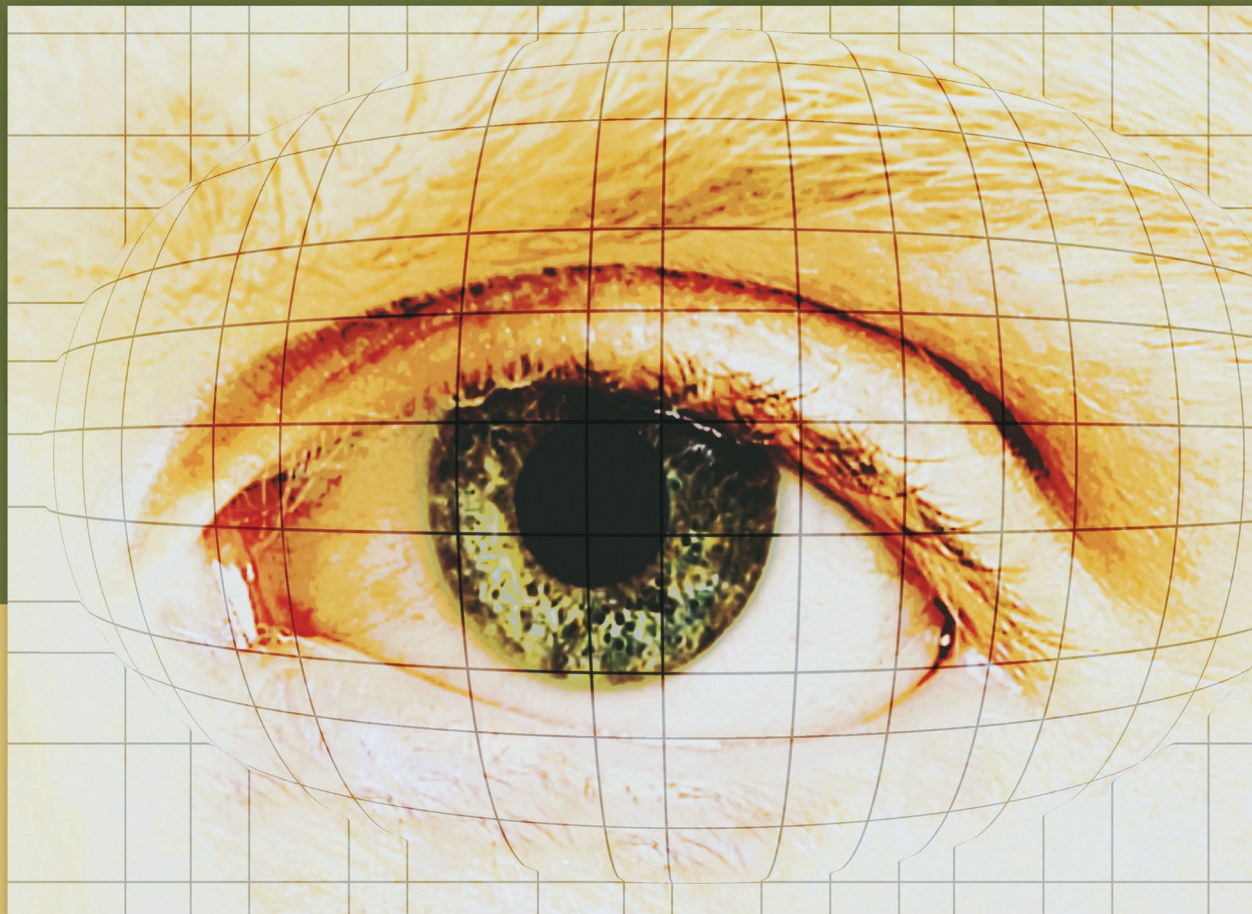


Ronald Jay Cohen | Mark E. Swerdlik | Edward D. Sturman



Psychological Testing and Assessment

An Introduction to Tests and Measurement

Eighth Edition

Psychological Testing and Assessment: A Timeline Spanning 2200 B.C.E to the Present

Note: This is a brief, decidedly *non*comprehensive overview of historical events perceived important by the authors. Consult other authoritative historical sources for more detailed and comprehensive descriptions of these and other events.

2200 B.C.E.

Proficiency testing is known to have been conducted in China. The Emperor has public officials evaluated periodically.

1115 B.C.E.

Open and competitive civil service examinations in China are common during the Chang Dynasty. Proficiency is tested in areas such as arithmetic, writing, geography, music, agriculture, horsemanship, and cultural rites and ceremonies.

400 B.C.E.

Plato suggests that people should work at jobs consistent with their abilities and endowments—a sentiment that will be echoed many times through the ages by psychologists, human resource professionals, and parents.

175 B.C.E.

Claudius Galenus (otherwise known as Galen) designs experiments to show that the brain, not the heart, is the seat of the intellect.

200

The so-called Dark Ages begin and society forces science to take a (temporary) backseat to faith and superstition.

1484

Interest in individual differences centers primarily on question such as “Who is in league with Satan?” and “Are they in *voluntary* or *involuntary* league?” *The Hammer of Witches* is a primitive, diagnostic manual of sorts with tips on interviewing and identifying persons suspected of having strayed from the righteous path.

1550

The Renaissance witnesses a rebirth in philosophy, and German physician Johann Weyer writes that those accused of being witches may have been suffering from mental or physical disorders. For the faithful, Weyer is seen as advancing Satan’s cause.

1600

The pendulum begins to swing away from a religion-dominated view of the world to one that is more philosophical and scientific in nature.

1700

The cause of philosophy and science is advanced with the writings of the French philosopher René Descartes, the German philosopher Gottfried Leibniz, and a group of English philosophers (John Locke, George Berkeley, Dave Hume, and David Hartley) referred to collectively as “the British empiricists.” Descartes, for example, raised intriguing questions regarding the relationship between the mind and the body. These issues would be explored in a less philosophical and more physical way by Pierre Cabanis, a physiologist. For humanitarian purposes, Cabanis personally observed the state of consciousness of guillotine victims of the French Revolution. He concluded that the mind and body were so intimately linked that the guillotine was probably a painless mode of execution.

1734

Christian von Wolff authors two books, *Psychologia Empirica* [*Empirical Psychology*] (1732) and *Psychologia Rationalis*

[*Rational Psychology*] (1734), which anticipate psychology as a science. A student of Gottfried Leibniz, von Wolff also elaborated on Leibniz’s idea that there exist perceptions below the threshold of awareness, thus anticipating Freud’s notion of the unconscious.

1780

Franz Mesmer “mesmerizes” not only Parisian patients but some members of the European medical community with his use of what he once referred to as “animal magnetism” to effect cures. Mesmerism (or *hypnosis* as we know it today) would go on to become a tool of psychological assessment; the hypnotic interview is one of many alternative techniques for information gathering.

1823

The *Journal of Phrenology* is founded to further the study of Franz Joseph Gall’s notion that ability and special talents are localized in concentrations of brain fiber that press outward. Extensive experimentation eventually discredits phrenology, and the journal folds by the early twentieth century. By the mid-twentieth century, evaluation of “bumps” in paper profiles would be preferable to examination of bumps on the head for obtaining information about ability and talents.

1829

In *Analysis of the Phenomena of the Human Mind*, English philosopher James Mill argued that the structure of mental life consists of sensation and ideas. Mill anticipates an approach to experimental psychology called *structuralism*, the goal of which would be to explore the components of the structure of the mind.

1848

In Vermont, an accidental discharge of explosives sends a three-foot iron rod through the skull of railway construction foreman Phineas Gage, destroying much of the front part of the left side of his brain. With medical intervention, Gage survives. However, once viewed as a competent and capable worker, after the accident he is seen as fitful, irreverent, and “no longer Gage.” Because his intellect seemed unaffected, the case was significant for calling attention to the role of the brain in personality and its assessment.

1859

The publication of Charles Darwin’s *On the Origin of Species by Means of Natural Selection* advances the then-radical notion that humans descended from apes. The work raises questions about how animals and humans compare with regard to variables such as state of consciousness. Darwin also writes of natural selection and the survival of the fittest of the species. These ideas may have greatly influenced Freud, whose psychoanalytic theory of personality emphasized the importance of instinctual sexual and aggressive urges.

1860

The German physiologist Gustav Fechner publishes *Elements of Psychophysics*, in which he explored the way people respond to stimuli such as light and sound. The work prompts experimentation in the areas of human and animal perception.

1869

Sir Francis Galton, half-cousin to Charles Darwin, publishes *Hereditary Genius*, which was noteworthy both for (a) its claim that genius is inherited, and (b) its pioneering use of

the statistical technique that Karl Pearson would later call *correlation*. Galton would subsequently make numerous and varied contributions to measurement with his inventions and innovations.

1879

Wilhelm Max Wundt founded the first experimental psychology clinic in Leipzig, Germany; psychology is a science in its own right, not simply a branch of philosophy. A structuralist, Wundt relies heavily on a tool of assessment called *introspection* (wherein subjects verbally try to faithfully describe their conscious experience of a stimulus). The structuralists focus attention on sensory-related abilities and reaction time.

1885

Herman Ebbinghaus publishes *Memory: A Contribution to Experimental Psychology*, in which he describes his use of nonsense syllables to research and evaluate human memory. His many keen insights on learning (and forgetting) curves proves that higher-order mental processes such as memory—not just reaction time or sensory reaction to stimuli—can be effectively assessed.

1890

American psychologist James McKeen Cattell coins the term *mental test* in a publication. He would go on to found several publications, most notably *Science* and *Psychological Review*. In 1921, he formed the Psychological Corporation with the goal of “useful applications of psychology.” Also in 1890, New York became the first state to assume responsibility for its mentally ill citizens. Related legislation changed the name of so called “lunatic asylums” to state hospitals—the place where the indigent mentally ill would be afforded medically supervised evaluation and treatment.

1892

Psychiatrist Emil Kraepelin, who studied with Wundt, publishes research that employed a word association test. Also in 1892, the American Psychological Association (APA) is founded with 31 members, thanks primarily to the efforts of its first president, G. Stanley Hall. For a fascinating account, see Samuel Willis Fernberger’s article, “The American Psychological Association: 1892–1942” in the January 1943 issue (volume 50) of the *Psychological Review*.

1895

Alfred Binet and Victor Henri publish articles calling for the measurement of cognitive abilities such as memory, as well as other human abilities such as social comprehension. Interestingly, Binet also wondered aloud about the possible uses of inkblots to study personality.

1896

Lightner Witmer establishes the first psychological clinic in the United States, at the University of Pennsylvania. Subsequently, in 1907, Witmer founded a journal called *Psychological Clinic*. Witmer wrote “Clinical Psychology,” the first article in that journal.

1904

Charles Spearman, a student of Wundt at Leipzig, begins to lay the foundation for the concept of test reliability. He also begins building the mathematical framework for factor analysis.

1905

Alfred Binet and Theodore Simon publish a 30-item “measuring scale of intelligence” designed to help identify mentally retarded Paris schoolchildren. The notion of

measuring intelligence would find a receptive worldwide audience.

1910

How is your handwriting? If you were a student at this time, you might have had it checked by one of the first standardized tests ever—a test authored by E. L. Thorndike. His article entitled “Handwriting” (*Teachers College Record*, volume 11, issue 2) provides 16 handwriting samples arranged in order of merit.

1912

This was the year that the now-familiar term “IQ” (intelligence quotient) came into being. William Stern devised a formula whereby “mental age” as determined by Binet’s test was the dividend, the divisor was the testtaker’s chronological age, and the quotient, multiplied by 100 was the IQ. Although “IQ” remains a fixture in the world’s vocabulary, contemporary measures of intelligence are no longer devised by such ratios. Also in 1912, another IQ-related milestone (of sorts): Henry Herbert Goddard’s book, *The Kallikak Family: A Study in the Heredity of Feeble-Mindedness*, was published. Goddard’s own life and controversial career is presented in our *Close-up* in Chapter 2.

1913

Swiss psychiatrist Hermann Rorschach, the son of an art teacher, publishes papers on how analysis of patients’ artwork can provide insights into personality. In 1921 his now-famous monograph, *Psychodiagnostics*, would evolve into a test that has become an icon for psychological tests in the public eye: the Rorschach Inkblot test. Also in 1913, John Watson’s now-famous *Psychological Review* article “Psychology as the Behaviorist Views It” becomes known as the “behaviorist manifesto.” Of course, as the behaviorist views it, behavioral observation is a key tool of assessment.

1914

World War I serves as a boon to the testing movement since thousands of recruits must be quickly screened for intellectual functioning, as well as emotional fitness.

1916

After years of research, Lewis M. Terman, working at Stanford University, publishes the Stanford Revision of the Binet-Simon Intelligence Scale. This American adaptation and revision of the test first developed in France would become widely known as the Stanford-Binet.

1920

Army Mental Tests, edited by then Majors Clarence S. Yoakum and Robert M. Yerkes (both psychologists with distinguished careers), is published by Holt. This edited volume provides in great detail information about the Army Alpha and Beta tests developed during the “great war” at a time “in this supreme struggle [when] it became clear . . . that the proper utilization of man power, and more particularly of mind or brain power, would assure ultimate victory” (p. vii).

1926

The College Board sponsors the development of the Scholastic Aptitude Test (SAT) and administers the test for the first time.

(continued on inside back cover)

Psychological Testing and Assessment

An Introduction to Tests and Measurement

EIGHTH EDITION

Ronald Jay Cohen

RJ COHEN CONSULTING

Mark E. Swerdlik

ILLINOIS STATE UNIVERSITY

Edward D. Sturman

STATE UNIVERSITY OF NEW YORK—PLATTSBURGH

ONLINE LEARNING CENTER
www.mhhe.com/cohentesting8





PSYCHOLOGICAL TESTING AND ASSESSMENT:
AN INTRODUCTION TO TESTS AND MEASUREMENT, EIGHTH EDITION

Published by McGraw-Hill, a business unit of The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY, 10020. Copyright © 2013 by The McGraw-Hill Companies, Inc. All rights reserved. Printed in the United States of America. Previous editions © 2010, 2005, and 2002. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of The McGraw-Hill Companies, Inc., including, but not limited to, in any network or other electronic storage or transmission, or broadcast for distance learning.

Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 0 QDB/QDB 1 0 9 8 7 6 5 4 3 2

ISBN: 978-0-07-803530-2

MHID: 0-07-803530-9

Senior Vice President, Products & Markets: *Kurt L. Strand*

Vice President, General Manager, Products & Markets: *Michael Ryan*

Vice President, Content Production & Technology Services: *Kimberly Meriwether David*

Publisher: *Michael Sugarman*

Senior Sponsoring Editor: *Debra B. Hash*

Marketing Coordinator: *Angela R. FitzPatrick*

Project Manager: *Melissa M. Leick*

Buyer: *Nicole Baumgartner*

Design Coordinator: *Brenda A. Rolwes*

Cover Image: *Ingram Publishing RF*

Media Project Manager: *Sridevi Palani*

Typeface: *10/12 Palatino Roman*

Compositor: *Laserwords Private Limited*

Printer: *Quad/Graphics*

All credits appearing on page or at the end of the book are considered to be an extension of the copyright page.

Library of Congress Cataloging-in-Publication Data

Cohen, Ronald Jay.

Psychological testing and assessment : an introduction to tests and measurement / Ronald Jay Cohen, Mark E. Swerdlik, Edward Sturman. — 8th ed.

p. cm.

ISBN 978-0-07-803530-2 (alk. paper)

1. Psychological tests. 2. Psychometrics. I. Swerdlik, Mark E.

II. Sturman, Edward. III. Title.

BF176.C63 2013

150.28'7—dc23

2012012933

The Internet addresses listed in the text were accurate at the time of publication. The inclusion of a website does not indicate an endorsement by the authors or McGraw-Hill, and McGraw-Hill does not guarantee the accuracy of the information presented at these sites.

This book is dedicated with love to the memory of Edith and Harold Cohen.



Contents

Preface xiii

PART I | *An Overview*

1 Psychological Testing and Assessment 1

TESTING AND ASSESSMENT 1

Psychological Testing and Assessment Defined 1

THE TOOLS OF PSYCHOLOGICAL ASSESSMENT 6

The Test 6

The Interview 9

The Portfolio 10

Case History Data 11

Behavioral Observation 12

Role-Play Tests 14

Computers as Tools 14

Other Tools 17

WHO, WHAT, WHY, HOW, AND WHERE? 18

Who Are the Parties? 18

In What Types of Settings Are Assessments Conducted, and Why? 22

How Are Assessments Conducted? 27

Where to Go for Authoritative Information: Reference Sources 33

MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Stephen Finn 5

CLOSE-UP Should Observers Be Parties to the Assessment Process? 20

EVERYDAY PSYCHOMETRICS Everyday Accommodations 32

SELF-ASSESSMENT 36

2 Historical, Cultural, and Legal/Ethical Considerations 38

A HISTORICAL PERSPECTIVE 38

Antiquity to the Nineteenth Century 38

The Twentieth Century 43

CULTURE AND ASSESSMENT 45

Evolving Interest in Culture-Related Issues 46

Some Issues Regarding Culture and Assessment 51

Tests and Group Membership 55

LEGAL AND ETHICAL CONSIDERATIONS 58

The Concerns of the Public 58

The Concerns of the Profession 66

The Rights of Testtakers 72

CLOSE-UP The Controversial Career of Henry Herbert Goddard	48
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Nathaniel V. Mohatt	54
EVERYDAY PSYCHOMETRICS Life-or-Death Psychological Assessment	69
SELF-ASSESSMENT	76

PART || *The Science of Psychological Measurement*

3 A Statistics Refresher 77

SCALES OF MEASUREMENT	78
Nominal Scales	79
Ordinal Scales	80
Interval Scales	81
Ratio Scales	81
Measurement Scales in Psychology	82
DESCRIBING DATA	83
Frequency Distributions	83
Measures of Central Tendency	89
Measures of Variability	92
Skewness	96
Kurtosis	97
THE NORMAL CURVE	98
The Area Under the Normal Curve	99
STANDARD SCORES	102
z Scores	102
T Scores	103
Other Standard Scores	103
CORRELATION AND INFERENCE	106
The Concept of Correlation	106
The Pearson r	107
The Spearman Rho	110
Graphic Representations of Correlation	111
Meta-Analysis	115
EVERYDAY PSYCHOMETRICS Consumer (of Graphed Data), Beware!	88
CLOSE-UP The Normal Curve and Psychological Tests	100
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Benoît Verdon	116
SELF-ASSESSMENT	116

4 Of Tests and Testing 118

SOME ASSUMPTIONS ABOUT PSYCHOLOGICAL TESTING AND ASSESSMENT	118
Assumption 1: Psychological Traits and States Exist	118
Assumption 2: Psychological Traits and States Can Be Quantified and Measured	120

Assumption 3: Test-Related Behavior Predicts Non-Test-Related Behavior	122
Assumption 4: Tests and Other Measurement Techniques Have Strengths and Weaknesses	122
Assumption 5: Various Sources of Error Are Part of the Assessment Process	122
Assumption 6: Testing and Assessment Can Be Conducted in a Fair and Unbiased Manner	123
Assumption 7: Testing and Assessment Benefit Society	123
WHAT'S A "GOOD TEST"?	124
Reliability	124
Validity	125
Other Considerations	125
NORMS	128
Sampling to Develop Norms	129
Types of Norms	135
Fixed Reference Group Scoring Systems	139
Norm-Referenced Versus Criterion-Referenced Evaluation	139
CULTURE AND INFERENCE	142
EVERYDAY PSYCHOMETRICS Putting Tests to the Test	126
CLOSE-UP How "Standard" Is <i>Standard</i> in Measurement?	130
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Steve Julius and Dr. Howard W. Atlas	142
SELF-ASSESSMENT	143
5 Reliability	145
THE CONCEPT OF RELIABILITY	145
Sources of Error Variance	147
RELIABILITY ESTIMATES	150
Test-Retest Reliability Estimates	150
Parallel-Forms and Alternate-Forms Reliability Estimates	151
Split-Half Reliability Estimates	152
Other Methods of Estimating Internal Consistency	154
Measures of Inter-Scorer Reliability	159
USING AND INTERPRETING A COEFFICIENT OF RELIABILITY	160
The Purpose of the Reliability Coefficient	160
The Nature of the Test	161
The True Score Model of Measurement and Alternatives to It	164
RELIABILITY AND INDIVIDUAL SCORES	173
The Standard Error of Measurement	175
The Standard Error of the Difference Between Two Scores	178
EVERYDAY PSYCHOMETRICS The Reliability Defense and the Breathalyzer Test	165
CLOSE-UP Item Response Theory (IRT)	170
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Bryce B. Reeve	174
SELF-ASSESSMENT	180

6 Validity 181

THE CONCEPT OF VALIDITY	181
Face Validity	183
Content Validity	184
CRITERION-RELATED VALIDITY	190
What Is a Criterion?	190
Concurrent Validity	191
Predictive Validity	192
CONSTRUCT VALIDITY	198
Evidence of Construct Validity	199
VALIDITY, BIAS, AND FAIRNESS	204
Test Bias	204
Test Fairness	206
MEET AN ASSESSMENT PROFESSIONAL	Meet Dr. Adam Shoemaker 185
CLOSE-UP	Base Rates and Predictive Validity 193
EVERYDAY PSYCHOMETRICS	Adjustment of Test Scores by Group Membership: Fairness in Testing or Foul Play? 208
SELF-ASSESSMENT	210

7 Utility 211

WHAT IS UTILITY?	212
Factors That Affect a Test's Utility	212
UTILITY ANALYSIS	218
What Is a Utility Analysis?	218
How Is a Utility Analysis Conducted?	219
Some Practical Considerations	232
METHODS FOR SETTING CUT SCORES	235
The Angoff Method	236
The Known Groups Method	236
IRT-Based Methods	237
Other Methods	238
EVERYDAY PSYCHOMETRICS	Rethinking the "Costs" of Testing— and of Not Testing 214
MEET AN ASSESSMENT PROFESSIONAL	Meet Dr. Erik Viirre 216
CLOSE-UP	Utility Analysis: An Illustration 220
SELF-ASSESSMENT	239

8 Test Development 240

TEST CONCEPTUALIZATION	241
Some Preliminary Questions	241
Pilot Work	243
TEST CONSTRUCTION	244
Scaling	244
Writing Items	251
Scoring Items	260

TEST TRYOUT	261
What Is a Good Item?	262
ITEM ANALYSIS	262
The Item-Difficulty Index	263
The Item-Reliability Index	264
The Item-Validity Index	265
The Item-Discrimination Index	265
Item-Characteristic Curves	268
Other Considerations in Item Analysis	269
Qualitative Item Analysis	272
TEST REVISION	275
Test Revision as a Stage in New Test Development	275
Test Revision in the Life Cycle of an Existing Test	276
The Use of IRT in Building and Revising Tests	280
EVERYDAY PSYCHOMETRICS	Psychometrics in the Classroom 245
CLOSE-UP	Designing an Item Bank 257
MEET AN ASSESSMENT PROFESSIONAL	Meet Dr. Scott Birkeland 269
SELF-ASSESSMENT	284

PART  *The Assessment of Intelligence*

9 Intelligence and Its Measurement 285

WHAT IS INTELLIGENCE?	285
Intelligence Defined: Views of the Lay Public	286
Intelligence Defined: Views of Scholars and Test Professionals	287
Factor-Analytic Theories of Intelligence	291
The Information-Processing View	295
MEASURING INTELLIGENCE	297
Types of Tasks Used in Intelligence Tests	297
Theory in Intelligence Test Development and Interpretation	299
INTELLIGENCE: SOME ISSUES	301
Nature versus Nurture	301
The Stability of Intelligence	304
The Construct Validity of Tests of Intelligence	306
Other Issues	306
A PERSPECTIVE	318
MEET AN ASSESSMENT PROFESSIONAL	Meet Dr. Barbara C. Pavlo 298
EVERYDAY PSYCHOMETRICS	Being Gifted 307
CLOSE-UP	Culture-Fair, Culture-Loaded 315
SELF-ASSESSMENT	319

10 Tests of Intelligence 320

THE STANFORD-BINET INTELLIGENCE SCALES	323
The Stanford-Binet Intelligence Scales: Fifth Edition	324

THE WECHSLER TESTS	329
The Wechsler Adult Intelligence Scale–Fourth Edition (WAIS-IV)	330
The Wechsler Intelligence Scale for Children–Fourth Edition (WISC-IV)	336
The Wechsler Preschool and Primary Scale of Intelligence–Third Edition (WPPSI-III)	338
Short Forms	341
The Wechsler Tests in Perspective	342
OTHER MEASURES OF INTELLIGENCE	346
Tests Designed for Individual Administration	346
Tests Designed for Group Administration	347
Measures of Cognitive Style and Specific Intellectual Abilities	353
CLOSE-UP Factor Analysis	343
EVERYDAY PSYCHOMETRICS The Armed Services Vocational Aptitude Battery (ASVAB): A Test You Can Take	350
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Rebecca Anderson	353
SELF-ASSESSMENT	356

11 Assessment for Education 357

THE ROLE OF TESTING AND ASSESSMENT IN EDUCATION	357
Response to Intervention (RtI)	358
Dynamic Assessment	361
ACHIEVEMENT TESTS	363
Measures of General Achievement	363
Measures of Achievement in Specific Subject Areas	365
APTITUDE TESTS	369
The Preschool Level	370
The Elementary-School Level	376
The Secondary-School Level	377
The College Level and Beyond	379
DIAGNOSTIC TESTS	381
Reading Tests	382
Math Tests	383
PSYCHOEDUCATIONAL TEST BATTERIES	383
The Kaufman Assessment Battery for Children (K-ABC) and the Kaufman Assessment Battery for Children, Second Edition (KABC-II)	384
The Woodcock-Johnson III (WJ III)	386
OTHER TOOLS OF ASSESSMENT IN EDUCATIONAL SETTINGS	388
Performance, Portfolio, and Authentic Assessment	388
Peer Appraisal Techniques	390
Measuring Study Habits, Interests, and Attitudes	391
MEET AN ASSESSMENT PROFESSIONAL Meet Eliane Hack, M.A.	360
CLOSE-UP Tests of Minimum Competency	366
EVERYDAY PSYCHOMETRICS First Impressions	372
SELF-ASSESSMENT	392

PART IV *The Assessment of Personality*

12 Personality Assessment: An Overview 393

PERSONALITY AND PERSONALITY ASSESSMENT 393
 Personality 393
 Personality Assessment 394
 Traits, Types, and States 394
PERSONALITY ASSESSMENT: SOME BASIC QUESTIONS 399
 Who? 399
 What? 405
 Where? 406
 How? 407
DEVELOPING INSTRUMENTS TO ASSESS PERSONALITY 417
 Logic and Reason 417
 Theory 418
 Data Reduction Methods 418
 Criterion Groups 421
PERSONALITY ASSESSMENT AND CULTURE 435
 Acculturation and Related Considerations 435
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Eric A. Zillmer 407
EVERYDAY PSYCHOMETRICS Some Common Item Formats 412
CLOSE-UP Assessing Acculturation and Related Variables 437
SELF-ASSESSMENT 439

13 Personality Assessment Methods 440

OBJECTIVE METHODS 440
 How Objective Are Objective Methods of Personality Assessment? 441
PROJECTIVE METHODS 442
 Inkblots as Projective Stimuli 443
 Pictures as Projective Stimuli 449
 Words as Projective Stimuli 457
 Sounds as Projective Stimuli 460
 The Production of Figure Drawings 461
 Projective Methods in Perspective 464
BEHAVIORAL ASSESSMENT METHODS 470
 The Who, What, When, Where, Why, and How of It 472
 Varieties of Behavioral Assessment 476
 Issues in Behavioral Assessment 484
A PERSPECTIVE 487
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Anthony Bram 469
EVERYDAY PSYCHOMETRICS Confessions of a Behavior Rater 477
CLOSE-UP Personality, Life Outcomes, and College Yearbook Photos 485
SELF-ASSESSMENT 488

PART V *Testing and Assessment in Action*

14 Clinical and Counseling Assessment 489

AN OVERVIEW	489
The Diagnosis of Mental Disorders	491
The Interview in Clinical Assessment	495
Case History Data	501
Psychological Tests	501
CULTURALLY INFORMED PSYCHOLOGICAL ASSESSMENT	504
Cultural Aspects of the Interview	506
Cultural Considerations and Managed Care	507
SPECIAL APPLICATIONS OF CLINICAL MEASURES	508
The Assessment of Addiction and Substance Abuse	509
Forensic Psychological Assessment	510
Diagnosis and evaluation of emotional injury	517
Profiling	517
Custody Evaluations	518
CHILD ABUSE AND NEGLECT	521
THE PSYCHOLOGICAL REPORT	524
The Barnum Effect	525
Clinical versus Mechanical Prediction	527
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Joel Goldberg	503
CLOSE-UP Assessment of Dangerousness and the Secret Service	513
EVERYDAY PSYCHOMETRICS Elements of a Typical Report of Psychological Assessment	526
SELF-ASSESSMENT	529

15 Neuropsychological Assessment 530

THE NERVOUS SYSTEM AND BEHAVIOR	530
Neurological Damage and the Concept of Organicity	531
THE NEUROPSYCHOLOGICAL EVALUATION	535
When a Neuropsychological Evaluation Is Indicated	535
General Elements of a Neuropsychological Evaluation	537
The Physical Examination	540
NEUROPSYCHOLOGICAL TESTS	541
Tests of General Intellectual Ability	543
Tests to Measure the Ability to Abstract	544
Tests of Executive Function	545
Tests of Perceptual, Motor, and Perceptual-Motor Function	549
Tests of Verbal Functioning	553
Tests of Memory	553
Neuropsychological Test Batteries	556
OTHER TOOLS OF NEUROPSYCHOLOGICAL ASSESSMENT	561
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Jeanne P. Ryan	543

CLOSE-UP Fixed versus Flexible Neuropsychological Test Batteries and the Law	558
EVERYDAY PSYCHOMETRICS Medical Diagnostic Aids and Neuropsychological Assessment	562
SELF-ASSESSMENT	564

16 Assessment, Careers, and Business 565

CAREER CHOICE AND CAREER TRANSITION	565
Measures of Interest	566
Measures of Ability and Aptitude	569
Measures of Personality	575
Other Measures	579
SCREENING, SELECTION, CLASSIFICATION, AND PLACEMENT	581
The Résumé and the Letter of Application	582
The Application Form	582
Letters of Recommendation	583
Interviews	584
Portfolio Assessment	584
Performance Tests	584
Physical Tests	588
COGNITIVE ABILITY, PRODUCTIVITY, AND MOTIVATION MEASURES	591
Measures of Cognitive Ability	591
Productivity	592
Motivation	594
JOB SATISFACTION, ORGANIZATIONAL COMMITMENT, AND ORGANIZATIONAL CULTURE	598
Job Satisfaction	598
Organizational Commitment	599
Organizational Culture	599
OTHER TOOLS OF ASSESSMENT FOR BUSINESS APPLICATIONS	602
Consumer Psychology	602
The Measurement of Attitudes	602
Surveys	604
Motivation Research Methods	607
CLOSE-UP Validity Generalization and the GATB	572
MEET AN ASSESSMENT PROFESSIONAL Meet Dr. Chris Gee	599
EVERYDAY PSYCHOMETRICS Assessment of Corporate and Organizational Culture	601
SELF-ASSESSMENT	612

References R-1

Credits C-1

Name Index N-1

Glossary/Index I-1

Preface

In the late 1970s, when work first began on our introductory measurement text, there were only a few textbooks available on the subject of measurement in psychology. All of these books provided students with a basic grounding in psychometrics, but in our opinion none of them did a very satisfactory job of it. More specifically, these books all had a variety of shortcomings that really needed to be addressed.

Problems with the Available Measurement Textbooks

From our perspective, the available measurement textbooks had many problems:

- *Reading those books was a challenge.* The books seemed to us to be written for instructors to teach from. By contrast, we believed that such books should be written for students to learn from.
- *The authors of competing books had little or no actual experience in test administration and test interpretation.* The writing in the existing books was academic enough. However, the writing betrayed a total lack of any “hands-on,” working knowledge of the subject matter. One might read the entire text, cover to cover, and never find a shred of evidence that the writer had ever administered a psychological test, personally interpreted findings from a test, or dealt face-to-face with an assessee (or any other interested party).
- *Coverage of certain key subject areas was nonexistent.* Today it’s commonplace to cover topics such as legal/ethical issues in assessment, forensic assessment, neuropsychological assessment, and psychological assessment in business. But back in those days, any coverage of these topics in the existing measurement books was the exception rather than the rule. Cohen et al. pioneered such coverage, and Cohen et al. continued to lead the way in standard-setting coverage in other areas such as culture-related issues in assessment. As late as the previous (seventh) edition of our text, we introduced a chapter on the subject of test utility—this at a time when most other competing textbooks did not even list the terms *utility* or *test utility* in their subject index.
- *The other books contained way too much description of tests.* Granted, a book on the subject of testing must contain a description of some tests. However, back then, portions of existing texts were as replete with descriptions of tests as *Tests in Print*.
- *The art program in those books did not adequately support what was written.* What passed for an art program in the books that existed at the time was some number-intensive graphs and tables, as well as some photos of test materials. Photos seemed to be inserted more to break up text than to complement it. By contrast, we believed that supplemental art could be effectively used to reinforce learning. More specifically, it could be used to stimulate the reader’s imagination and help solidify meaningful visual associations to the text.
- *Coverage of the heritage and history of the assessment enterprise was scant.* In the books that existed prior to the publication of our own, little or no effort was made

to convey a sense of how all of the facts being presented fit within the grand scheme or context of the subject matter. Tests of intelligence were presented with little or no discussion of what was meant by *intelligence*. Tests of personality were presented with little or no discussion of what was meant by *personality*. By contrast, we would make an effort to place such material not only in a historical context, but in a logical context. Our appreciation for the importance of history and context is emphasized by the fact that the first feature readers are greeted by as they open our book—as well as the last feature they may see before closing it—is a listing of noteworthy historical landmarks set within the front and back covers of our book.

- *Books existing at the time made implicit assumptions—not true in all cases—about the level of preparation students had coming in to a measurement course.* All of the measurement textbooks that came before Cohen et al. were written based on the assumption that every student taking the course was up to speed on all of the statistical concepts that would be necessary to build on learning about psychometrics. In theory, at least, there was no reason not to assume this; statistics was a prerequisite to taking testing. In practice, a different picture emerged. It was simply not the case that all students were adequately and equally prepared to begin learning statistics-based measurement concepts. Our remedy for this problem was to include a “Statistics Refresher” chapter early on, just prior to building on students’ statistics-based knowledge.

Our Vision for a New Generation of Students

We envisioned something better for students who would be using our book, and for the succeeding generations of students. At a minimum, our book would impart a sound grounding in basic psychometrics, as well as all that was minimally necessary for students to achieve a clear conceptual understanding of the assessment enterprise. We would strive to present material in meaningful—even familiar, when possible—contexts. This meant that an unprecedented effort would be made to “breathe life” into all of the numbers, equations, models, and other statistics-related material. Features like *Everyday Psychometrics* and *Close-Ups* helped in this objective, as did subsequent pedagogical aids such as *Meet an Assessment Professional*. Further, we envisioned:

- *a book that students could read easily and learn from.* Not only would our book be one that was easy for instructors to teach from, it would be one that students could read (dare we say even *enjoy* reading?). We attempted to accomplish this most challenging goal by working hard to make the content and the writing as engaging as possible. We introduced concepts and terminology in an orderly way, building on subject matter already presented. Throughout, we tried to maintain a level of writing that was scholarly, engaging, and easily managed by most students.
- *a book that would convey to students the “feel” of actually working with tools of assessment.* It was important to us that this textbook convey something of the “hands-on feel” of actually using psychological tests. In contrast to many of the people writing textbooks about testing and assessment when our book was first published—and still true today—our authorship team has a great deal of experience in administering, scoring, and interpreting tests in clinical, counseling,

school, research, and business-related contexts. We felt that students could profit from our sharing of this experience. In recent editions, this vision has also been supported by introducing students to other “real life” test users in each chapter’s *Meet an Assessment Professional* feature.

- ***a book that would provide students with a working overview of relevant legal and ethical issues in psychological testing and assessment.*** Students taking an overview course in measurement should have a solid grounding in legal/ethical issues pertinent to the assessment enterprise. Accordingly, discussion of legal/ethical issues, which sets a context for all that follows, was placed early on in the book (Chapter 2), rather than at the end of the book (where it was, and still is, in the organization of some competing textbooks).
- ***a book that would provide students with an overview of areas where tools of assessment are employed.*** A clear need existed for coverage of various applied areas of test use (such as neuropsychological and forensic applications), and this type of material broke new ground in our first edition, and has continued to evolve with changes in applications of assessment ever since.
- ***a book that contained meaningful and useful illustrations.*** Our view is and has been that illustrations need to supplement learning, strengthen associations, entertain, and help “humanize” the material being presented.

More About This Book’s Heritage

Our book was originally published by a small, independent publisher. To give the reader an idea of how small that publisher was, the company had a sales force of about five people (which included the president of the company and the acquisitions editor who acquired the book). By comparison, the existing books were published by publishers with sales forces of over 100 dedicated sales people. The “marketing” of the first edition of our book consisted of the publisher buying a list of instructors who taught a measurement course and then sending out a sample copy of the book to everyone on that list.

It did not take a huge marketing campaign on the part of this small publisher to make this book a hit among instructors and their students. One after another, in one unsolicited letter after another, instructors voiced appreciation for our perspective on the discipline, our organization of topics, our selection of subject matter to be covered, and our lucid presentation of the material. By the time we began work on the second edition, our textbook was the one that was being emulated by all of the others. It still is.

Today, coverage of many of the topics we first deemed to be essential in a measurement text is now “standard” in measurement textbooks. We assure you that such material—a statistics refresher, coverage of behavioral assessment, coverage of legal and ethical issues, and so on—were by no means standard when the first edition of our book was published.

The Organization of This Book and Related Variables

Textbook authors—past as well as present—are confronted with many choices, especially regarding variables related to presentation of the material, such as *organization* of topic areas, *content* selected, *art* used to supplement the text, *pedagogical tools* to reinforce

learning, and the *writing style* or *voice* used to “speak to” readers. We believe these variables are all critically important vis-à-vis how much students ultimately take away from the textbook they are assigned. Let’s briefly consider each of these areas and the choices that textbook authors have to make with respect to them.

Organization

From the first edition of our book forward, we have organized the information to be presented into five major sections. We have no illusions that this organization will reach the iconic status of another “big five,” but this organization has been proven to work well for both students and instructors alike. Part I, *An Overview*, contains two chapters that do just that. Chapter 1 provides a comprehensive overview of the field, including some important definitional issues, a general description of tools of assessment, and related important information couched as answers to questions regarding the *who*, *what*, *why*, *how*, and *where* of the enterprise.

The foundation for the material to come continues to be laid in the second chapter of the overview, which deals with historical, cultural, and legal/ethical issues. The material presented in Chapter 2 clearly sets a context for everything that will follow. To relegate such material to the back of the book (as a kind of elective topic, much like the way that legal/ethical issues are treated in some books), or to ignore presentation of such material altogether (as most other books have done with regard to cultural issues in assessment), is, in our estimation, a grave error. “Back page infrequency” (to borrow an MMPI-2 term) is too often the norm, and relegation of this critically important information to the back pages of a textbook too often translates to a potential shortchanging of students with regard to key cultural, historical, and legal/ethical information.

Part II, *The Science of Psychological Measurement*, contains Chapters 3 through 8. These six chapters were designed to build—logically and sequentially—on the student’s knowledge of psychometric principles. Part II begins with a chapter reviewing basic statistical principles and ends with a chapter on test construction. In between, there is extensive discussion of assumptions inherent in the enterprise, the elements of good test construction, as well as the concepts of norms, correlation, inference, reliability, and validity. In Chapter 7, entitled, “Utility,” readers will find information they need to know about this important concept, including the many factors that can affect a test’s utility. The Chapter 7 *Close-Up* provides a step-by-step, informative illustration of a hypothetical utility analysis. Students will come away from this chapter with a working knowledge of not only what utility is, but how an index of utility is derived.

Let’s note here that topics such as utility and utility analysis can get extremely complicated. However, we have never shied away from the presentation of complicated subject matter. For example, we were the first introductory textbook to present detailed information related to factor analysis. As more commercial publishers and other test users have adopted the use of item response theory (IRT) in test construction, our coverage of IRT has kept pace. As more test reviews have begun to evaluate tests not only in terms of variables such as reliability and validity but in terms of *utility*, we saw a need for the inclusion of a chapter on that topic. By the way, we could find no comparable coverage—nor any reference at all, for that matter, to *utility* or *test utility*—in any competing textbook at the time we decided to devote a chapter to that subject.

Of course, no matter how “difficult” the concepts we present are, we never for a moment lose sight of the appropriate level of presentation, or who the students are who have been assigned our text. This book is designed for students taking a first course in psychological testing and assessment. Our objective in presenting material on methods such as IRT and utility analysis is simply to acquaint the introductory student with

these techniques. The depth of the presentation in these and other areas has always been guided and informed by extensive reviews from a geographically diverse sampling of instructors who teach measurement courses. For users of this textbook, what currently tends to be required is a conceptual understanding of commonly used IRT methods. We believe our presentation of this material effectively conveys such an understanding. Moreover, it does so without unnecessarily burdening students with level-inappropriate formulas and calculations.

Part III of this book, *The Assessment of Intelligence*, contains three chapters, two that deal with the concept of intelligence and tests of intelligence, and a third chapter that deals more generally with school-related assessments. Readers will find that Chapter 11 presents a most informative, updated discussion of the response to intervention (RtI) approach to diagnosing learning abilities.

Part IV, *The Assessment of Personality*, contains two chapters, which respectively overview how personality assessments are conducted, and the various methods used.

Part V, *Testing and Assessment in Action*, is designed to convey to students a sense of how tests and other tools of assessment are actually used in clinical, counseling, business, and other settings. Instructors who teach from Chapter 15 (Neuropsychological Assessment) will be pleased to find updated material on fMRI as a tool of assessment.

On occasion we have recognized the need to reorganize material. If such reorganization better serves the interest of a logical presentation of material, we will implement it without hesitation. In this edition, for example, instructors who have used previous editions will notice that the coverage of correlation, formerly in Chapter 4, has been moved to Chapter 3. This change in the organization of material was prompted by a persuasive suggestion made by an instructor who teaches an undergraduate measurement course.

Content

In addition to a logical organization that sequentially builds on student learning, we view *content* selection as another key element of our appeal. The multifaceted nature and complexity of the discipline affords textbook authors wide latitude in terms of what material to elaborate on, what material to ignore, and what material to highlight, exemplify, or illustrate. We take full advantage of the wide range of choices available and include not only what students *must* know, but what students (and instructors) might *like* to know as well. As such, we pepper the text with sometimes unexpected, intriguing facts and perspectives. Our objective here has always been to enhance the memorability of the material, while enriching students' appreciation for it.

So, for example, in the context of discussing projective techniques, students are introduced to one (surprising) pioneer in projective testing, B. F. Skinner (yes, *that* B. F. Skinner!). In Chapter 12, students are privy to an “inside look” at the detention center at Guantanamo Bay, Cuba, through the eyes of a consulting psychologist. In the *Close-Up* in Chapter 2, we introduce students to the fascinating but controversial life and times of Henry Herbert Goddard. How many of us knew (or could ever imagine) that the first coach of the University of Southern California (USC) football team was none other than H. H. Goddard?

And speaking of *Close-Ups*—the pedagogical tool employed in each chapter since the first edition to focus-in and supplement knowledge on a particular assessment-related topic—we believe that students will find a wealth of useful information in the wide array of topics covered in our eighth-edition *Close-Ups*. For example, the *Close-Up* in Chapter 1 tackles the growing controversy regarding the issue of third-party presence during test administration. The Chapter 5 *Close-Up* introduces

students to item response theory (IRT). In Chapter 12, the *Close-Up* presents timely material on measures of acculturation.

The Art Program

Complementing a judicious selection of manuscript content is an equally thoughtful (and thought-provoking) selection of illustrations. See, for example, the series of photos used to illustrate a computer-assisted method of quantifying back stress (Chapter 1), the turn-of-the-century photo of the immigrants being tested at Ellis Island (Chapter 2), and the dramatic photo capturing hockey violence in the context of discussion of the trait of aggression and a questionnaire designed to measure that trait (Chapter 12). In the world of textbooks, such photos may not seem very revolutionary. And maybe they are not. However, in the world of *measurement* textbooks, our innovative art program was indeed revolutionary (and by all accounts, still is). In some instances, figures have been used to pay homage to some very well-known (as well as some not-so-well-known) contributors to the field (such as L. L. Thurstone and Lev Vygotsky), as well as more contemporary psychologists (such as John L. Holland and John E. Exner) who have recently left us.

Pedagogical Tools

The objective of incorporating timely, relevant, and intriguing illustrations of assessment-related material is furthered by several *pedagogical tools* built into the text. We have already made reference to our strategic use of *Close-Ups*. Another pedagogical tool we innovated seven editions ago is *Everyday Psychometrics*. In each chapter of the book, relevant, practical, and “everyday” examples of the material being discussed are highlighted in an *Everyday Psychometrics* box. For example, in the *Everyday Psychometrics* presented in Chapter 1 (“Everyday Accommodations”), students will be introduced to accommodations made in the testing of persons with handicapping conditions. In Chapter 4, the *Everyday Psychometrics* feature (“Putting Tests to the Test”) equips students with a working overview of the variables they need to be thinking about when reading about a test and evaluating how satisfactory the test really is for a particular purpose. In Chapter 5, the subject of the *Everyday Psychometrics* is the reliability of the instrumentation used by law enforcement authorities to measure alcoholic intoxication.

A pedagogical tool called *Meet an Assessment Professional* was introduced in the previous (seventh) edition. This feature provides a forum through which everyday users of psychological tests from various fields can share insights, experiences, and advice with students. The result is that in each chapter of our book, students are introduced to a different test user and provided with an intriguing glimpse of their professional life—this in the form of an excerpt from their *Meet and Assessment Professional (MAP)* essay. *MAP* essays in their entirety are presented on our companion instructional website: www.mcgrawhill.com/cohentesting8, which, by the way, also contains a wealth of other course-enhancing, assessment-related information for students.

Beyond introducing students to accomplished test users who are earning a living doing assessment-related work, each *MAP* essay serves to underscore the practical value of learning about psychological tests. For example, in Chapter 4, students will meet a team of test users, Drs. Steve Julius and Howard Atlas, who have pressed psychometric knowledge into the service of professional sports. They provide a unique and fascinating account of how application of their knowledge of was used to improve the on-court of achievement of the Chicago Bulls. New to this edition are *MAP* essays from Stephen Finn, the well-known proponent of therapeutic assessment (Chapter 1);

Nathaniel Mohatt, a valued contributor in the area of cultural issues in assessment (Chapter 2); Benoît Verdon, a French clinical psychologist (Chapter 3); Eliane Hack, a working school psychologist (Chapter 11); Anthony Bram, an independent practitioner who is also a clinical instructor at Harvard Medical School (Chapter 13); Joel Goldberg, an independent practitioner who is Director of Clinical Training at York University's Department of Psychology (Chapter 14); Jeanne Ryan, a practicing neuropsychologist who is Clinical Director at the Neuropsychology Clinic at State University of New York at Plattsburgh (Chapter 15); and Chris Gee, a corporate psychologist who is Director of Research Services at the Self Management Group (Chapter 16).

There are other pedagogical tools that readers (as well as other textbook authors) may take for granted—but we do not. Consider, in this context, the various tables and figures found in every chapter. In addition to their more traditional use, we view tables as space-saving devices in which a lot of information may be presented. For example, in the first chapter alone, tables are used to provide succinct but meaningful comparisons between *testing* and *assessment*, the *pros* and *cons* of computer-assisted psychological assessment, and the *pros* and *cons* of using various sources of information about tests.

Critical thinking may be defined as “the active employment of judgment capabilities and evaluative skills in the thought process” (Cohen, 1994, p. 12). *Generative thinking* may be defined as “the goal-oriented intellectual production of new or creative ideas” (Cohen, 1994, p. 13). The exercise of both of these processes, we believe, helps optimize one’s chances for success in the academic world as well as in more applied pursuits. In the early editions of this textbook, questions designed to stimulate critical and generative thinking were raised “the old-fashioned way.” That is, they were right in the text, and usually part of a paragraph. Acting on the advice of reviewers, we made this special feature of our writing even more special in the sixth edition of this book; we raised these critical thinking questions in the margins with a *Just Think* heading. Perhaps with some encouragement from their instructors, motivated students will, in fact, give thoughtful consideration to these (critical and generative thought-provoking) *Just Think* questions.

In addition to critical thinking and generative thinking questions called out in the text, other pedagogical aids in this book include original cartoons created by the authors, original illustrations created by the authors (including the model of memory in Chapter 15), and original acronyms created by the authors.¹ Each chapter ends with a *Self-Assessment* feature that students may use to test themselves with respect to key terms and concepts presented in the text. By the way, many of the same terms listed in the *Self-Assessment* exercise are used as the response keyed correct in the corresponding crossword puzzles presented for student distribution in this text’s companion website.

Writing Style

What type of *writing style* or author *voice* works best with students being introduced to the field of psychological testing and assessment? Instructors familiar with the many measurement books that have come (and gone) may agree with us that the “voice” of too many authors in this area might best be characterized as humorless and academic to

1. By the way, our use of the French word for black (*noir*) as an acronym for levels of measurement (nominal, ordinal, interval, and ratio) now appears in other textbooks. So if, as they say, “imitation is the sincerest form of flattery,” we’ll use this occasion to express our gratitude to fellow textbook authors for paying us their highest compliments.

the point of arrogance or pomposity. Students do not tend to respond well to textbooks written in such styles, and their eagerness and willingness to spend study time with these authors (and even their satisfaction with the course as a whole) may easily suffer as a consequence.

In a writing style that could be characterized as somewhat informal and—to the extent possible, given the medium and particular subject being covered—“conversational,” we have made every effort to convey the material to be presented as clearly as humanly possible. In practice, this means:

- keeping the vocabulary of the presentation appropriate (without ever “dumbing-down” or trivializing the material);
- presenting so-called difficult material in step-by-step fashion where appropriate, and always preparing students for its presentation by placing it in an understandable context;
- italicizing the first use of a key word or phrase and then bolding it when a formal definition is given;
- providing a relatively large glossary of terms to which students can refer;
- supplementing material where appropriate with visual aids, tables, or other illustrations.
- incorporating timely, relevant, and intriguing illustrations of assessment-related material in the text as well as in the online materials.

In addition, we have interspersed some elements of humor in various forms (original cartoons, illustrations, and vignettes) throughout the text. The judicious use of humor to engage and maintain student interest is something of a novelty among measurement textbooks. Where else would one turn for pedagogy that employs an example involving a bimodal distribution of test scores from a new trade school called *The Home Study School of Elvis Presley Impersonators*? As readers learn about face validity, they discover why it “gets no respect” and how it has been characterized as “the Rodney Dangerfield of psychometric variables.” Numerous other illustrations could be cited here. But let’s reserve those smiles as a pleasant surprise when readers happen to come upon them.

Also in the interest of engaging and maintaining student interest, we do not hesitate to draw on popular culture for examples. *The X-Factor*, *Iron Chef*, *The Apprentice*, *South Park*, and *Survivor* are television shows that many students watch, and a surprise reference to one of them to illustrate an assessment-related point can pair pleasant feelings of recognition with learning—perhaps more solidly involving students in the material. In the course of learning how to write a good matching-type item, for example, students are challenged to identify what actors Pierce Brosnan, Sean Connery, Daniel Craig, Timothy Dalton, George Lazenby, David Niven, and Roger Moore all have in common.

“Humanization” of Material

Perhaps our treatment of the discipline—and how radically different that treatment is from other textbooks on the subject—is best characterized by our dedicated and persistent efforts to *humanize* the presentation. While other authors in this discipline impress us as blindly intent on viewing the field as Greek letters to be understood and formulas to be memorized, we view an introduction to the field to be about *people* as much as anything else. Students are more motivated to learn this material when they can place it in a human context. Many psychology students simply do not respond well to endless presentations of psychometric concepts and formulas. In our opinion, to *not* bring a human face to

the field of psychological testing and assessment, is to risk perpetuating all of those unpleasant (and now unfair) rumors about the course that first began circulating long before the time that the senior author himself was an undergraduate.

Our effort to humanize the material is evident in the various ways we have tried to bring a face (if not a helping voice) to the material. The inclusion of *Meet an Assessment Professional* (much like the *Test Developer Profiles* in past editions) is a means toward that end, as it quite literally “brings a face” to the enterprise. Our inclusion of interesting biographical facts on historical figures in assessment is also representative of efforts to humanize the material. Consider in this context the photo and brief biographical statement of MMPI-2 senior author James Butcher in Chapter 12 (p. 428). Whether through such images of historical personages or by other means, our objective has been made to truly involve students via intriguing, real-life illustrations of the material being discussed. See, for example:

- the discussion of life-or-death psychological assessment and the ethical issues involved (pp. 69–71)
- the intriguing hypotheses that have been advanced regarding the relationship between categorical cutoffs and human emotion (p. 8)
- the candid “confessions” of a behavior rater in the *Everyday Psychometrics* feature in Chapter 12 (pp. 477–478)
- the research that has been conducted linking life outcomes and personality to evaluations of college yearbook photos (p. 485)
- the discussion of the utility of tests to measure aggressiveness (p. 395) and dangerousness (pp. 513–514)
- the material on the use of tests by the military to select pilots and NASA to select astronauts (pp. 585–587)

We believe that assessment is, after all, a uniquely human, problem-solving enterprise in which data from a variety of tools (tests among them) is gathered, skillfully assimilated, and professionally interpreted. The process of assessment may be distinguished from, and contrasted with, the administration of tests. The latter process, otherwise known as *testing*, is one that may result simply in a test score; it can be, and often is, relatively mechanistic and devoid of any problem-solving efforts. In a bygone era, no distinction was made between the terms *testing* and *assessment*. Consequently, textbooks might be titled *Psychological Testing* even if, in substance, these books were actually much broader in scope (dealing with the use of various tools of assessment and the application of measurement principles). Today, to equate *testing* with *assessment* seems to us to be anachronistic. Moreover, such an equation confuses a problem-solving process that requires professional skill, knowledge of measurement theory, and knowledge of applicable legal, ethical, and cultural issues, with a process more akin to summing the number of responses keyed correct on an answer sheet.

So how has our “humanization” of the material in this discipline been received by some of its more “hard core” and “old school” practitioners? Very well, thank you—at least from all that we have heard, and the dozens of reviews that we have read over the years. What stands out prominently in my own mind was the reaction of one particular psychometrician whom I happened to meet at an APA convention not long after the first edition of this text was published. Lee J. Cronbach was quite animated as he shared with me his delight with the book, and how refreshingly different he thought that it was from anything comparable that had been published. I was so grateful to Lee for his encouragement, and felt so uplifted by that meeting, that I subsequently requested a photo from Lee for use in the second edition. The photo he sent was indeed published

in the second edition of this book—this despite the fact that at that time, Lee had a measurement book that could be viewed as a direct competitor to ours. Regardless, I felt it was important not only to acknowledge Lee’s esteemed place in measurement history, but to express my sincere gratitude in this way for his kind words and for what I perceived as his most valued “seal of approval.”

What Has Changed: New to This Edition

Of course, this eighth edition of *Psychological Testing and Assessment* has been thoroughly updated. It contains updated discussions of any instruments highlighted, as well as updated information about new test-related legislation, judicial decisions, and administrative regulations. Additionally, new and updated coverage is also presented on a wide variety of assessment-related topics, a small sampling of which would include the presentation of new material on: the new, fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-V)*; new neuropsychological tools such as the fMRI; response to intervention (RtI) as a measure for learning disabilities; Robert Sternberg’s concept of *successful intelligence*; Maria Kozhevnikov’s concept of *visual-object intelligence*; a new measure for evaluating the internal consistency of a test called the *average proportional distance* (developed by our own Eddy Sturman); historical information on assessment in ancient Egypt; conditions that might prompt a neuropsychological evaluation; and the utility of the nominal distinction between *objective* versus *projective* assessment.

We have expanded our discussions of the Flynn effect, the MMPI-2, the timeline followback (TLFB) method in behavioral assessment, competency to stand trial, systematic versus random error in measurement, test validation strategies, classical test theory as contrasted to item response theory, nature versus nurture as factors in measured intelligence, and the purpose of educational assessment within the context of current legislation. In addition, instructors who have used previous editions will find many more “Just Thinks” to engage students, as well as more glossary terms (and corresponding “Self-Assessment” terms) added as pedagogical tools.

What Has Not Changed

We have written at length about how this book was conceived, and how it has changed over time. What has *not* changed and what will not change is our dedicated resolve to provide a leading-edge, much-emulated-but-never-duplicated, measurement textbook that

- “speaks to” students with an author voice that humanizes the discipline, making it all the more understandable;
- introduces students to the assessment enterprise and overviews the wide range of instruments and procedures they may encounter;
- familiarizes students with the reasoning behind the construction of tests and the rationale of various approaches to assessment;
- leaves students with a sense of what constitutes the appropriate uses of tests and the various legal and ethical issues involved;
- leaves students with a sense of what constitutes the appropriate and inappropriate uses of tests;

- compels students to think critically, generatively, and actively about issues related to testing and assessment; and,
- provides instructors with a text that has timely, new elements in each edition, and a classroom-proven package of ancillaries (including a new eighth-edition *Instructor's Manual*, complete with a new and revised *Test Item Bank*), as well as intriguing supplementary material posted for students at our online learning center, www.mhhe.com/cohentesting8.

It took about 12 years from initial conception of this book to the publication of our first edition. History (in the form of user reviews) records that, for the most part, we “got it right.” Consequently, we’re not about to stop “getting it right” now.

Meet the Authors

New to this edition is Edward Sturman, Ph.D., an Associate Professor of Psychology at the State University of New York, Plattsburgh. Dr. Sturman is the co-coordinator of the Psychology program at the Queensbury branch campus, where he has taught many courses, including a seminar in Psychological Assessment. Dr. Sturman has developed several psychological tests, including the Mood Disorders Insight Scale (MDIS) and the Involuntary Subordination Questionnaire (ISQ), which have been linked to the course and outcome of mood disorders. He has also conducted research into the assessment of competency and developed a new method to evaluate the reliability of tests. His research findings have been published in well-regarded psychological journals and presented at major psychological conferences. Prior to his current teaching position, Dr. Sturman worked at the Self-Management Group as a consultant investigating the link between personality and performance in competitive environments, including sales and management positions at large corporations. His current research is primarily focused on the vulnerability of various personality styles to mental disorder as well as the evolutionary underpinnings of mental disorder. Dr. Sturman thanks his students, and in particular, Michelle Mann-Saumier, Kylie McKeighan, Joyalina David, Jeff Merrigan, and Jennifer Burch Dean for their work on his contribution to this book.

Mark E. Swerdlik, Ph.D., ABPP, is Professor of Psychology at Illinois State University, where he has taught the undergraduate psychological measurement course, conducted professional seminars addressing legal/ethical issues in assessment, and supervised practicum students in assessment. He has served as an editorial board member of several journals, written test reviews for several journals, reviewed test-scoring software for a major test publisher, and served as a reviewer for the *Mental Measurements Yearbook*. In various professional capacities, he has participated in the standardization of many psychological tests, including, for example, the WISC-R, the WISC-III, the Kaufman Assessment Battery for Children (K-ABC), the Stanford-Binet IV, the Peabody Picture Vocabulary Test (PPVT), the Kaufman Test of Educational Achievement, the Vineland Adaptive Behavior Scale, the Psychological Processing Checklist (PPC), and the Psychological Processing Checklist-Revised (PPC-R). As a licensed clinical psychologist, a nationally certified school psychologist, independent practitioner, and consultant, Dr. Swerdlik administers and interprets psychological tests, and conducts seminars to train fellow professionals in proper test administration, scoring, and interpretation procedures. He has also served as a program evaluator for many programs, a partial listing of which would include the Heart of Illinois Low Incidence Association (HILA), the Autism/Pervasive Developmental Delays Training and Technical Assistance Project,

and the Illinois National Guard Statewide Reintegration Program for Combat Veterans (for veterans who served in Iraq and Afghanistan, from 2006 to the present).

Ronald Jay Cohen, Ph.D., ABPP, is a Diplomate of the American Board of Professional Psychology in Clinical Psychology, and a Diplomate of the American Board of Assessment Psychology (ABAP). He is a New York State licensed psychologist, and a “scientist-practitioner” and “scholar-professional” in the finest traditions of each of those terms. During a long and gratifying professional career in which he has published numerous journal articles and books, Dr. Cohen has had the privilege of personally working alongside some of the luminaries in the field of psychological assessment, including David Wechsler (while Cohen was a clinical psychology intern at Bellevue Psychiatric Hospital in New York City) and Doug Bray (while working as an assessor for AT&T in its Management Progress Study). After serving his clinical psychology internship at Bellevue, Dr. Cohen was appointed Senior Psychologist there, and his clinical duties entailed not only psychological assessment but the supervision and training of others in this enterprise. Subsequently, as an independent practitioner in the New York City area, Dr. Cohen taught various courses at local universities on an adjunct basis, including undergraduate and graduate courses in psychological assessment. Asked by a colleague to conduct a qualitative research study for an advertising agency, Dr. Cohen would quickly become a sought-after qualitative research consultant with a client list of major companies and organizations—among them Paramount Pictures, Columbia Pictures, NBC Television, the Campbell Soup Company, Educational Testing Service, and the College Board. Dr. Cohen’s approach to qualitative research, referred to by him as *dimensional qualitative research*, has been emulated and written about by qualitative researchers around the world. Working as a consultant to one major company that wanted to learn more about its corporate culture, Dr. Cohen developed the Discussion of Organizational Culture (a qualitative research instrument discussed in Chapter 16). It was Dr. Cohen’s work in the area of qualitative assessment that led him to found the scholarly journal *Psychology & Marketing*, which in 2012 celebrated some 30 years of consecutive publishing with Dr. Cohen as editor-in-chief.

Thanks

The authors have been focused and diligent in their efforts to bring you a leading-edge measurement textbook that involves students in the subject matter and imparts a wealth of academic and applied information essential to understanding psychological testing and assessment. Of course, there are a number of people who must be thanked for helping to make this eighth edition the exciting and fresh revision that it is. Professor Jennifer Kisamore of the University of Oklahoma wrote the early drafts of the chapter on test utility as well as early drafts of the corresponding chapter in the instructor’s manual for the previous edition of this book. Dr. Bryce Reeve of the National Institutes of Health wrote the *Meet an Assessment Professional* essay for Chapter 5. We were so intrigued by his recounting of his “real life” use of item response theory (IRT) that we asked him to expand his contribution in the form of the *Close-Up* that is also presented in that chapter. Let’s take this opportunity to thank each and every one of the 17 professionals who wrote intriguing, often inspiring, *Meet an Assessment Professional* essays, one for each chapter in this book.² Thanks are also due Susan Cohen, the senior author’s wife,

2. There are 16 chapters in the book but 17 *Meet an Assessment Professional* contributors. How did this happen? Two sports psychologists to the Chicago Bulls who collaborated on their essay for Chapter 4 “came as a team.”

who helped in the creation of the chapter-by-chapter crossword puzzles presented on this text's companion website, and Ramya Thirumavalavan for her skillful production editing. Thanks also to all of the reviewers who supplied helpful input:

Wendy Dunn, *Coe College*

Jameson Hirsch, *East Tennessee State University*

Sharon Rae Jenkins, *University of North Texas*

Jeff Maney, *Midwestern University*

Kerry Schwanz, *Coastal Carolina University*

Noam Shpancer, *Otterbein College*

Brian Carey Sims, *North Carolina A&T State University*

Randi Smith, *Metropolitan State College of Denver*

Of course, the present authorship team has sole responsibility for all material in this eighth edition, and we take complete responsibility for any possible errors that may have somehow found their way in to it.

And on a Personal Note . . .

I think back to the time when we were just wrapping up work on the sixth edition of this book. At that time, I received the unexpected and most painful news that my mother had suffered a massive and fatal stroke. It is impossible to express the sense of sadness and loss experienced by myself, my brother, and my sister, as well as the countless other people who knew this gentle, loving, and much-loved person. To this day, we continue to miss her counsel, her sense of humor, and just knowing that she's there for us. We continue to miss her genuine exhilaration, which in turn exhilarated us, and the image of her welcoming, outstretched arms whenever we came to visit. Her children were her life, and the memory of her smiling face, making each of us feel so special, survives as a private source of peace and comfort for us all. She always kept a copy of this book proudly displayed on her coffee table, and I am very sorry that a copy of more recent editions did not make it to that most special place. My dedication of this book is one small way I can meaningfully acknowledge her contribution, as well as that of my beloved, deceased father, to my personal growth. As in the sixth edition, I am using my parents' wedding photo in the dedication. They were so good together in life. And so there Mom is, reunited with Dad. Now, that is something that would make her very happy.

As the reader might imagine, given the depth and breadth of the material covered in this textbook, it requires great diligence and effort to create and periodically re-create an instructional tool such as this that is timely, informative, and readable. Thank you, again, to all of the people who have helped through the years. Of course, I could not do it myself were it not for the fact that even through eight editions, this truly Herculean undertaking remains a labor of love.

Ronald Jay Cohen, Ph.D., ABPP

Diplomate, American Board of Professional Psychology (Clinical)

Diplomate, American Board of Assessment Psychology

Psychological Testing and Assessment

All fields of human endeavor use measurement in some form, and each field has its own set of measuring tools and measuring units. For example, if you're recently engaged or thinking about becoming engaged, you may have learned about a unit of measure called the carat. If you've been shopping for a computer, you may have learned something about a unit of measurement called a byte. As a student of psychological measurement, you need a working familiarity with some of the commonly used units of measure in psychology as well as knowledge of some of the many measuring tools employed. In the pages that follow, you will gain that knowledge as well as an acquaintance with the history of measurement in psychology and an understanding of its theoretical basis.

Testing and Assessment

The roots of contemporary psychological testing and assessment can be found in early twentieth-century France. In 1905, Alfred Binet and a colleague published a test designed to help place Paris schoolchildren in appropriate classes. Binet's test would have consequences well beyond the Paris school district. Within a decade an English-language version of Binet's test was prepared for use in schools in the United States. When the United States declared war on Germany and entered World War I in 1917, the military needed a way to screen large numbers of recruits quickly for intellectual and emotional problems. Psychological testing provided this methodology. During World War II the military would depend even more on psychological tests to screen recruits for service. Following the war, more and more tests purporting to measure an ever-widening array of psychological variables were developed and used. There were tests to measure not only intelligence but also personality, brain functioning, performance at work, and many other aspects of psychological and social functioning.

Psychological Testing and Assessment Defined

The world's receptivity to Binet's test in the early twentieth century spawned not only more tests but more test developers, more test publishers, more test users, and the emergence of what, logically enough, has become known as a testing enterprise. "Testing" was the term used to refer to everything from the administration of a test (as in "Testing in progress") to the interpretation of a test score ("The testing indicated that . . ."). During World War I the term "testing" aptly described the group screening of thousands

of military recruits. We suspect that it was then that the term gained a powerful foothold in the vocabulary of professionals and laypeople. The use of “testing” to denote everything from test administration to test interpretation can be found in postwar textbooks (such as Chapman, 1921; Hull, 1922; Spearman, 1927) as well as in various test-related writings for decades thereafter. However, by World War II a semantic distinction between testing and a more inclusive term, “assessment,” began to emerge.

During World War II the U.S. Office of Strategic Services (OSS), a predecessor to today’s Central Intelligence Agency (CIA), used a variety of procedures and measurement tools—psychological tests among them—in selecting military personnel for highly specialized positions involving espionage, intelligence gathering, and the like. For example, one of the tools employed was a very uncomfortable, group-on-one interview technique to evaluate how well candidates might respond to Gestapo-like interrogation. With a light harshly pointed at their face, interviewees would have to draw on their own creativity, persuasive abilities, and other resources to satisfactorily defend and explain a given scenario to a group of increasingly hostile interviewers. Candidates might have to explain why they were in a particular building that was off-limits to them, and doing something that they were not authorized to do, such as looking at or removing classified files. Candidates were evaluated on a number of variables, such as their ability to maintain noncontradictory responses. Today, such an assessment method, or any assessment method that has potential for harming the persons being assessed, would be likely to raise serious ethical concerns.

As summarized in *Assessment of Men* (OSS Assessment Staff, 1948) and elsewhere (Murray & MacKinnon, 1946), the assessment data generated were subjected to thoughtful integration and evaluation by highly trained assessment center staff. The OSS model—using an innovative variety of evaluative tools along with data from the evaluations of highly trained assessors—would later inspire what is now referred to as the assessment center approach to personnel evaluation (Bray, 1982).

Military, clinical, educational, and business settings are but a few of the many contexts that entail behavioral observation and active integration by assessors of test scores and other data. In such situations, the term *assessment* may be preferable to *testing*. In contrast to testing, assessment acknowledges that tests are only one type of tool used by professional assessors (along with other tools, such as the interview), and that the value of a test, or of any other tool of assessment, is intimately linked to the knowledge, skill, and experience of the assessor.

The semantic distinction between psychological testing and psychological assessment is blurred in everyday conversation. Somewhat surprisingly, the distinction between the two terms still remains blurred in some published “psychological testing” textbooks. Yet the distinction is important. Society at large is best served by a clear definition of and differentiation between these two terms as well as related terms such as *psychological test user* and *psychological assessor*. Clear distinctions between such terms may also help avoid the turf wars now brewing between psychology professionals and members of other professions seeking to use various psychological tests. In many psychological evaluation contexts, conducting an assessment requires greater education, training, and skill than simply administering a test.

We define **psychological assessment** as the gathering and integration of psychology-related data for the purpose of making a psychological evaluation that is accomplished through the use of tools such as tests, interviews, case studies, behavioral observation, and specially designed apparatuses and measurement procedures. We define **psychological testing** as the process of measuring psychology-related variables by means of devices or procedures designed to obtain a sample of behavior. Some of the differences between these two processes are presented in Table 1–1.¹

JUST THINK . . .

Describe a situation in which testing is more appropriate than assessment. By contrast, describe a situation in which assessment is more appropriate than testing.

Table 1–1
Testing in Contrast to Assessment

In contrast to the process of administering, scoring, and interpreting psychological tests (psychological testing), psychological assessment may be conceived as a problem-solving process that can take many different forms. How psychological assessment proceeds depends on many factors, not the least of which is the reason for assessing. Different tools of evaluation—psychological tests among them—might be marshaled in the process of assessment, depending on the particular objectives, people, and circumstances involved as well as on other variables unique to the particular situation.

Admittedly, the line between what constitutes testing and what constitutes assessment is not always as clear as we might like it to be. However, by acknowledging that such ambiguity exists, we can work to sharpen our definition and use of these terms. It seems useful to distinguish the differences between testing and assessment in terms of the objective, process, and outcome of an evaluation and also in terms of the role and skill of the evaluator. Keep in mind that, although these are useful distinctions to consider, exceptions can always be found.

Testing	Assessment
<i>Objective</i>	
Typically, to obtain some gauge, usually numerical in nature, with regard to an ability or attribute.	Typically, to answer a referral question, solve a problem, or arrive at a decision through the use of tools of evaluation.
<i>Process</i>	
Testing may be individual or group in nature. After test administration, the tester will typically add up “the number of correct answers or the number of certain types of responses . . . with little if any regard for the how or mechanics of such content” (Maloney & Ward, 1976, p. 39).	Assessment is typically individualized. In contrast to testing, assessment more typically focuses on how an individual processes rather than simply the results of that processing.
<i>Role of Evaluator</i>	
The tester is not key to the process; practically speaking, one tester may be substituted for another tester without appreciably affecting the evaluation.	The assessor is key to the process of selecting tests and/or other tools of evaluation as well as in drawing conclusions from the entire evaluation.
<i>Skill of Evaluator</i>	
Testing typically requires technician-like skills in terms of administering and scoring a test as well as in interpreting a test result.	Assessment typically requires an educated selection of tools of evaluation, skill in evaluation, and thoughtful organization and integration of data.
<i>Outcome</i>	
Typically, testing yields a test score or series of test scores.	Typically, assessment entails a logical problem-solving approach that brings to bear many sources of data designed to shed light on a referral question.

1. Especially when discussing general principles related to the creation of measurement procedures, as well as the creation, manipulation, or interpretation of data generated from such procedures, the word *test* (as well as related terms, such as *test score*) may be used in the broadest and most generic sense; that is, “test” may be used in shorthand fashion to apply to almost any procedure that entails measurement (including, for example, situational performance measures). Accordingly, when we speak of “test development” in Chapter 8, many of the principles set forth will apply to the development of other measurements that are not, strictly speaking, “tests” (such as situational performance measures, as well as other tools of assessment). Having said that, let’s reemphasize that a real and meaningful distinction exists between the terms *psychological testing* and *psychological assessment*, and that effort should continually be made not to confuse the meaning of these two terms.

The process of assessment In general, the process of assessment begins with a referral for assessment from a source such as a teacher, school psychologist, counselor, judge, clinician, or corporate human resources specialist. Typically one or more referral questions are put to the assessor about the assessee. Some examples of referral questions are: “Can this child function in a regular classroom?,” “Is this defendant competent to stand trial?,” and “How well can this employee be expected to perform if promoted to an executive position?”

The assessor may meet with the assessee or others before the formal assessment in order to clarify aspects of the reason for referral. The assessor prepares for the assessment by selecting the tools of assessment to be used. For example, if the assessment occurs in a corporate or military setting and the referral question concerns the assessee’s leadership ability, the assessor may wish to employ a

JUST THINK . . .

What qualities makes a good leader? How might these qualities be measured?

measure (or two) of leadership. Typically the assessor’s own past experience, education, and training play a key role in the specific tests or other tools to be employed in the assessment. Sometimes an institution in which the assessment is taking place has prescribed guidelines for which instruments can and cannot be used. In most

every assessment situation, particularly situations that are relatively novel to the assessor, the tool selection process may be informed by some research in preparation for the assessment. For example, in the assessment of leadership, the tool selection procedure might be informed by publications dealing with behavioral studies of leadership (Derue et al., 2011) psychological studies of leaders (Kouzes & Posner, 2007), cultural issues in leadership (Byrne & Bradley, 2007), or whatever aspect of leadership the assessment will be focused on (e.g., Carnevale et al., 2011; Elliott, 2011; Kalshoven et al., 2011).

Subsequent to the selection of the instruments or procedures to be employed, the formal assessment will begin. After the assessment, the assessor writes a report of the findings that is designed to answer the referral question. More feedback sessions with the assessee and/or interested third parties (such as the assessee’s parents and the referring professional) may also be scheduled.

Different assessors may approach the assessment task in different ways. Some assessors approach the assessment with minimal input from assessees themselves. Other assessors view the process of assessment as more of a collaboration between the assessor and the assessee. For example, in one approach to assessment, referred to (logically enough) as **collaborative psychological assessment**, the assessor and assessee may work as “partners” from initial contact through final feedback (Finello, 2011; Fischer, 1978, 2004, 2006). One variety of collaborative assessment includes an element of therapy as part of the process. Stephen Finn and his colleagues (Finn, 2003, 2011; Finn & Martin, 1997; Finn & Tonsager, 2002) have described a collaborative approach to assessment called **therapeutic psychological assessment**. Here, therapeutic self-discovery and new understandings are encouraged throughout the assessment process. Learn more about this approach, and “say hello” to Dr. Stephen Finn himself, in this chapter’s *Meet an Assessment Professional*. This feature was designed to help readers acquire a firsthand acquaintance with “real life” test use. In each chapter of this book, you will be introduced to a different test user, and a unique perspective on the assessment enterprise. All assessment professionals featured were asked to write an essay sharing their own thoughts and experiences. A brief excerpt of each essay is presented in each chapter, and the complete version of each essay is available online at www.mhhe.com/cohentesting8.

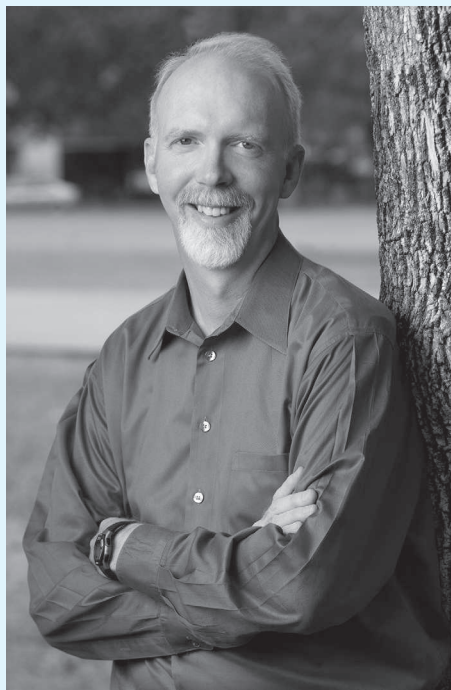
Another approach to assessment that seems to have picked up momentum in recent years, most notably in educational settings, is referred to as *dynamic assessment*

Meet Dr. Stephen Finn

In Therapeutic Assessment, we use a variety of psychological instruments including tests of cognitive functioning (e.g., the WAIS-IV), self-report tests of personality and symptomatology (e.g., the MMPI-2-RF), and performance-based personality tests (e.g., the Rorschach). We select the tests we use based on our initial session with a client. In that meeting, we help clients formulate personalized “assessment questions” they wish to have answered, such as “Why do I have such a difficult time making eye contact?” or “Why have I never been able to have an intimate relationship?” We then select tests that will help address the clients’ questions as well as those questions given to us by any referring professionals. For example, with the questions just mentioned, we might propose that a client take the MMPI-2-RF and the Rorschach, because our experience is that the combination of a self-report and performance-based personality test is helpful in helping us understand these types of issues. In our initial session, we would also collect comprehensive background information about the concerns reflected in the client’s questions. For example, we would ask about when it is most difficult or easiest for the client to make eye contact, when this problem began, and what the client already has tried to address this problem. We would also ask about previous attempts to have intimate relationships.

We believe that at their best, psychological tests serve as “empathy magnifiers”—helping us to get “in our clients’ shoes” and understand puzzles, quandaries, or stuck points in their lives that they have not been able to address in other ways. We administer tests in a standardized fashion early in our assessments, and find that the information they provide yields very useful hypotheses about why clients have the problems they do. Often, through our tests, we are able to help people understand puzzling, even self-destructive or off-putting behaviors that other mental health professionals have not been able to understand or ameliorate. And we consciously use tests to identify people’s strengths as well as their struggles.

We involve clients as collaborators and “co-experimenters” during our testing sessions. For example, with the client mentioned above, we might discuss actual MMPI-2 items suggesting



Stephen Finn, Ph.D., Founder, Center for Therapeutic Assessment, Austin, Texas.

that the client felt worthless and ashamed. Or we might ask the client to think with us about the following Rorschach responses: “A bat that is flying with terribly damaged wings—I don’t know how it’s continuing to fly” and “A mangy dog—the kind no one would ever take home from the animal shelter.” We might even ask the client to experiment in session with making more eye contact with the assessor, and to pay attention to what he or she feels. All of these interactions might lead to discussions about the client’s feeling inadequate and ashamed, how such feelings came to be, and how this is all related to the client’s assessment questions about making eye contact and having intimate relationships.

At the end of the assessment, we would show and talk to the client about the actual test scores, and we would discuss “next steps” the client could take to address the problems in living that were the focus of the assessment. Often a therapeutic assessment is a good entry into further psychological treatment . . .

Read more of what Dr. Finn had to say—his complete essay—at www.mhhe.com/cohentesting8.

(Poehner & van Compernelle, 2011). The term *dynamic* may suggest that a psychodynamic or psychoanalytic approach to assessment is being applied. However, that is not the case. As used in the present context, *dynamic* is used to describe the interactive, changing, or varying nature of the assessment. In general, **dynamic assessment** refers to an interactive approach to psychological assessment that usually follows a model of (1) evaluation, (2) intervention of some sort, and (3) evaluation. Dynamic assessment is most typically employed in educational settings, although it may be employed in correctional, corporate, neuropsychological, clinical, and most any other setting as well.

Intervention between evaluations, sometimes even between individual questions posed or tasks given, might take many different forms, depending upon the purpose of the dynamic assessment (Haywood & Lidz, 2007). For example, an assessor may intervene in the course of an evaluation of an assessee's abilities with increasingly more explicit feedback or hints. The purpose of the intervention may be to provide assistance with mastering the task at hand. Progress in mastering the same or similar tasks is then measured. In essence, dynamic assessment provides a means for evaluating how the assessee processes or benefits from some type of intervention (feedback, hints, instruction, therapy, and so forth) during the course of evaluation. In some educational contexts, dynamic assessment may be viewed as a way of measuring not just learning but "learning potential," or "learning how to learn" skills. Computers are one tool used to help meet the objectives of dynamic assessment (Wang, 2011). There are others . . .

The Tools of Psychological Assessment

The Test

A **test** may be defined simply as a measuring device or procedure. When the word *test* is prefaced with a modifier, it refers to a device or procedure designed to measure a variable related to that modifier. Consider, for example, the term *medical test*, which refers to a device or procedure designed to measure some variable related to the practice of medicine (including a wide range of tools and procedures, such as X-rays, blood tests, and testing of reflexes). In a like manner, the term **psychological test** refers to a device or procedure designed to measure variables related to psychology (for example, intelligence, personality, aptitude, interests, attitudes, or values). Whereas a medical test might involve analysis of a sample of blood, tissue, or the like, a psychological test almost always involves analysis of a sample of behavior. The behavior sample could range from responses to a pencil-and-paper questionnaire, to oral responses to questions related to the performance of some task. The behavior sample could be elicited by the stimulus of the test itself, or it could be naturally occurring behavior (observed by the assessor in real time as it occurs, or recorded).

Psychological tests and other tools of assessment may differ with respect to a number of variables, such as content, format, administration procedures, scoring and interpretation procedures, and technical quality. The *content* (subject matter) of the test will, of course, vary with the focus of the particular test. But even two psychological tests purporting to measure the same thing—for example, personality—may differ widely in item content. This is so because two test developers might have entirely different views regarding what is important in measuring "personality"; different test developers employ different definitions of "personality." Additionally, different test developers come to the test development process with different theoretical orientations. For example, items on a psychoanalytically oriented personality test may

have little resemblance to those on a behaviorally oriented personality test, yet both are personality tests. A psychoanalytically oriented personality test might be chosen for use by a psychoanalytically oriented assessor, and an existentially oriented personality test might be chosen for use by an existentially oriented assessor.

The term **format** pertains to the form, plan, structure, arrangement, and layout of test items as well as to related considerations such as time limits. *Format* is also used to refer to the form in which a test is administered: computerized, pencil-and-paper, or some other form. When making specific reference to a computerized test, the format may also involve the form of the software: PC- or Mac-compatible. The term *format* is not confined to tests. *Format* is also used to denote the form or structure of other evaluative tools and processes, such as the guidelines for creating a portfolio work sample.

Tests differ in their *administration procedures*. Some tests, particularly those designed for administration on a one-to-one basis, may require an active and knowledgeable test administrator. The test administration may involve demonstration of various kinds of tasks demanded of the assessee, as well as trained observation of an assessee's performance. Alternatively, some tests, particularly those designed for administration to groups, may not even require the test administrator to be present while the testtakers independently complete the required tasks.

Tests differ in their *scoring and interpretation procedures*. To better understand how and why, let's define *score* and *scoring*. Sports enthusiasts are no strangers to these terms. For them, these terms refer to the number of points accumulated by competitors and the process of accumulating those points. In testing and assessment, we may formally define **score** as a code or summary statement, usually but not necessarily numerical in nature, that reflects an evaluation of performance on a test, task, interview, or some other sample of behavior. **Scoring** is the process of assigning such evaluative codes or statements to performance on tests, tasks, interviews, or other behavior samples. In the world of psychological assessment, many different types of scores exist. Some scores result from the simple summing of responses (such as the summing of correct/incorrect or agree/disagree responses), and some scores are derived from more elaborate procedures.

Scores themselves can be described and categorized in many different ways. For example, one type of score is the *cut score*. A **cut score** (also referred to as a *cutoff score* or simply a *cutoff*) is a reference point, usually numerical, derived by judgment and used to divide a set of data into two or more classifications. Some action will be taken or some inference will be made on the basis of these classifications. Cut scores on tests, usually in combination with other data, are used in schools in many contexts. For example, they may be used in grading, and in making decisions about the class or program to which children will be assigned. Cut scores are used by employers as aids to decision making about personnel hiring, placement, and advancement. State agencies use cut scores as aids in licensing decisions. There are probably more than a dozen different methods that can be used to formally derive cut scores (Dwyer, 1996). If you're curious about what some of those different methods are, stay tuned; we cover that in an upcoming chapter.

Sometimes no formal method is used to arrive at a cut score. Some teachers use an informal "eyeball" method to proclaim, for example, that a score of 65 or more on a test means "pass" and a score of 64 or below means "fail." Whether formally or informally derived, cut scores typically take into account, at least to some degree, the values of those who set them. Consider, for example, two teachers who teach the same course at the same college. One teacher might set a cut score for passing the course

JUST THINK . . .

Imagine you wanted to develop a test for a personality trait you termed "goth." How would you define this trait? What kinds of items would you include in the test? Why would you include those kinds of items?

that is significantly higher (and more difficult for students to attain) than the other teacher. There is also another side to the human equation as it relates to cut scores, one that is seldom written about in measurement texts. This phenomenon concerns the emotional consequences of “not making the cut” and “just making the cut” (see Figure 1–1).

Tests differ widely in terms of their guidelines for scoring and interpretation. Some tests are self-scored by the testtakers themselves, others are scored by computer, and others require scoring by trained examiners. Some tests, such as most tests of intelligence, come with test manuals that are explicit not only about scoring criteria but also about the nature of the interpretations that can be made from the scores. Other tests, such as the Rorschach Inkblot Test, are sold with no manual at all. The (presumably qualified) purchaser buys the stimulus materials and then selects and uses one of many available guides for administration, scoring, and interpretation.



Figure 1–1
Emotion Engendered by Categorical Cutoffs

People who just make some categorical cutoff may feel better about their accomplishment than those who make the cutoff by a substantial margin. But those who just miss the cutoff may feel worse than those who miss it by a substantial margin. Evidence consistent with this view was presented in research with Olympic athletes (Medvec et al., 1995; Medvec & Savitsky, 1997). Bronze medalists were—somewhat paradoxically—happier with the outcome than silver medalists. Bronze medalists might say to themselves “At least I won a medal” and be happy about it. By contrast, silver medalists might feel frustrated that they tried for the gold and missed winning it.

Tests differ with respect to their **psychometric soundness** or technical quality. Synonymous with the antiquated term *psychometry*, **psychometrics** may be defined as the science of psychological measurement. Variants of these words include the adjective *psychometric* (which refers to measurement that is psychological in nature) and the nouns **psychometrist** and **psychometrician** (both terms referring to a professional who uses, analyzes, and interprets psychological test data). One speaks of the psychometric soundness of a test when referring to how consistently and how accurately a psychological test measures what it purports to measure. Assessment professionals also speak of the psychometric *utility* of a particular test or assessment method. In this context, **utility** refers to the usefulness or practical value that a test or other tool of assessment has for a particular purpose. These concepts are elaborated on in subsequent chapters. Now, returning to our discussion of tools of assessment, meet one well-known tool that, as they say, “needs no introduction.”

JUST THINK . . .

How might one test of intelligence have more utility than another test of intelligence in the same school setting?

The Interview

In everyday conversation, the word *interview* conjures images of face-to-face talk. But the interview as a tool of psychological assessment typically involves more than talk. If the interview is conducted face-to-face, then the interviewer is probably taking note of not only the content of what is said but also the way it is being said. More specifically, the interviewer is taking note of both verbal and nonverbal behavior. Nonverbal behavior may include the interviewee’s “body language,” movements, and facial expressions in response to the interviewer, the extent of eye contact, apparent willingness to cooperate, and general reaction to the demands of the interview. The interviewer may also take note of the way the interviewee is dressed. Here, variables such as neat versus sloppy, and appropriate versus inappropriate, may be noted.

Because of a potential wealth of nonverbal information to be gained, interviews are ideally conducted face-to-face. However, face-to-face contact is not always possible and interviews may be conducted in other formats. In an interview conducted by telephone, for example, the interviewer may still be able to gain information beyond the responses to questions by being sensitive to variables such as changes in the interviewee’s voice pitch or the extent to which particular questions precipitate long pauses or signs of emotion in response. Of course, interviews need not involve verbalized speech, as when they are conducted in sign language. Interviews may also be conducted by various electronic means, as would be the case with online interviews, e-mail interviews, and interviews conducted by means of text messaging. In its broadest sense, then, we can define an **interview** as a method of gathering information through direct communication involving reciprocal exchange.

JUST THINK . . .

What type of interview situation would you envision as ideal for being carried out entirely through the medium of text messaging?

Interviews differ with regard to many variables, such as their purpose, length, and nature. Interviews may be used by psychologists in various specialty areas to help make diagnostic, treatment, selection, or other decisions. So, for example, school psychologists may use an interview to help make a decision about the appropriateness of