

# Contents

List of Figures	xiii
List of Tables	xvii
Preface	xxi
Notes to Students	xxv

## **PART ONE. BASIC CONCEPTS 1**

### **1. Who Uses Tests? And for What Purposes? 3**

Testing and Social Policy	5
<i>Reformers' Hopes for Tests</i>	5
<i>Critics' Discontents</i>	8
How the Testing Enterprise Has Evolved	10
Purchasing Tests	16
<i>Sources of Information</i>	16
<i>Who May Obtain Tests?</i>	18
Decisions for Which Tests Are Used	24
<i>Classification</i>	25
<i>Promoting Self-Understanding</i>	27
<i>Program Evaluation</i>	28
<i>Scientific Inquiry</i>	29

## **2. Varieties of Tests and Test Interpretations 31**

- What Is a Test? 32
  - Psychometric and Impressionistic Styles of Testing* 34
- Classification of Procedures 37
  - Tests of Maximum Performance* 38
  - Measures of Typical Response* 41
- Testing in the Computer Age 46
  - Item Banks* 46
  - Automated Administration* 47
- Inferences, Short and Long 51

## **3. Administering Tests 55**

- Two Specimen Tests 57
  - Block Design* 57
  - Mechanical Comprehension* 59
- Procedure for Test Administration 63
  - Testing a Group* 64
  - Policies Regarding Guessing* 66
  - Establishing Rapport in Testing an Individual* 69
- Testing as a Social Relationship 72
  - Characteristics of the Tester* 74
  - How Testers Distort Results* 76
- Motivation of the Test Taker 79
- Advance Preparation and Coaching 82
  - Pros and Cons of Coaching* 84
  - Dynamic Assessment* 87
- Rights of the Test Taker 89
  - Providing for Physical Handicaps* 95

## **4. Scores and Score Conversions 97**

- Obtaining Raw Scores 99
  - Scoring Procedures* 99
  - Limitations of Raw-Score Scales* 102
- Domain-Referenced Interpretations and the Alternatives 104
  - Three Kinds of interpretation* 104
  - Reading Power: Illustration of a Content Scale* 106
  - Setting Standards* 107
- Conversions of Raw Scores 109
  - Percentiles* 109
  - Standard Scores* 114
  - The Normal Distribution* 119
  - Comparison of Score Conversions* 121
  - Profiles* 123

Norms	125
<i>Characteristics Desired in Norms</i>	125
<i>Differentiated Norms</i>	131
<i>Norms by Calibration</i>	134
Unidimensional Scaling	136
<i>Stretching the Score Scale</i>	137
<i>Applications</i>	140
<i>Cautions</i>	142

## 5. How to Judge Tests: Validation 144

The Need for Critical Scrutiny	145
<i>Sources of Information and Criticism</i>	147
<i>The Testing Standards</i>	148
Methods of Validation Inquiry	150
<i>Introduction to Criterion-oriented Inquiries</i>	152
<i>Introduction to Content-related Inquiries</i>	157
<i>Construct Validation: Explaining Scores and Their Relations</i>	159
Validation Against a Criterion: A Closer Look	160
<i>The Criterion</i>	160
<i>Correlation and Regression Coefficients</i>	161
<i>Validity in Classification</i>	167
Checking Out Test Content	170
<i>Selection of Content</i>	170
<i>Format</i>	174
<i>Statistical Properties of Items</i>	176
More on Construct Validation	178
<i>Inquiries Contributing to Explanation</i>	180
<i>Strong and Weak Construct Validation</i>	183
Validation as Persuasive Argument	185

## 6. How to Judge Tests: Reliability and Other Qualities 190

Errors of Measurement	191
<i>True-Score Theory</i>	192
<i>Generalizability Theory</i>	195
Estimation Methods	197
<i>Simple Correlational Studies</i>	197
<i>Estimation from a Two-Way Score Table</i>	199
<i>Internal-Consistency Formulas</i>	202
Interpreting Findings	206
<i>Considering the Range of Scores</i>	206
<i>Number of Observations: Effect on Accuracy</i>	207

<i>Accuracy for Particular Individuals</i>	210
<i>Effect of Measurement Error on Correlations</i>	213
Further Considerations in Choosing Tests	214
Evaluating a Test	217

## **PART TWO. TESTS OF ABILITY 223**

### **7. General Ability: Appraisal Methods 225**

The Emergence of Mental Testing	227
<i>Measurement Before Binet</i>	227
<i>Binet Defines Intelligence</i>	229
<i>Contrasting British Ideas</i>	231
<i>Developments in the United States</i>	232
<i>Appreciating the Diversity of Abilities</i>	236
Score Scales of Mental Tests	240
Illustrative Mental Tests	243
<i>The Wechsler and Its Cousins</i>	243
<i>The Kaufman Batteries</i>	253
<i>The Cognitive Abilities Test and the Stanford-Binet</i>	258
Tests for Admission to Advanced Education	264
Testing in the Early Years	269

### **8. The Meanings of General Ability 274**

Reliability of Scores and Score Differences	276
<i>Test and Subtest Scores</i>	276
<i>Differences and Gains</i>	278
<i>Divergence of Constructs</i>	280
<i>Findings on Part Scores of Mental Tests</i>	280
Changes with Age	282
<i>Change in Level of Performance</i>	283
<i>Consistency of Individual Differences</i>	284
<i>Ability and Later Accomplishment</i>	288
Three Lines of Inquiry	293
How Unitary Is General Ability?	294
<i>Consistency Across Tasks</i>	294
<i>The Fluid-Crystallized Spectrum</i>	296
<i>How Central Is Speed?</i>	301
<i>Ability to Learn</i>	303
Performance Through the Clinician's Eyes	305
Information Processing: The View from the Laboratory	307
<i>Two Illustrative Studies</i>	308
<i>Piaget's Account of Intellectual Development</i>	311
<i>Modeling Bead-Chain Performance</i>	312
<i>Sternberg's Componential Analysis</i>	316
Marks of Excellence	318

## 9. Influences on Intellectual Development 320

- How Nature and Nurture Combine 321
  - The Continuity of Development* 321
  - What Kinship Correlations Do Not Mean* 324
- Physiological Bases for Intellectual Differences 325
  - Hemispheric Specialization* 326
  - Evidence from PET Scans and Electric Potentials* 327
- Cultural Influences 328
  - How Culture Channels Motivation* 328
  - Adapting Testing to the Culture* 329
  - Cultural Shaping of Profiles* 331
- Cultural Loading of Ability Tests 334
  - Comparisons of U.S. Ethnic Groups* 334
  - Should Cultural Experience Count?* 336
  - Detecting Questionable Items* 337
- Intervention to Modify Abilities 340
- Testing and Educational Differentiation 347
  - Decisions About Learners with Intellectual Handicaps* 348
  - Fitting Instruction to the Child* 352
  - Problems in Evaluating Assignment Schemes* 356
- Achievement Tests in Educational Management 358
  - The Role of Objectives* 359
  - Standard Tests* 362
  - An Advanced Placement Test* 365
  - Constructed-Response vs. Choice-Response Tests* 368

## 10. Multiple Abilities and Their Role in Counseling 370

- Sorting Tests by Factor Analysis 372
  - How Factor Analysis Is Carried Out* 372
  - Pictures of the Makeup of CogAT* 375
  - Confirmatory Analysis* 378
  - Multilevel Structures* 379
- Variables and Tests in the Thurstone Tradition 383
  - Guilford's "Structure of Intellect"* 386
  - Divergent Thinking* 387
- Multiscore Aptitude Batteries 390
  - The Differential Aptitude Tests* 390
  - The Armed Services Vocational Aptitude Battery* 393
  - The General Aptitude Test Battery* 394
  - Limitations on Interpretation of Profiles* 399
- Possibilities in Microanalysis 404

## **11. Personnel Selection 406**

- Developing a Selection Procedure 407
  - Job Analysis* 409
  - Choice of Tests for Tryout* 411
  - Experimental Trial* 412
  - The Criterion* 413
  - Translating Predictions into Decisions* 417
- Two Examples 421
  - Selecting Power-Plant Operators* 421
  - An Aptitude Test for Programmers* 423
- Interpreting Predictive Validities 428
  - What Level of Validity Is Acceptable?* 428
  - Adjusting Validity Coefficients* 432
- Validity Generalization 433
- Using Multiple Predictors 437
  - Linear Composites* 437
  - Patterns and Configurations* 440
  - Cross-Validation* 443
- Classification Decisions 445
- Fairness in Selection 446
  - Some History and Law* 446
  - Psychometric Issues* 450
  - Main Points Reviewed* 455

## **PART THREE. MEASURES OF TYPICAL RESPONSE 459**

### **12. Interest Inventories 461**

- Functions of Interest Measurement 462
- Dimensionalizing Interests 464
- Choices in Constructing Inventories 467
  - Simplicity vs. Complexity* 468
  - Item Form* 469
  - Organizing Items into Scales* 470
  - Score Reporting* 473
- Predictive Significance 476
  - Stability* 476
  - Agreement of Inventories with Claimed Interests* 478
  - Relations with Criteria* 478
- Interest Measures in Counseling 480
- Technical Issues 483
  - Two Types of Criterion Key* 483
  - Changes from Decade to Decade* 484
  - Erasing Sex Differences* 485

### 13. General Problems in Studying Personality 487

- The Range of Inquiries 488
- Alternatives to Self-Report 490
  - Should Methods Agree?* 493
- Conceptualizations of Individuality 497
  - Traits as Response Tendencies in Situations* 499
  - Systems of Traits* 503
  - Clinical Inferences* 506
  - Interpreting Without Theory: Criterion Keying* 510
- Stability of Personality 512
- Self-Description: Report of Typical Behavior? 513
  - Limitations of Inventory Responses* 515
  - Coping with Biases and Distortions* 519
- Ethical Issues 522
  - Respecting the Dignity of Those Tested* 523
  - How Permissible Is Deception?* 524
  - Limits on Scientific Inquiry* 525

### 14. Personality Measurement Through Self-Report 528

- Specimen Instruments 530
  - Evaluative Scales* 530
  - Multiscore Descriptions* 537
- Validity for Evaluating Disorders 549
  - Screening* 549
  - Classification* 552
  - Forecasting Response to Psychological Intervention* 554
- Predictions in the General Population 555
  - Occupational Predictions* 555
  - Educational Predictions* 560
- Structure of the Individual 564
- Human Judges as Test Interpreters 569
- The Computer as Test Interpreter 573
  - Reports at Three Levels* 573
  - Hazards and Safeguards* 578

### 15. Judgments and Systematic Observations 582

- Ratings by Supervisors and Professional Observers 583
  - Sources of Error* 583
  - Improvement of Ratings* 586
- Information from Acquaintances and Associates 595
  - Mothers' Reports on Children* 595
  - Peer Ratings* 601

Participants' Reports on Settings 603  
Direct Observation 606  
    *Choice of the Situations* 606  
    *The Schedule for Sampling* 607  
    *Observer Error and Its Reduction* 610

## **16. Inferences from Performance 612**

Performance Tests 613  
    *Procedures* 613  
    *Problems of Design and Validity* 620  
Two Stylistic Tests 623  
    *The Bender Gestalt* 623  
    *The Rorschach* 628  
Thematic Tests 631  
    *The TAT and Its Meaning* 631  
    *Correlates of Responses* 633  
Integrative Assessment 636  
    *History of Assessment Programs* 636  
    *Specialized Tests for Organizational Behavior* 639  
    *Validity, Past and Present* 641  
Toward an Ecological View 646  
  
Appendix A. Selected Publishers and Test  
Distributors 649  
  
Appendix B. Classified Index of Tests and  
Inventories 651  
  
References 657  
  
Glossary 701  
  
Name Index 709  
  
Subject Index 719