

Contents

PART ONE 1

Chapter 1: The World of Psychological Testing 3

Introduction 4

Major Categories of Tests 5

Some Additional Ways to Categorize Tests 8

Uses and Users of Tests 10

Major Issues: Assumptions and Questions 13

Basic Assumptions 14

Fundamental Questions 16

The Differential Perspective 17

The Historical Perspective 18

Remote Background: Up to 1840 18

Setting the Stage: 1840–1880 20

The Roots: 1880–1915 22

The Flowering: 1915–1940 25

Consolidation: 1940–1965 27

Just Yesterday: 1965–2000 28

And Now: 2000–Present 30

Major Forces 31

The Scientific Impulse 31

Concern for the Individual 32

Practical Applications 32

Statistical Methodology 33

The Rise of Clinical Psychology 33

Computers 34

By Way of Definition 37

Summary 40 • Key Terms 41 • Exercises 41

Chapter 2: Sources of Information About Tests 43

Two Common Problems Requiring Information About Tests 43

A Test's Introductory Kit 45

Comprehensive Lists of Tests 46

Systematic Reviews 48

Electronic Listings 49

ETS Test Collection on the Web 49

Buros Reviews Available Electronically 50

| | |
|---|--|
| Health and Psychosocial Instruments (HaPI) | 51 |
| PsycTESTS | 51 |
| Special-Purpose Collections | 51 |
| Books About Single Tests | 52 |
| Textbooks on Testing | 53 |
| Journals | 53 |
| Publishers' Catalogs and Personnel | 55 |
| Other Users | 56 |
| Strengths and Shortcomings of the Sources | 56 |
| Summary | 58 • Key Terms 59 • Exercises 59 |
| Chapter 3: Test Norms | 61 |
| Purpose of Norms | 62 |
| Review of Statistics: Part 1 | 63 |
| Variables | 63 |
| Types of Scales | 64 |
| Organization of Raw Data | 66 |
| Central Tendency | 67 |
| Variability | 68 |
| z-Scores | 69 |
| Shapes of Distributions | 70 |
| The Raw Score | 71 |
| The Special Case of Theta (θ) | 72 |
| Types of Norms | 74 |
| Percentile Ranks and Percentiles | 76 |
| Standard Scores | 81 |
| Developmental Norms | 87 |
| Examples of Norm Tables | 91 |
| Norm Groups | 94 |
| National Norms | 94 |
| International Norms | 95 |
| Convenience Norm Groups | 95 |
| User Norms | 96 |
| Subgroup Norms | 96 |
| Local Norms | 96 |
| Institutional Norms | 98 |
| Criterion-Referenced Interpretation | 98 |
| The Standardization Group: Determining Its Usefulness | 100 |
| Summary | 104 • Key Terms 105 • Exercises 105 |
| Chapter 4: Reliability | 107 |
| Introduction | 108 |
| Four Important Distinctions | 108 |

| | |
|---|------------------------|
| Review of Statistics: Part 2—Correlation and Prediction | 110 |
| Bivariate Distribution and Correlation Coefficients | 110 |
| Regression Line | 113 |
| Factors Affecting Correlation Coefficients | 115 |
| Major Sources of Unreliability | 119 |
| Test Scoring | 120 |
| Test Content | 122 |
| Test Administration Conditions | 122 |
| Personal Conditions | 122 |
| Conceptual Framework: True Score Theory | 123 |
| Methods of Determining Reliability | 126 |
| Test-Retest Reliability | 126 |
| Inter-Scorer Reliability | 127 |
| Alternate Form Reliability | 129 |
| Internal Consistency Reliability | 129 |
| Three Important Conclusions | 135 |
| The Standard Error of Measurement | 135 |
| Standard Errors: Three Types | 138 |
| Some Special Issues in Reliability | 139 |
| Reliability in Item Response Theory | 141 |
| Generalizability Theory | 142 |
| Factors Affecting Reliability Coefficients | 143 |
| How High Should Reliability Be? | 144 |
| Summary | 145 • Key Terms |
| 146 • Exercises | 147 |
| | |
| Chapter 5: Validity | 149 |
| Introduction | 150 |
| Refining the Definition of Validity | 150 |
| Construct Underrepresentation and Construct-Irrelevant Variance | 151 |
| The Basic Issue | 153 |
| The Traditional and Newer Classifications of Types of Validity Evidence | 154 |
| The Issue of Face Validity | 155 |
| Content Validity | 155 |
| Application to Achievement Tests | 156 |
| Instructional Validity | 158 |
| Application to Employment Tests | 159 |
| Content Validity in Other Areas | 160 |
| Problems with Content Validity | 160 |
| Criterion-Related Validity | 161 |
| External, Realistic Criterion | 162 |
| Contrasted Groups | 164 |
| Correlations with Other Tests | 166 |
| Special Considerations for Interpreting Criterion-Related Validity | 168 |
| Combining Information from Different Tests | 174 |

| | |
|--|--|
| Decision Theory: Basic Concepts and Terms | 178 |
| Hits, False Positives, and False Negatives | 178 |
| Base Rate | 179 |
| Sensitivity and Specificity | 180 |
| Construct Validity | 182 |
| Internal Structure | 183 |
| Factor Analysis | 183 |
| Response Processes | 185 |
| Effect of Experimental Variables | 186 |
| Developmental Changes | 186 |
| Consequential Validity | 187 |
| Test Bias as Part of Validity | 188 |
| The Practical Concerns | 189 |
| Integrating the Evidence | 189 |
| In the Final Analysis: A Relative Standard | 190 |
| Summary | 190 • Key Terms 192 • Exercises 192 |

Chapter 6: Test Development, Item Analysis, and Fairness 195

| | |
|--|-----|
| Introduction | 196 |
| Defining the Test's Purpose | 196 |
| Preliminary Design Issues | 198 |
| Origin of New Tests | 199 |
| Item Preparation | 200 |
| Types of Test Items | 201 |
| Selected-Response Items | 202 |
| Scoring Selected-Response Items | 203 |
| Constructed-Response Items | 205 |
| Scoring Constructed-Response Items | 205 |
| The Pros and Cons of Selected-Response versus Constructed-Response Items | 209 |
| Suggestions for Writing Selected-Response Items | 210 |
| Suggestions for Writing Constructed-Response Items | 211 |
| Some Practical Considerations in Writing Items | 211 |
| Item Analysis | 212 |
| Item Tryout | 213 |
| Item Statistics | 214 |
| Item Difficulty | 214 |
| Item Discrimination | 214 |
| Examples of Item Statistics | 216 |
| Item Statistics in Item Response Theory | 219 |
| Factor Analysis as an Item Analysis Technique | 223 |
| Item Selection | 224 |
| Standardization and Ancillary Research Programs | 228 |

Preparation of Final Materials and Publication 229

Fairness and Bias 230

Gaining Perspective on the Issue of Fairness 230

Methods of Studying Test Fairness 233

Accommodations and Modifications 240

Some Tentative Conclusions 242

Summary 243 • Key Terms 244 • Exercises 244

PART TWO 247

Chapter 7: Intelligence: Theories and Issues 249

Intelligence: Areas of Study 250

The Meaning of Intelligence 250

The Real-World Correlates of Intelligence 251

What to Call Them? 253

Theories of Intelligence 253

Two Classical Theories 254

Hierarchical Models 258

Developmental Theories 262

Information Processing and Biological Theories of Intelligence 263

Current Status of Testing in Relation to the Theories 269

Group Differences in Intelligence 270

Differences by Sex 272

Differences by Age 273

Differences by Socioeconomic Level 275

Differences by Racial/Ethnic Group 275

Heredity and Environment 276

Common Misconceptions 277

Methodology and Terms 277

Major Results 278

Summary 279 • Key Terms 280 • Exercises 280

Chapter 8: Individual Tests of Intelligence 283

Some Cases 284

Uses and Characteristics of Individual Intelligence Tests 284

Typical Items in an Individual Intelligence Test 287

The Wechsler Scales: An Overview 289

Historical Introduction 289

Wechsler's Concept of Intelligence 291

Wechsler Adult Intelligence Scale—Fourth Edition 291

Wechsler Intelligence Scale for Children—Fourth Edition 300

WISC versus WAIS 300

| | |
|--|--|
| Psychometric Characteristics of WISC-IV | 301 |
| The Stanford-Binet | 302 |
| Transition to New Structure: SB4 and SB5 | 303 |
| Psychometric Characteristics of SB5 | 305 |
| Brief Individually Administered Tests of Mental Ability | 305 |
| Peabody Picture Vocabulary Test | 305 |
| Two Other Entries | 310 |
| A Test of a Specific Mental Ability: The Wechsler Memory Scale | 310 |
| Intellectual Disability and Mental Retardation: Changing Terminology | 315 |
| The Concept of Adaptive Behavior | 316 |
| Definition of Intellectual Disability | 317 |
| Infant and Early Childhood Tests | 323 |
| Other Applications | 324 |
| Trends in Individual Intelligence Testing | 325 |
| Summary | 327 • Key Terms 328 • Exercises 328 |
| Chapter 9: Group Tests of Mental Ability | 331 |
| Some Cases | 332 |
| Uses of Group-Administered Mental Ability Tests | 332 |
| Common Characteristics of Group Mental Ability Tests | 333 |
| Mental Ability Tests in School Testing Programs | 336 |
| Otis-Lennon School Ability Test | 337 |
| College Admissions Tests | 345 |
| The SAT | 346 |
| The ACT | 348 |
| Graduate and Professional School Selection | 354 |
| Graduate Record Examinations: General Test | 355 |
| Military and Business Selection Tests | 360 |
| Armed Services Vocational Aptitude Battery | 361 |
| Wonderlic Personnel Test | 363 |
| Culture-Fair Tests of Mental Ability | 365 |
| Raven's Progressive Matrices | 365 |
| Other Culture-Fair Tests and Some Conclusions | 366 |
| Intelligence Tests for Micro-Cultures | 369 |
| Generalizations About Group Mental Ability Tests | 369 |
| Summary | 371 • Key Terms 372 • Exercises 372 |
| Chapter 10: Neuropsychological Assessment | 375 |
| Case Examples | 376 |
| Focus on the Brain: The Road to Clinical Neuropsychology | 376 |
| Two Main Approaches to Neuropsychological Assessment | 381 |

| | |
|---|------------|
| Fixed Battery Approach | 381 |
| Luria-Nebraska Neuropsychological Battery | 382 |
| Halstead-Reitan Neuropsychological Test Battery | 383 |
| Flexible Battery Approach | 383 |
| Mental Status | 385 |
| Intelligence | 388 |
| Achievement | 388 |
| Attention/Concentration | 388 |
| Language | 389 |
| Visuospatial/Perceptual | 390 |
| Memory | 391 |
| Motor Functioning | 393 |
| Executive Functions | 393 |
| Personality/Psychological State | 396 |
| Supplementary Information | 398 |
| Medical History | 399 |
| Psychiatric History | 399 |
| Psychosocial History | 399 |
| School Records | 400 |
| Collateral Information | 400 |
| Behavioral Observations | 400 |
| Case Examples Revisited | 401 |
| Summary | 406 |
| • Key Terms | 407 |
| • Exercises | 407 |

Chapter 11: Achievement Tests 409

| | |
|---|-----|
| Introduction | 410 |
| The Ability-Achievement Continuum | 410 |
| The Psychologist's Interface with Achievement Tests | 411 |
| A Broad Classification of Achievement Tests | 412 |
| A Typical School Testing Program | 413 |
| The Accountability Movement and Standards-Based Education | 413 |
| Achievement Batteries | 415 |
| Stanford Achievement Test | 415 |
| Typical Uses and Special Features | 419 |
| Achievement Batteries at the College Level | 420 |
| Single-Area Achievement Tests | 421 |
| Examples | 421 |
| Typical Uses and Special Features | 423 |
| Licensing and Certification Tests | 424 |
| Examples | 424 |
| Typical Uses and Special Features | 425 |
| A Primer on Establishing Cutoff Scores | 426 |
| State, National, and International Achievement Tests | 427 |
| State Testing Programs | 427 |

| | |
|--|--|
| A National Testing Program: NAEP | 428 |
| International Testing Programs: TIMSS, PIRLS, and PISA | 429 |
| Special Features | 430 |
| Individually Administered Achievement Tests | 430 |
| Examples | 430 |
| Typical Uses and Special Features | 431 |
| Some Nagging Questions About Achievement Tests | 433 |
| Summary | 435 • Key Terms 436 • Exercises 436 |

Chapter 12: Objective Personality Tests 439

| | |
|--|--|
| Introduction | 440 |
| Uses of Objective Personality Tests | 441 |
| A Functional Classification of Objective Personality Tests | 442 |
| Comprehensive Inventories: Common Characteristics | 443 |
| Specific Domain Tests: Common Characteristics | 446 |
| The Special Problems of Response Sets and Faking | 447 |
| Strategies for Dealing with Response Sets and Faking | 449 |
| Major Approaches to Personality Test Development | 451 |
| Content Method | 452 |
| Criterion-Keying Approach | 453 |
| Factor Analysis | 455 |
| Theory-Driven Approach | 456 |
| Combined Approaches | 456 |
| Examples of Comprehensive Inventories | 457 |
| The Edwards Personal Preference Schedule (EPPS): An Example of a Theory-Based Test | 457 |
| The NEO Personality Inventory-3: An Example of a Factor-Analytic Test | 459 |
| IPIP: Build Your Own Personality Inventory | 463 |
| Specific Domain Tests | 464 |
| Piers-Harris Children's Self-Concept Scale | 464 |
| Measures Within Positive Psychology | 466 |
| An Example: Subjective Well-Being | 468 |
| Trends in the Development and Use of Objective Personality Tests | 469 |
| Summary | 470 • Key Terms 471 • Exercises 471 |

Chapter 13: Clinical Instruments and Methods 473

| | |
|---|-----|
| Introduction | 474 |
| The Clinical Interview as Assessment Technique | 475 |
| Unstructured, Semistructured, and Structured Interviews | 475 |
| The DSM and ICD | 476 |
| Structured Clinical Interview for DSM-IV Axis I Disorders | 476 |
| The Employment Interview: A Sidebar | 478 |

| | |
|--|--|
| Examples of Comprehensive Self-Report Inventories | 479 |
| The Minnesota Multiphasic Personality Inventory (MMPI) | 479 |
| MMPI-2 RF (Restructured Form) | 487 |
| The Millon Clinical Multiaxial Inventory (MCMI) and the Millon Family Symptom Checklist-90-R | 488 |
| Examples of Specific Domain Tests | 493 |
| The Beck Depression Inventory (BDI) | 494 |
| The Eating Disorder Inventory (EDI) | 495 |
| State-Trait Anxiety Inventory | 498 |
| Behavior Rating Scales | 500 |
| Multiscore Systems | 501 |
| Single-Area Scales | 502 |
| Behavioral Assessment | 502 |
| Direct or Naturalistic Observation | 504 |
| Analogue Behavioral Observation | 504 |
| Behavioral Interviewing | 505 |
| Self-Monitoring and Self-Report | 505 |
| Cognitive-Behavioral Assessment | 505 |
| Physiological Measures | 506 |
| Concluding Comments on Behavioral Assessment Methods | 506 |
| Trends in the Development and Use of Clinical Instruments | 506 |
| Summary | 508 • Key Terms 509 • Exercises 509 |
| Chapter 14: Projective Techniques | 511 |
| General Characteristics of Projective Techniques and the Projective Hypothesis | 512 |
| Uses of Projective Techniques | 513 |
| Indicators for the Use of Projectives | 515 |
| Administration and Scoring of Projective Techniques: A Forewarning | 515 |
| The Rorschach Inkblot Test | 516 |
| The Materials | 516 |
| Administration and Scoring | 519 |
| The Coding System | 521 |
| Sequence of Scores and Structural Summary | 524 |
| Evaluation of the Rorschach | 524 |
| What's Next: Here Comes R-PAS | 525 |
| Thematic Apperception Test (TAT) | 525 |
| Rotter Incomplete Sentences Blank (RISB) | 528 |
| Human Figure Drawings | 531 |
| The Future of Projective Techniques | 534 |
| Training of Psychologists | 534 |
| Emergence of Managed Care | 534 |
| Objective Scoring, Norm-Referenced Interpretation, and Psychometric Quality | 535 |
| Summary | 536 • Key Terms 537 • Exercises 537 |

Chapter 15: Interests and Attitudes 539

- Introduction 540
- Orientation to Career Interest Testing 540
 - Strong and Kuder 540
 - Traditional Approaches 541
 - Uses of Career Interest Tests 542
 - A Forewarning on Names 543
 - Holland Themes and the RIASEC Codes 544
- Strong Interest Inventory 546
- Kuder Career Interests Assessments 552
- Self-Directed Search (SDS) 555
- Some Generalizations About Career Interest Measures 559
- Attitude Measures 560
 - Likert Scales 562
 - Thurstone Scales 564
 - Guttman Scales 566
 - Public Opinion Polls and Consumer Research 567
- Summary 568 • Key Terms 569 • Exercises 569**

Chapter 16: Ethical and Legal Issues 571

- Ethics versus Law 572
- Ethical Issues 572
 - Background on Professional Ethics 573
 - Sources of Ethical Principles for Testing 574
- Generalizations About Ethical Use of Tests 577
 - Competence 577
 - Informed Consent 577
 - Knowledge of Results 578
 - Confidentiality 578
 - Test Security 578
 - Test Development and Publication 579
 - Automated Scoring/Interpretation Systems 579
 - Unqualified Persons 579
 - Test User Qualifications 579
- Legal Issues 581
 - Areas of Application: An Overview 581
 - Definition of Laws 582
 - Laws Related to Testing 583
 - The Fourteenth Amendment 585
 - The Civil Rights Acts of 1964 and 1991 586
 - Rehabilitation Act of 1973 and Americans with Disabilities Act of 1990 586
 - The Handicapped/Disabled in Education: P.L. 94-142 and IDEA 587
 - FERPA and HIPPA 589
 - EEOC Guidelines 589

| | |
|---|--|
| No Child Left Behind Act | 592 |
| Illustrative Court Cases | 593 |
| Griggs v. Duke Power | 593 |
| Debra P. v. Turlington and GI Forum v. TEA | 594 |
| Larry P. v. Riles, PASE v. Hannon, and Crawford v. Honig | 595 |
| Karraker v. Rent-A-Center | 596 |
| Atkins v. Virginia | 597 |
| New Haven Firefighters Case | 598 |
| Forensic Application of Tests | 598 |
| Two Legal Terms | 599 |
| Three Areas of Special Concern | 600 |
| And Beyond | 600 |
| Some Generalizations About the Interface of Testing and the Law | 600 |
| Summary | 602 • Key Terms 603 • Exercises 603 |

Appendixes 605

| | |
|--|-----|
| A. Test Reviewing and Selection | 605 |
| B. How to Build a (Simple) Test | 613 |
| C. Contact Information for Major Test Publishers | 621 |
| D. Sample Data Sets | 623 |
| E. Answers to Selected Exercises | 625 |

Glossary 629

References 641

Name Index 665

Subject Index 669