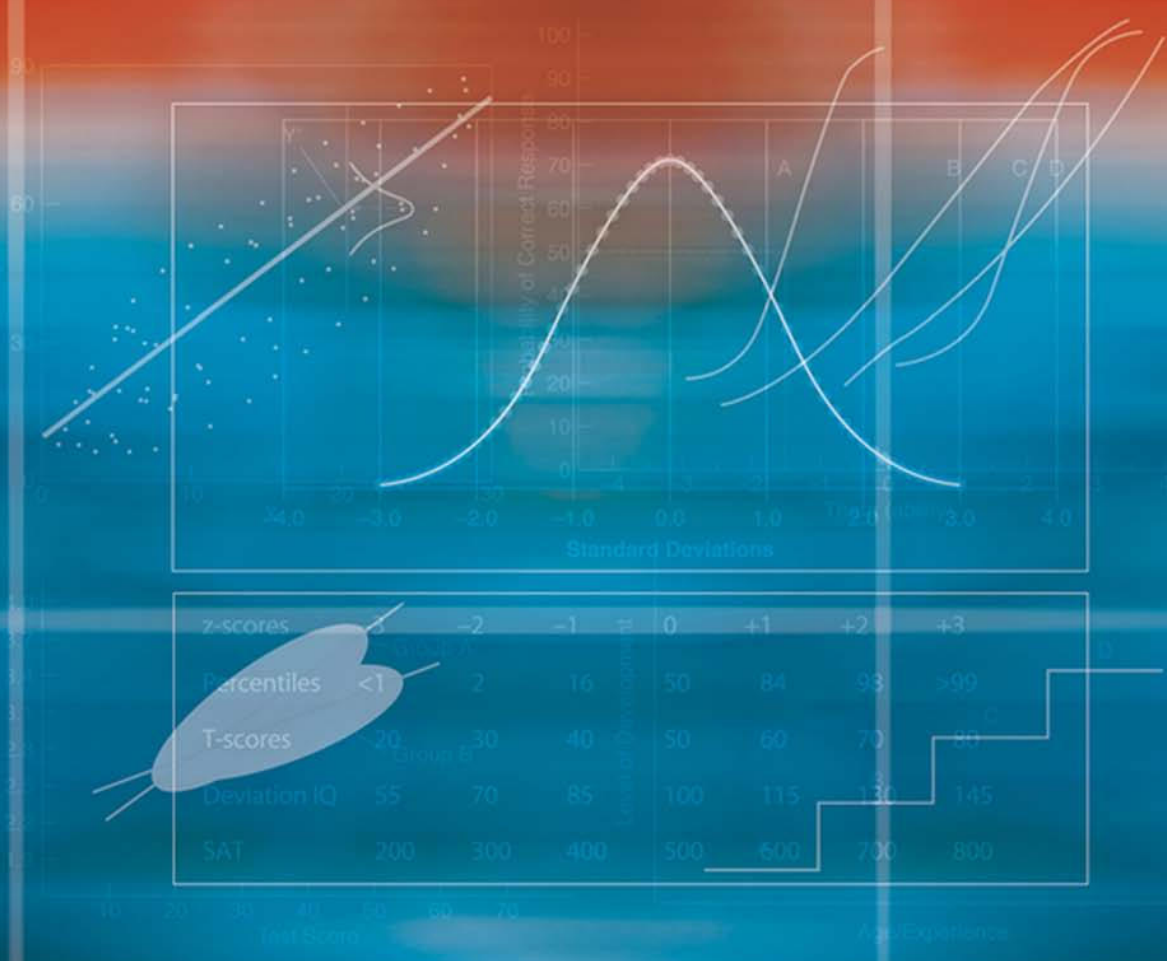


THIRD EDITION

# PSYCHOLOGICAL TESTING

## A Practical Introduction



THOMAS P. HOGAN

WILEY



# PSYCHOLOGICAL TESTING

A Practical Introduction

**Thomas P. Hogan**

University of Scranton

**Third Edition**

**WILEY**

PUBLISHER:	Jay O'Callaghan
ACQUISITIONS EDITOR:	Robert Johnston
ASSISTANT EDITOR:	Brittany Cheetham
EDITORIAL ASSISTANT:	Kristen Mucci
MARKETING MANAGER:	Margaret Barrett
COVER DESIGNER:	Kenji Ngieng
ASSOCIATE PRODUCTION MANAGER:	Joyce Poh
PRODUCTION EDITOR:	Jolene Ling
COVER ILLUSTRATIONS CREDIT:	Thomas P. Hogan

This book was set by Thomson Digital. Cover and text printed and bound by Edwards Brothers Malloy.

This book is printed on acid free paper.

Founded in 1807, John Wiley & Sons, Inc. has been a valued source of knowledge and understanding for more than 200 years, helping people around the world meet their needs and fulfill their aspirations. Our company is built on a foundation of principles that include responsibility to the communities we serve and where we live and work. In 2008, we launched a Corporate Citizenship Initiative, a global effort to address the environmental, social, economic, and ethical challenges we face in our business. Among the issues we are addressing are carbon impact, paper specifications and procurement, ethical conduct within our business and among our vendors, and community and charitable support. For more information, please visit our website: [www.wiley.com/go/citizenship](http://www.wiley.com/go/citizenship).

Copyright © 2015, 2007, 2003 John Wiley & Sons, Inc. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc. 222 Rosewood Drive, Danvers, MA 01923, website [www.copyright.com](http://www.copyright.com). Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, (201)748-6011, fax (201)748-6008, website <http://www.wiley.com/go/permissions>.

Evaluation copies are provided to qualified academics and professionals for review purposes only, for use in their courses during the next academic year. These copies are licensed and may not be sold or transferred to a third party. Upon completion of the review period, please return the evaluation copy to Wiley. Return instructions and a free of charge return mailing label are available at [www.wiley.com/go/returnlabel](http://www.wiley.com/go/returnlabel). If you have chosen to adopt this textbook for use in your course, please accept this book as your complimentary desk copy. Outside of the United States, please contact your local sales representative.

Library of Congress Cataloging-in-Publication Data

Hogan, Thomas P.

Psychological testing : a practical introduction / Thomas P. Hogan, University of Scranton. – Third edition.  
pages cm

Includes bibliographical references and index.

ISBN 978-1-118-55412-8 (pbk.)

1. Psychological tests. I. Title.

BF176.H64 2013

150.28<sup>7</sup>—dc23

2013018015

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

## Preface

### Purpose and Emphases

---

This book provides an introduction to the field of psychological testing. It is intended for the student of psychology and allied disciplines. The book attempts to offer a practical approach, with an emphasis on active learning strategies. Practicality is attained by treating tests in their real-life, contemporary application in psychology. Active learning strategies are emphasized by supplying the student with TRY IT exercises sprinkled throughout the text that require application of the concepts and procedures presented in the text. Too many textbooks on psychological testing are written as reference works, almost like encyclopedias, rather than as textbooks. A textbook should be primarily a learning device. Reference works are useful, but not as textbooks, except perhaps for very advanced students. Research on student learning has amply demonstrated that active engagement with the material maximizes learning. Liberal use is made of Internet resources. Much information formerly inaccessible to students of psychological testing, short of Herculean efforts on the instructor's part, is now readily available on the Internet. The text encourages utilization of these resources. In addition to the exercises incorporated directly into the text, each chapter begins with a list of learning objectives and concludes with a summary of major points, a list of key terms, and additional exercises. Key Point Summaries accompany major chunks of material within chapters. These intermediate summaries should help students stay organized. All these features should aid student learning.

The list of objectives at the beginning of each chapter should serve as “advance organizers” to help focus students’ attention. The summary at the end of each chapter will aid in bringing closure, sometimes after a difficult journey through the material. The list of key terms should supplement the summary points. An ample supply of exercises is given at the end of each chapter. Their character varies. Some emphasize technical matters, others are “thought” questions, and still others call on students to find information through Internet resources. No one is expected to complete all these exercises. However, the instructor can make a judicious selection to be performed by individuals or small groups of students. The end-of-chapter exercises are more challenging than the TRY IT examples embedded in the main text, but I have tried to design most of these exercises so that each can be completed in 10 to 20 minutes, some in as little as 2 to 3 minutes. Students should find that completing at least a few of these exercises helps understanding and retention of the material.

The book’s emphasis on practicality does not imply a lack of rigor or avoidance of difficult topics. On the contrary, the text meets hard material head-on. Psychological testing is not an easy subject. The text does not attempt to “dumb down” this material. It does, however, attempt to show the actual practice of psychological testing and give examples of the concepts and procedures of testing in contemporary applications.

## Student Background

---

When preparing the text, I assumed that students had completed a course in elementary statistics, including the following topics: methods for tabular and graphic summaries of raw data, including shapes of distributions; measures of central tendency, variability, and (zero-order) correlation and regression; the elements of sampling theory, interval estimation, and hypothesis testing. I have also assumed that the students have forgotten a substantial amount of this material. Hence, the text provides “refreshers” on various topics from elementary statistics throughout the first several chapters. Because of students’ predisposition to avoid anything appearing in an appendix, I have incorporated these statistics refreshers into the main text. The unusual student who actually remembers most of what was covered in elementary statistics can skip the refreshers. Instructors will need to use their own judgment on how much time to spend on these refreshers. In addition, the text covers certain topics from multivariate statistics, especially multiple correlation and factor analysis. However, most students will not have had exposure to these topics, and no prior knowledge of these topics is assumed.

## Organization

---

The book divides naturally into two major sections. Part I covers the basic concepts of psychological testing. Chapter 1 presents an overview of the field, including the typical uses of psychological tests. Chapter 2 provides a much more complete treatment of sources of information about tests than is available elsewhere. I have done this for two reasons. First, students, as well as others, frequently ask the question “Is there a test that measures?” I want students using this text to know how to answer that common question. Second, in later chapters, students are asked to use the sources of information covered in Chapter 2 to find real-life examples of concepts, procedures, and examples of tests.

Chapters 3–6 cover the foundational topics of norms, reliability, validity, test development, and test fairness. In each of these chapters, most of the material should be covered with all students. The remaining material is somewhat more advanced or more technical. Instructors may include all of this more advanced material, some of it, or none of it, depending on their areas of interest and on the special needs of their students. Chapter 6 on test development is complemented by Appendix B on steps in building a simple test. Some instructors require students to build a simple test as a class exercise, and Appendix B should be helpful in these cases. The small data sets in Appendix D may be used with standard statistical packages for practical exercises related to the concepts in Chapters 3–6. Selected end-of-chapter exercises call for use of these data sets. Some instructors will want to use their own data sets instead.

The balance to be struck between classical test theory (CTT) and item response theory (IRT) in Chapters 3–6 presents a particular challenge. CTT is hard enough for novices to swallow; IRT often evokes a choking response. In current practice, IRT procedures are now applied routinely. Students will encounter IRT procedures, as well as CTT procedures, in nearly all recently developed tests. Even the beginning student needs familiarity with the concepts and language of IRT. Therefore, in

accordance with my intent to be practical, Chapters 3–6 provide a good dosage of IRT, as well as CTT, procedures. Of course, instructors will need to strike the right balance for the needs of their own students.

Part II provides an introduction to the major categories of psychological tests. For each category, a chapter outlines the major conceptual and procedural approaches and gives a *few* examples of tests in the category. I have tried to resist the temptation to list test after test with brief descriptions of each because I do not believe that students learn anything useful from such cataloging. In fact, they probably learn nothing at all from such lists. Ordinarily, when a test is introduced in Part II, I have tried to describe it in enough detail and in such a way that the introductory student can learn something from it. I have made just a few exceptions to this rule for extraordinary cases. Selection of examples for the chapters in Part II has been guided *primarily* by frequency of actual test usage—this in accordance with the intention to be practical—and secondarily by the desire to illustrate some variety in approaches to testing within a category. While covering the chapters in Part II, I hope instructors will have a supply of introductory kits to show their students. It would be convenient to have kits for all the major examples used in the text. However, if these are not available but kits for alternative tests are, it is probably preferable to use these alternatives as examples rather than the examples given in the text. Particularly for novices in the field, it is not a very meaningful experience to read about a test without seeing it and literally feeling it.

Part II concludes with a chapter (16) on ethical and legal issues. The chapter does not fit thematically with the other chapters in Part II. However, the chapter is clearly necessary. I did not think students could appreciate some of the issues treated in this chapter until they had completed all other chapters. Not wanting a Part III with only one chapter, rather like the forbidden one-sentence paragraph, I simply included the chapter on ethical and legal issues at the end of Part II.

## New to the Third Edition

---

The third edition preserves the major features of the second edition, while introducing the following major changes:

- **Significant expansion in the treatment of test fairness/test bias:** Test fairness has risen steadily in the psychometric constellation of crucial topics. Earlier editions of the book covered the topic but in scattered sections. These sections have now been “pulled together,” coordinated, and expanded. Test fairness logically belongs with validity. However, as noted in the text, much of the practical work on fairness occurs during the test development process, so I decided to place the material on test fairness in the chapter (6) on test development.
- **Coordination with the new *Standards for Educational and Psychological Testing*:** The new Standards, under revision for nearly five years, appeared about the same time as this new edition and, in fact, the new edition was intentionally targeted to reflect the new Standards. Citations to the new Standards appear throughout the text, especially in Chapters 3–6.

- **Thorough updating for new versions of widely used tests:** Staying abreast of today's dizzying pace of new test development was perhaps the most daunting part of preparing this new edition. Roughly half of all the tests covered in this third edition are new or newly revised tests. Included among these (limiting ourselves to initials here) are the following:

WAIS-IV	ASVAB
WMS-IV	NBAP-D
PPVT-4	MMSE-2
WPT	NEO PI-3
GRE	MMPI-2 RF

plus the latest developments for school testing programs in light of NCLB and for the Rorschach. Numerous embellishments appear for other tests covered in the text. The reader will also encounter a substantial addition to the references, with special emphasis on the research literature of the past half-dozen years, accompanied by a judicious paring of earlier references.

Appendix C includes updated contact information for major test publishers. Virtually all contact with publishers these days begins with the Internet. Therefore, the updated information gives publisher URLs and omits addresses and telephone numbers.

Added to the sample data sets in Appendix D is a simple Excel spreadsheet that allows the student to generate item characteristic curves (ICCs) by varying item parameters in the three-parameter model.

The Instructor's Manual and Test Bank have been thoroughly updated to reflect the changes in this new edition.

## To the Student

---

If you will work in some field of psychology, you will almost certainly deal with psychological tests. Thus, learning about tests has very practical applications. I hope that this text will help you understand the basic issues in psychological testing.

Here are suggestions for using the text effectively.

- Review the objectives at the beginning of each chapter to alert you to what's important in that chapter.
- Carefully observe the "key terms." They are called "key" because they are key! Every key term is in boldface type in the text and listed at the end of the chapter. And, if it's in the list of key terms at chapter's end, it's in the Glossary.
- Complete all of the TRY IT exercises scattered through the text. Doing so will help the material "sink in." Note that each one will only take a minute or so.
- Use the intermediate "Key Point Summaries" throughout the text to help you stay focused and organized. When you reach one of these, pause to go back over what has just been covered.
- Use the Summary points at the conclusion of each chapter to review the major themes and points covered in the chapter.



- Complete at least some of the Exercises at the end of each chapter. You are not expected to do all of them, but do some. Like the TRY IT exercises, the end-of-chapter exercises will help make the material practical.
- Finally, note that psychological testing is not an easy subject. Study hard!

## Acknowledgments

---

Accounting for the myriad contributions to preparation of this book is a daunting and humbling task. So many people have done so much to assist me. I am very grateful, especially to the following. To all my students over many years, for their willingness to suggest ways to present concepts about testing in effective ways, with special thanks to Allyson Kiss and Matthew Sabia for assistance with research and manuscript preparation for this third edition. To the many publishers who granted permission for reproductions and to their staff who gave helpful advice about their test products. I am especially grateful to the following individuals who provided feedback on their actual use of the book as well as comments on the plan for revisions of this and earlier editions: Ira H. Bernstein, University of Texas—Arlington; Jeffrey B. Brookings, Wittenberg University; Douglas Maynard, SUNY—New Paltz; Robert Resnick, Randolph Macon College; and Marie D. Thomas, California State University—San Marcos. In addition, my thanks to the following individuals who provided useful comments and suggestions on revised chapters for this and earlier editions: Julie Alvarez—Tulane University; David Bush—Utah State University; Mark Lenzenweger, State University of New York—Binghamton; Stuart McKelvie—Bishop’s University; John Suler, Rider University; Stefan Schulenberg, University of Mississippi; and David Trumpower, Marshall University. All of the above individuals helped to create an improved textbook.

On a longer term basis, my continuing gratitude goes to my academic mentor, the renowned Anne Anastasi; to Dorothea McCarthy, who arranged for my first job in the field of testing; and to Joseph Kubis for his pedagogical flair. To my professional mentors, Dena Wadell and Roger Lennon, who showed me the interface of theory and practice. To my professional colleagues, beginning with William Tsushima, whose early help was more important than he can possibly imagine. Very special thanks to my University of Scranton colleague, John Norcross, who served as a sounding board on a host of issues. And, of course, to Brooke Cannon and Matthew Eisenhard for their excellent contribution in the form of the chapter (10) on neuropsychological assessment. Finally, I want to thank my wife, Peg, and our children for moral support (and some really helpful suggestions) throughout the endeavor.

With all that help, you’d think the book would be perfect. Alas, that is probably not the case. I must take responsibility for any imperfections that may have crept into the work.

Thomas P. Hogan  
May 2013  
Scranton, Pennsylvania



# Brief Contents

## PREFACE III

## PART ONE 1

---

- CHAPTER 1. THE WORLD OF PSYCHOLOGICAL TESTING 3
- CHAPTER 2. SOURCES OF INFORMATION ABOUT TESTS 43
- CHAPTER 3. TEST NORMS 61
- CHAPTER 4. RELIABILITY 107
- CHAPTER 5. VALIDITY 149
- CHAPTER 6. TEST DEVELOPMENT, ITEM ANALYSIS, AND FAIRNESS 195

## PART TWO 247

---

- CHAPTER 7. INTELLIGENCE: THEORIES AND ISSUES 249
  - CHAPTER 8. INDIVIDUAL TESTS OF INTELLIGENCE 283
  - CHAPTER 9. GROUP TESTS OF MENTAL ABILITY 331
  - CHAPTER 10. NEUROPSYCHOLOGICAL ASSESSMENT 375
  - CHAPTER 11. ACHIEVEMENT TESTS 409
  - CHAPTER 12. OBJECTIVE PERSONALITY TESTS 439
  - CHAPTER 13. CLINICAL INSTRUMENTS AND METHODS 473
  - CHAPTER 14. PROJECTIVE TECHNIQUES 511
  - CHAPTER 15. INTERESTS AND ATTITUDES 539
  - CHAPTER 16. ETHICAL AND LEGAL ISSUES 571
- 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



# 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
<b>Summary</b>	<b>58</b>
<b>Key Terms</b>	<b>59</b>
<b>Exercises</b>	<b>59</b>

### **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 ( $\theta$ )	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
<b>Summary</b>	<b>104</b>
<b>Key Terms</b>	<b>105</b>
<b>Exercises</b>	<b>105</b>

### **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
<b>Summary</b>	<b>145</b>
<b>Key Terms</b>	<b>146</b>
<b>Exercises</b>	<b>147</b>
<b>Chapter 5: Validity</b>	<b>149</b>
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
<b>Summary</b>	<b>190</b>
<b>• Key Terms</b>	<b>192</b>
<b>• Exercises</b>	<b>192</b>
<b>Chapter 6: Test Development, Item Analysis, and Fairness</b>	<b>195</b>
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
<b>Summary</b>	<b>243</b>
<b>Key Terms</b>	<b>244</b>
<b>Exercises</b>	<b>244</b>

## **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
<b>Summary</b>	<b>279</b>
<b>Key Terms</b>	<b>280</b>
<b>Exercises</b>	<b>280</b>

### **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
<b>Summary</b>	<b>327</b>
<b>Key Terms</b>	<b>328</b>
<b>Exercises</b>	<b>328</b>
<b>Chapter 9: Group Tests of Mental Ability</b>	<b>331</b>
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
<b>Summary</b>	<b>371</b>
<b>Key Terms</b>	<b>372</b>
<b>Exercises</b>	<b>372</b>
<b>Chapter 10: Neuropsychological Assessment</b>	<b>375</b>
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
<b>Summary</b>	<b>406</b>
<b>Key Terms</b>	<b>407</b>
<b>Exercises</b>	<b>407</b>

## **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
<i>Summary</i>	<i>435</i>
<i>Key Terms</i>	<i>436</i>
<i>Exercises</i>	<i>436</i>
<b>Chapter 12: Objective Personality Tests</b>	<b>439</b>
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
<i>Summary</i>	<i>470</i>
<i>Key Terms</i>	<i>471</i>
<i>Exercises</i>	<i>471</i>
<b>Chapter 13: Clinical Instruments and Methods</b>	<b>473</b>
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
<b>Summary</b>	<b>508</b>
<b>Key Terms</b>	<b>509</b>
<b>Exercises</b>	<b>509</b>

## **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
<b>Summary</b>	<b>536</b>
<b>Key Terms</b>	<b>537</b>
<b>Exercises</b>	<b>537</b>

**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
<b>Summary</b>	<b>602</b>
<b>Key Terms</b>	<b>603</b>
<b>Exercises</b>	<b>603</b>

**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**





# PART ONE

Part one of this book provides an overview of the field of psychological testing, then concentrates on fundamental principles and procedures applicable to all types of tests. Chapter 1 introduces the student to the field as it exists today. This chapter also sketches how the field came to be this way. Chapter 2 reviews sources of information that students may use to help find out more about tests. These sources are used frequently in later chapters to identify tests employed for particular purposes. Hence, it is important for the student to learn the use of these sources.

Chapters 3–6 present the fundamental principles used to judge all types of tests. These chapters cover norms (3), reliability (4), validity (5), and test development, including fairness (6). This is not easy material. However, it is essential that the student learn these basic concepts, since they are the basis for treatment of specific tests in Part two of the book. The student should complete the little TRY IT exercises sprinkled throughout the text to help the material “sink in.” Although Chapters 1 and 2 may be read in a leisurely manner, Chapters 3–6 require intense concentration and much review. Studying these chapters properly will pay big dividends when the later chapters arrive.



# CHAPTER

# 1

---

## The World of Psychological Testing

### *Objectives*

1. List the major categories of tests, giving at least one example for each category.
2. Identify the major uses and users of tests.
3. Summarize major assumptions and fundamental questions involved in testing.
4. Outline significant features of the major periods in the history of testing.
5. Identify the six major forces influencing the development of testing.
6. Give a definition of a “test.”

## Introduction

---

This chapter provides an overview of the world of testing. Of course, everyone knows, at least roughly, what we mean by a “test” or “testing.” Everyone has at least some familiarity with a variety of tests, for example, college admission tests, final examinations in courses, vocational interest inventories, and perhaps some personality measures. However, as we begin formal study of this world, it is important to develop both a more comprehensive and a more precise understanding of the field. “More comprehensive” so that we consider all types of tests and all relevant issues: We do not want to miss anything important. “More precise” so that we begin to acquire the technical expertise needed by professionals within the broader fields of psychology and allied disciplines: We will not be satisfied with just a passing acquaintance with these topics.

This is an ambitious agenda for one chapter. However, this opening chapter seeks only to provide an overview of these matters. The remaining chapters supply the details. There are a variety of ways to accomplish our goal of providing an overview and orientation to the field. No single way is best. We will use five perspectives or approaches to introduce the field, viewing it, as it were, from different angles or through different lenses. First, we outline the major categories of tests. Most of these categories correspond to chapters in the latter half of this book. In the process of describing these major categories, we mention examples of some of the more widely used tests. Second, we identify the major uses and users of tests. Who actually uses these tests and for what purposes? Third, we outline the primary issues that we worry about in testing. Notice that this outline—the list of principal worries—corresponds to the chapters in the first half of the book. Fourth, we trace the historical roots of contemporary testing. We mark off major periods in this history and identify some major forces that have shaped the field. Fifth, we examine some of the attempts to define *test*, *testing*, and some related terms. When we finish viewing the field through these five perspectives, we should have a good overview of the field of testing.

### Key Points Summary 1.1

#### *Five Ways to Introduce the Field*

1. Categories of Tests
2. Uses and Users of Tests
3. Issues: Assumptions and Questions
4. Historical Periods and Forces
5. By Definition

## Major Categories of Tests

---

We begin our exploration of the world of testing by identifying the major categories of tests. Any such classification is necessarily fuzzy around the edges. Categories often blend into one another rather than being sharply different. Nevertheless, an organizational scheme helps us to comprehend the breadth of the field. Key Points Summary 1.2 provides the classification scheme we use throughout the book. In fact, Chapters 8–15 follow this organization. This introductory chapter just touches on the major categories. Each category receives in-depth treatment later.

The first major division encompasses **mental ability tests**. In the world of psychological testing, the term *mental ability* includes a wide variety of cognitive functions, such as memory, spatial visualization, and creative thinking. Historically, the area has centered on intelligence, broadly defined. This category subdivides into individually administered intelligence tests, group-administered intelligence tests, and a variety of other ability tests, that is, other than intelligence tests. An example of an individually administered intelligence test is the *Wechsler Adult Intelligence Scale*,<sup>1</sup> abbreviated WAIS. Another classic example in this category is the *Stanford-Binet Intelligence Scale*. These tests are administered to individual examinees, one-on-one, by trained psychologists to provide an index of the overall mental ability of individuals. An example of a group-administered intelligence test is the *Otis-Lennon School Ability Test (OLSAT)*. This test is administered to groups of students, usually in classroom settings, to gauge mental ability to succeed in typical school subjects. Another example of tests in this category is the SAT<sup>2</sup> used to predict success in college.

### **TRY IT!**.....

To see how we cover a category in more depth later, flip to page 291. Quickly scan pages 291–300. You will see how subsequent chapters give details about tests mentioned in this opening chapter.

.....

There are many other types of mental ability tests—nearly an infinite variety—including tests of memory, quantitative reasoning, creative thinking, vocabulary, and spatial ability. Sometimes these mental functions are included in the tests of general mental ability, but sometimes they are tested separately.

The next major category includes **achievement tests**. These tests attempt to assess a person's level of knowledge or skill in a particular domain. We cover here only professionally developed, standardized tests. We exclude the vast array of teacher-made tests used daily in the educational enterprise. Even excluding teacher-made

---

<sup>1</sup> In this opening chapter, we refer only to the first editions of tests. In subsequent chapters, we refer to the more recent editions and their corresponding initials, for example, WAIS-IV, MMPI-2, and so on.

<sup>2</sup> For many years, this test was titled the *Scholastic Aptitude Test*. The title was officially changed to the *Scholastic Assessment Test* in 1992 and later simply to the initials SAT. The old titles still appear in many publications. The reference here is specifically to the SAT I: Reasoning Test. The SAT II: Subject Tests is a series of tests in specific fields such as literature, French, and chemistry.

## Key Points Summary 1.2

### *Major Categories of Tests*

- Mental Ability Tests
  - Individually Administered
  - Group Administered
  - Other Abilities
- Achievement Tests
  - Batteries
  - Single Subject
  - Certification, Licensing
  - Government-sponsored Programs
  - Individual Achievement Tests
- Personality Tests
  - Objective Tests
  - Projective Techniques
  - Other Approaches
- Interests and Attitudes
  - Vocational Interests
  - Attitude Scales
- Neuropsychological Tests

tests, achievement tests are easily the most widely used of all types of tests. The first subdivision in this area includes achievement batteries used in elementary and secondary schools. Nearly everyone reading this book will have taken one or more of these achievement batteries. Examples include the *Stanford Achievement Test*, the *Metropolitan Achievement Tests*, and the *Iowa Tests of Basic Skills*. All these batteries consist of a series of tests in such areas as reading, mathematics, language, science, and social studies. The second subdivision includes single-subject tests that cover only one area, such as psychology, French, or geometry. An example of such a test—one that many readers of this book have taken or will take—is the *Graduate Record Examinations (GRE): Psychology Test*.

The third subdivision includes the incredible variety of tests used for purposes of certification and licensing in such fields as nursing, teaching, physical therapy, airline piloting, and so on. None of the tests in this category is a household name. But they have important consequences for people in specific vocational fields.

Fourth, various government agencies sponsor certain achievement testing programs. Most prominent among these are statewide achievement testing programs in such basic subjects as reading, writing, and mathematics. In fact, such state assessment programs have assumed enormous importance in recent years as a result of new

federal laws. In some states, high school graduation depends partly on performance on these tests. Other government-sponsored programs provide information about nationwide performance in a variety of areas. The best known of these efforts are the National Assessment of Educational Progress (NAEP) and the Trends in International Mathematics and Science Study (TIMSS), both of which are the subject of frequent reports in the media.

Finally, there are individually administered achievement tests. The first four types of achievement tests are typically group administered. However, some achievement tests are individually administered in much the same way as individually administered mental ability tests. The individually administered achievement tests aid in the diagnosis of such conditions as learning disabilities.

The next major category includes the variety of tests designed to yield information about the human personality. The first subdivision includes what we call **objective personality tests**. In testing parlance, objective simply means the tests are objectively scored, based on items answered in a true–false or similar format. Examples of these objective personality tests are the *Minnesota Multiphasic Personality Inventory*, abbreviated MMPI, the *Beck Depression Inventory* (BDI), and the *Eating Disorder Inventory* (EDI). The MMPI provides a profile showing how similar the examinee’s responses are to the responses of several clinical groups. The BDI and EDI, as suggested by their titles, try to measure depression and eating disorders, respectively. For both convenience and conceptual clarity, in subsequent chapters we divide these objective tests into those designed to measure personality traits within the normal range and those designed as clinical instruments to measure pathological or disabling conditions.

### **TRY IT!** .....

Part of becoming a professional in this field involves learning the initials for tests. The initials are used routinely in psychological reports and journal articles, often without reference to the full name of the test. Become accustomed to this! Without referring to the text, see if you can give the full test title for each of these sets of initials:

EDI \_\_\_\_\_                      WAIS \_\_\_\_\_  
 GRE \_\_\_\_\_                      MMPI \_\_\_\_\_

.....

The second major subdivision of personality tests includes **projective techniques**. With all these techniques, the examinee encounters a relatively simple but unstructured task. We hope that the examinee’s responses will reveal something about his or her personality. The most famous of these techniques is the *Rorschach Inkblot Test*—sometimes just called the Rorschach, other times called the inkblot test. Other examples are human figure drawings, sentence completion techniques, and reactions to pictures. We include under personality measures a third category, simply labeled “other approaches,” to cover the myriad of other ways psychologists have devised to satisfy our limitless fascination with the human personality.

The next major category of tests encompasses measures of interests and attitudes. The most prominent subdivision in this category includes **vocational interest measures**. These tests are widely used in high schools and colleges to help individuals explore jobs relevant to their interests. Examples of such tests are the *Strong Interest Inventory* (SII) and the *Kuder Career Search* (KCS). This category also includes numerous measures of attitudes toward topics, groups, and practices. For example, there are measures for attitude toward capital punishment, attitude toward the elderly, and so on.

Our final category includes **neuropsychological tests**. These are tests designed to yield information about the functioning of the central nervous system, especially the brain. From some perspectives, this should not be a separate category because many of the tests used for neuropsychological testing simply come from the other categories. Much neuropsychological testing employs ability tests and often uses personality tests, too. However, we use a separate category to capture tests used specifically to assess brain functions. Of special interest are tests of memory for verbal and figural material, psychomotor coordination, and abstract thinking.

**TRY IT!** .....

Here is a simple test used by neuropsychologists. It is called a Greek cross. Look at the figure for a moment. Then put it aside and try to draw it from memory. What behaviors and mental processes do you think are involved in this test?



## Some Additional Ways to Categorize Tests

Thus far, we have categorized tests according to their predominant type of content. In fact, this is the most common and, from most perspectives, the most useful way to classify tests. However, there are a number of other ways to classify tests. We will list them briefly. See Key Points Summary 1.3.

### Key Points Summary 1.3

#### *Some Additional Ways to Categorize Tests*

- Paper-and-Pencil versus Performance
- Speed versus Power
- Individual versus Group
- Maximum versus Typical Performance
- Norm-Referenced versus Criterion-Referenced



### Paper-and-Pencil versus Performance Tests

In a **performance test** the examinee completes some action such as assembling a product, delivering a speech, conducting an experiment, or leading a group. In a **paper-and-pencil test**, the examinee responds to a set of questions usually, as implied by the title, using paper and pencil. Many paper-and-pencil tests use multiple-choice, true-false, or similar item types. Traditional paper-and-pencil tests often appear now on a computer screen, with the answer marked by key stroke or mouse click.

### Speed versus Power Tests

The essential purpose of a **speed (or speeded) test** is to see how fast the examinee performs. The task is usually quite simple. The person's score is how many items or tasks can be completed in a fixed time, or how much time (e.g., in minutes or seconds) is required to complete the task. For example, how quickly can you cross out all the "e's" on this page? How quickly can you complete 50 simple arithmetic problems such as  $42 + 19$ ,  $24 \times 8$ , and so on? A **power test**, on the other hand, usually involves challenging material, administered with no time limit or a very generous limit. The essential point of the power test is to test the limits of a person's knowledge or ability (other than speed). The distinction is not necessarily all-or-none: pure speed versus pure power. Some power tests may have an element of speed. You can't take forever to complete the SAT. However, mental prowess and knowledge rather than speed are the primary determinants of performance on a power test. Some speed tests may have an element of power. You have to do some thinking and perhaps even have a plan to cross out all the "e's" on the page. However, crossing out "e's" is primarily a matter of speed, not rocket science.

### Individual versus Group Tests

This distinction refers simply to the mode of test administration. An **individual test** can be administered to only one individual at a time. The classic examples are individually administered intelligence tests. An examiner presents each question or task to the individual and records the person's response. A **group test** can be administered to many individuals at the same time, that is, to a group. Of course, individuals receive their own scores from a group-administered test. In general, any group-administered test can be administered to one individual at a time, when circumstances warrant, but individually administered tests cannot be given to an entire group at once.

### Maximum versus Typical Performance

Here is another useful distinction between types of tests. Some tests look for **maximum performance**. How well can examinees perform when at their best? This is usually the case with achievement and ability tests. On the other hand, we sometimes want to see a person's **typical performance**. This is usually the case with personality, interest, and attitude tests. For example, on a personality test we want to know how extroverted a person typically is, not how extroverted he can be if he is trying really hard to appear extroverted.