill.
  - C. **Intelligence:** Measures potential to solve problems, adapt to changing circumstances, and profit from experience.
- II. **Personality tests:** Measure typical behavior—traits, temperaments, and dispositions.
  - A. **Structured (objective):** Provides a self-report statement to which the person responds “True” or “False,” “Yes” or “No.”
  - B. **Projective:** Provides an ambiguous test stimulus; response requirements are unclear.

to the subject, who is asked, “What might this be?” Projective tests assume that a person’s interpretation of an ambiguous stimulus will reflect his or her unique characteristics (see Chapter 14).

See Table 1.1 for a brief overview of ability and personality tests.

**Psychological testing** refers to all the possible uses, applications, and underlying concepts of psychological and educational tests. The main use of these tests, though, is to evaluate individual differences or variations among individuals. Such tests measure individual differences in ability and personality and assume that the differences shown on the test reflect actual differences among individuals. For instance, individuals who score high on an IQ test are assumed to have a higher degree of intelligence than those who obtain low scores. Thus, the most important purpose of testing is to differentiate among those taking the tests. We shall discuss the idea of individual differences later in this chapter.

## Overview of the Book

This book is divided into three parts: *Principles*, *Applications*, and *Issues*. Together, these parts cover psychological testing from the most basic ideas to the most complex. Basic ideas and events are introduced early and stressed throughout to reinforce