Assessing Adolescent and Adult Intelligence

Third Edition

Alan S. Kaufman

Elizabeth O. Lichtenberger

JOHN WILEY & SONS, INC
Assessing Adolescent and Adult Intelligence

Third Edition

Alan S. Kaufman

Elizabeth O. Lichtenberger

JOHN WILEY & SONS, INC
For Nadeen, For Always
La Bohème, Act IV

MIMI:
Sono andati? Fingevo di dormire
perchè volli con te sola restare.
Ho tante cose che ti voglio dire,
o una sola,
ma grande come il mare,
come il mare profonda ed infinita.
Sei il mio amor e tutta la mia vita.

Have they left us? I was pretending to sleep
to be alone with you.
So many things to tell you,
but really just one,
that is as huge as the ocean,
and as deep and infinite.
You are my love and my whole life.

RODOLFO:
O mia bella Mimi!

MIMI:
Son bella ancora?

RODOLFO:
Bella come un’aurora.

Beautiful as the dawn in Springtime.

A. S. K.

For Mike, for everything

For everyone on this earth,
there will always be
one special someone to love.
For me,
there will always be
You.

E. O. L.
Preface to the Third Edition xix

PART I
Introduction to the Assessment of Adolescent and Adult Intelligence

CHAPTER 1 IQ Tests: Their History, Use, Validity, and Intelligent Interpretation 1

Outline of the Book 2
Wechsler’s Scales 3
Clinical Relevance of Theory 3
A Short History of IQ Tests 3
The Binet-Simon Scales 4
Terman’s Stanford-Binet 4
The World War I Tests 5
Wechsler’s Creativity 6

Surveys of Test Usage for Adults 7
Has Test Use Changed over the Years? 8
Test Usage of 1,500 Psychologists and Neuropsychologists 8
How Frequently Are Tests Used? 9
Administration Time and Implications for Reimbursement 10
For What Purposes Are Adults Given Intelligence Tests? 10
Conclusions 10

Validity of the IQ Construct for Adolescents and Adults 11
Prediction of Academic Achievement 12
Relationship of IQ to Education 13
IQ and Occupation 14
Prediction of Job Performance 16

The Intelligent Testing Philosophy 18
IQ Tasks Measure What the Individual Has Learned 19
IQ Tasks Are Samples of Behavior and Are Not Exhaustive 19
CONTENTS

IQ Tasks Like the WAIS-III, KAIT, and Wj III Assess Mental Functioning under Fixed Experimental Conditions 20
IQ Tasks are Optimally Useful When They Are Interpreted from an Information-Processing Model 20
Hypotheses Generated from IQ Test Profiles Should Be Supported with Data from Multiple Sources 21
Tying Together the Tenets of Intelligent Testing 22
Summary 22

CHAPTER 2 Heritability and Malleability of IQ and Attacks on the IQ Construct 24
The Heritability and Malleability of IQ 24
Heritability 24
Maternal Environment 33
Overview 35
Malleability 36
The Flynn Effect 37
The Abecedarian Project 42
Overview of Malleability of IQ 45
Attacks on the IQ Construct 46
Lezak’s Eulogy 46
The Learning Disabilities Challenge 53
Summary 59

CHAPTER 3 From the Wechsler-Bellevue I to the WAIS-III 61
Selection of the Subtests 62
Verbal Scale 62
Performance Scale 66
Item Content Changes from the W-B I to the WAIS to the WAIS-R and to the WAIS-III 69
Reliability Comparisons of the W-B I, WAIS, WAIS-R, and WAIS-III 72
Standardization of the W-B I, WAIS, WAIS-R, and WAIS-III 74
Comparison of the Construct Validity of the W-B I, WAIS, WAIS-R, and WAIS-III 76
Internal Consistency 76
Factor Analysis 78
Developmental Trends 80
Problems with Adolescent Data 82
Clinical Implications of the WAIS-III Norms at Ages 16–19 84
Overview of Developmental Trends 84
Correlations between the W-B I and WAIS, WAIS and WAIS-R, and the WAIS-R and WAIS-III 84
PART II
Individual Differences on Age, Socioeconomic Status, and Other Key Variables

CHAPTER 4 Individual Differences for Adolescents and Adults on Gender, Ethnicity, Urban–Rural Residence, and Socioeconomic Status 96

Gender Differences 96
Global Scales 96
Gender Differences on Separate Subtests 98
Clinical Implications of Gender Differences on Mental Tasks 100

Ethnic Differences in IQ 101
Differences between Caucasians and African Americans 101
Differences between Caucasians and Hispanics 106

Urban–Rural Residence Differences 109
Generational Changes in Urban–Rural Differences 109
Residence Differences on the WAIS-R Subtests 110

Occupational Differences 111
Best Estimate of IQ Differences for Adults in Different Occupations 111
IQ Variability within Occupational Groups 113
Occupational Status and Canonical Factors 115

Educational Attainment 115
CONTENTS

WAIS-III Mean Scores Earned by Adults Differing in Educational Attainment  115
WAIS-III Correlations with Educational Attainment  117
Relationship of Intelligence to Education on the Kaufman Tests  120
Clinical Implications of Educational Data  123
Summary  124

CHAPTER 5 Age and Intelligence across the Adult Life Span  127
Does IQ Decline with Advancing Age?
A Cross-Sectional Approach  128
Cross-Sectional Investigations of Wechsler’s Adult Scales  129
Cross-Sectional Investigations of Kaufman and Kaufman’s Adult Tests  150
Overview of Cross-Sectional Investigations  159
Cautions Associated with Cross-Sectional Investigations  161
Does IQ Decline with Advancing Age?
A Longitudinal Approach  163
Problems in Investigating Aging Longitudinally  163
Two Groundbreaking Longitudinal Investigations of IQ and Aging  167
Kaufman’s (1990, 2001) Longitudinal Investigations of Wechsler’s Adult Scales Using Independent Samples  171
Interpretation of the Aging Patterns for V-IQ versus P-IQ: Speed or Fluid/Crystallized Ability?  180
Age-Related Declines on Untimed Tasks such as Matrix Reasoning  181
Age-Related Changes on Wechsler’s Arithmetic, Picture Completion, and Picture Arrangement Subtests  181
Data from the Bonn Longitudinal Study  182
Speed versus Cognitive Ability on Digit Symbol  182
Creative Works  182
Integration of Speed, Fluid Ability, and Other Factors  183
Can Cognitive Decline Be Slowed Down?  187
Summary  189

PART III Integration and Application of WAIS-III Research

CHAPTER 6 Research on Administration, Scoring, and Relationships between Wechsler Scales  192
Administration and Scoring  192
Administration Time  192
Administration Errors  197
CONTENTS

The Nature and Frequency of Administration Errors 197
Scoring Errors 198
Implications for Administering and Scoring the WAIS-III 202

CHAPTER 7 Factor Analysis of the WAIS-III 218
The WAIS-III as a One-Factor Test 218
The WAIS-III as a Two-Factor Test 220
The WAIS-III as a Three-Factor Test 222
The WAIS-III as a Four-Factor Test 223
How Many Factors Underlie the WAIS-III? 224
Are There Two, Three, or Four
WAIS-III Factors? 225

Differences in WAIS-III Factor Structure Due
to Ethnicity 227
Age Differences in the WAIS-III Factor Structure 227

Comparison of WAIS-III and WAIS-R Factors 232
Comparison of WAIS-III and WISC-III Factors 232
The General Factor (g) 234
Subtest Specificity 235
Definition of Subtest Specificity 236
Subtest Specificity of WAIS-III Subtests 236

Evaluation of the WAIS-III 237
Kaufman and Lichtenberger (1999) 242
Sattler and Ryan (1999) 242
Summary 243

CHAPTER 8 Verbal–Performance IQ Discrepancies:
A Neuropsychological Approach 244
V–P Discrepancies and Brain Damage 244

Right- versus Left-Brain Lesions 245
Wechsler-Bellevue Studies 251
WAIS Studies 251
Overlapping Samples of Patients 252
WAIS-R Studies 252
Comparison of Results for Wechsler-Bellevue versus WAIS
versus WAIS-R 252
Todd, Coolidge, and Satz’s Large-Scale Study with the WAIS 253
## CONTENTS

*Smith's Large-Scale Investigations Using the Wechsler-Bellevue* 253  
Bilateral Brain Damage 254  
General Conclusions regarding Studies of Brain Damage 257  

The Nature of Brain Damage 258  
Lesions Caused by Stroke, Head Injury, and Tumors 259  
Temporal Lobe Epilepsy 263  
Temporal Lobe versus Parietal Lobe Lesions 266  
Frontal Lobe Lesions 266  
Frontal versus Posterior Lesions 266  

Acute versus Chronic Lesions 267  
Fitzhugh et al.'s Study 267  
Aggregated Data from Several Samples 268  
Correlational Studies 269  

Subtest Patterns for Left- and Right-Lesion Patients 269  
Mean Scaled Scores for Patients with Focal Brain Damage 269  
Discrimination between Right- and Left-Lesion Patients 270  
Hemispheric Functioning: Process versus Content 274  
Should V–P Discrepancies Be Used for Assessing Brain Damage? 275  

Gender of Patients with Lateralized Lesions 276  
Lansdell's Initial Observations 276  
McGlone's Research 277  
Well-Controlled Investigations 278  
Inglis and Lawson's Study of Gender Differences 278  
Turkheimer et al.'s Study of Gender Differences 278  
Inglis and Lawson's Meta-Analysis of Gender Difference Studies 279  
Bornstein and Matarazzo's Review of Gender Difference Studies 279  
Snow, Freedman, and Ford's Review of Gender Differences 280  
Turkheimer and Farace's Meta-Analysis of Gender Difference Studies 281  
Research Issues in the Investigation of Gender Differences 283  
Aggregated Data from Gender Difference Studies 283  
Proposed Explanations of the Interaction with Gender 284  
Summary of Evaluation of Hypotheses for Gender Differences in Lesioned Patients 292  

Ethnicity of Patients with Lateralized Lesions 293  
Munder's Careful Investigation of V–P Differences for Caucasians and African Americans 293  
Subtest Patterns for Brain-Damaged Males by Race 294  

Age of Patient Samples 296  
Brain Damage in Children and in Adults 296  
Empirical Analysis with Samples of Adolescents and Adults 297  

Educational Level of Patient Samples 297
CONTENTS

Research Implications of Base-Rate Education Data 299
Education Level and V–P for Brain-Damaged Samples 300
Education as a Variable for Normal versus Brain-Damaged People 300
Clinical Issues in the Interpretation of a Patient’s V–P Difference 302
Individual versus Group Data 303
Emotional Disturbance and Diffuse Brain Injury 304
Emotional Disturbance and Lateralized Lesions 304
V–P Interpretation in the Context of Patients’ Behaviors 307
Illustrative Case Report 308
Walt H., Age 21, Head Injury 308
Summary 314

CHAPTER 9 Verbal–Performance IQ Discrepancies: A Clinical Approach 317
Variables Believed to Be Correlates of High Performance IQ 317
Learning Disabilities 318
Delinquency and Psychopathic Behavior 327
Bilingualism 330
Autism 332
Mental Retardation 334
Overview of Variables Believed to Be Correlates of High Performance IQ 337
Variables Believed to Be Correlates of High Verbal IQ 337
Psychiatric Disorders 338
Alcohol Abuse 340
Motor Coordination Problems 341
Alzheimer’s-Type Dementia 344
High Average and Gifted Intellectual Functioning (Full Scale IQ of 110+) 350
Overview of Variables Believed to Be Correlates of High Verbal IQ 353
When V–P IQ Discrepancies Are Meaningless 354
IQs Are Not Unitary Constructs 354
Verbal Compensation for Performance Deficit 355
Effects of Retesting 356
Illustrative Case Reports 357
Chester P., Age 17, Possible Autism 357
Robert N., Age 36, Mild Retardation 362
Summary 366
PART IV
Interpretation of the WAIS-III Profile: IQs, Factor Indexes, and Subtest Scaled Scores

CHAPTER 10  Profile Interpretation: What the Subtests Measure  369

Ways of Grouping WAIS-III Subtests  369

Factor Analysis  370
Bannatyne’s Categories  370
Horn’s Modified Fluid–Crystallized Model  371
Cattell-Horn-Carroll (CHC) Theory  372
Baltes’s Life Span Two-Component Model  373
Osgood’s Psycholinguistic Approach and the Information-Processing Approach  374
Rapaport’s Pioneering Clinical Model  376
Dean’s Individual Ability Profile  377
Guilford’s Structure-of-Intellect Model  380

Abilities Measured by the 14 WAIS-III Subtests  382

Sources and Methods for Analyzing Each Subtest  382

Vocabulary  384
Similarities  385
Arithmetic  387
Digit Span  389
Information  391
Comprehension  393
Letter-Number Sequencing  395
Picture Completion  397
Digit Symbol-Coding  399
Block Design  401
Matrix Reasoning  403
Picture Arrangement  405
Symbol Search  407
Object Assembly  408

Summary  410

CHAPTER 11  WAIS-III Profile Interpretation: Steps 1–7  412

Considerations about Profile Interpretation  412
Considerations for Applying the Interpretive Steps when Only 11 Subtests Were Administered  413

Step 1: Interpret the Full Scale IQ  414

The Role of Full Scale IQ in Profile Interpretation  415

Step 2: Are the Verbal IQ versus the Performance IQ (or the Verbal Comprehension Index versus the Perceptual Organization Index) Significantly Different?  415
CONTENTS

Step 3: Are the Verbal IQ versus the Performance IQ (or the Verbal Comprehension Index versus the Perceptual Organization Index) Differences Abnormally Large? 417
Step 4: Is the V-IQ versus P-IQ Discrepancy Interpretable? 421
Step 5: Is VCI versus POI Difference Interpretable? 426
Step 6: Determine whether the Working Memory and Processing Speed Indexes Are Interpretable 428
Step 7: Interpret the Global Verbal and Nonverbal Dimensions, as well as the Small Factors, if They Were Found to Be Interpretable 428

General Interpretation of WAIS-III Indexes 430

Verbal Comprehension Index 430
Perceptual Organization Index 430
Working Memory Factor 431
Processing Speed Factor 433

Horn's and Bannatyne's Systems for Interpreting Global Verbal, Nonverbal, Working Memory, and Processing Speed Dimensions 434

Assessment of Verbal Skills by the Two Approaches 435
Assessment of Nonverbal Skills by the Two Approaches 435
Assessment of Additional Skills by the Two Approaches 435
Computing Standard Scores for the Bannatyne Categories and Horn Groupings 436
Reliability of Bannatyne Categories and Horn Groupings 437
Determining Strengths and Weaknesses within Each System 437
Strengths and Weaknesses within a Cluster 440
Selecting a System 443

Characteristic Profiles in Unique Populations 445

Summary 445

CHAPTER 12 WAIS-III Profile Interpretation: Steps 8 and 9 447

Generating WAIS-III Hypotheses 447
The Nine-Step Process 448
Step 8: Interpret Significant Strengths and Weaknesses of the Subtest Profile 448
Illustration of Step 8: Ryan N. 451
Step 9: Generating Hypotheses about the Fluctuations in the WAIS-III Profile 451
Basic Tenets of the Philosophy of Hypothesis Generation 453
Introduction to WAIS-III Subtest Interpretive Tables 455
Reliability Coefficients of Shared Abilities 455
CONTENTS

How to Use Information about Shared Abilities  455
  Guideline 1  459
  Guideline 2  459
  Guideline 3  459
  Guideline 4  476
  Guideline 5  479
Summary of Sequential Approach to WAIS-III Interpretation  479
A Simultaneous Approach to Hypothesis Generation  479
Common Configurations of the WAIS-III Verbal Scale  481
Common Configurations of the WAIS-III Performance Scale  486
Configurations of Both Scales  492
Illustrative Case Reports  496
  Nicole H., Age 34, Learning Problem  496
  Aimée L., Age 26, Memory Concerns  501
  Lauren J., Age 57, Decline in Job Performance  507
Summary  515
WAIS–III Interpretive Worksheet  517

PART V
Additional Measures of Adolescent and Adult IQ

CHAPTER 13 Kaufman Adolescent and Adult Intelligence Test (KAIT)  522
KAIT Theory  522
  Horn-Cattell Theory in the KAIT  522
  The Theories of Luria and Piaget in the KAIT  523
Composition of the KAIT  523
Other Tests and the KAIT  524
  KAIT and the Wechsler Scales  524
  KAIT and WJ-III  525
Types of Scores on the KAIT  525
KAIT’s Standardization and Psychometric Properties  525
KAIT Interpretation  527
  Step 1: Interpret the KAIT Composite IQ  527
  Step 2: Examine the KAIT Crystallized IQ versus the Fluid IQ  528
  Step 3: Is the KAIT Crystallized IQ versus Fluid IQ Difference Abnormal?  529
  Step 4: Determine if the Crystallized and Fluid IQs Are Interpretable  529
  Step 5: Interpret the Meaning of the Crystallized and Fluid Scales  530
  Step 6: Interpret KAIT Strengths and Weaknesses  531
CONTENTS

Step 7: Generate Hypotheses about Fluctuations in the KAIT Profile 533
Alternative KAIT Subtest Groupings 539
KAIT Memory Comparisons 540
KAIT Research on Clinical Profiles 541
Assessment of Learning Disabilities 541
Assessment of Alzheimer’s-Type Dementia 543
Depression and Pseudodementia 545
Personality and Interests as They Pertain to Career Choices 547
KAIT’s Construct Validity for African Americans and Hispanics 549
KAIT Case Study 550
Jeff H., Age 17, Possible Learning Disability 550
Summary 560

CHAPTER 14 Woodcock-Johnson Battery—Third Edition (WJ III) 561
Kevin S. McGrew, Richard Woodcock, and Laurie Ford
History and Evolution of the WJ III 561
Original Woodcock-Johnson (WJ) 561
Woodcock-Johnson—Revised (WJ-R) 562
Theoretical Foundations of the WJ III 563
The Cattell-Horn-Carroll (CHC) Theory of Cognitive Abilities 563
Use of the CHC Theory as the WJ III Design Blueprint 566
Description and Organization of the WJ III Tests and Clusters 570
Organization 570
The Cognitive Battery 571
The Achievement Battery 572
Interpretive Features of the WJ III 577
Types of Scores 578
Types of Profiles 579
Clinical and Selected Special Purpose Clusters 579
The Cognitive and Achievement Performance Models 581
Psychometric Characteristics 582
Norms 582
Reliability 583
Validity 584
Special Applications and Use with Special Populations 600
Assessment of Learning Disabilities 600
General Intellectual Ability/Achievement Discrepancy Model 600
Predicted Achievement–Achievement Discrepancy Model 602
Intra-Ability Discrepancy Model 603
Preliminary Research on Adult LD Identification with the WJ III 604
Neuropsychological Applications 607
CONTENTS

CHC Abilities across the Life Span 613
CHC Cross-Battery Applications 624
Summary 627

CHAPTER 15 Brief Tests of Intelligence and Related Abilities 629
Brief History of Brief Assessment 629
The Terman–Merrill Innovation 629
The Role of the Slosson, Shipley-Hartford, and Other Early Brief Tests 630
Silverstein’s Contributions to Wechsler Short Forms 631
Trends in Short-Form Development for WAIS and Its Successors 632
The Future of Short Forms 633
Let’s Stop Developing and Using Short Forms of Long Tests 633
The Normative Issue, Part I: Thompson’s Key Research 634
The Normative Issue, Part II: The Satz–Mogel Approach 635
The Statistical Issue: Correlated Error Variance 637
 Appropriateness of the New Brief Tests for African Americans and Hispanics 638
The Seven-Subtest WAIS-III 640
When to Administer Brief Tests 640
Three Well-Normed Recent Brief Tests of Verbal and Nonverbal Intelligence 641
Wechsler Abbreviated Scale of Intelligence (WASI) 641
Kaufman Brief Intelligence Test (K-BIT) 645
Wide Range Intelligence Test (WRIT) 647
Brief Tests of Either Nonverbal or Verbal Ability 650
Peabody Picture Vocabulary Test—Third Edition (PPVT-III) 650
General Abilities Measure for Adults (GAMA) 652
Matrix Analogies Test (MAT) 654
Brief Tests for Specialized Abilities 655
Kaufman Short Neuropsychological Assessment Procedure (K-SNAP) 655
Kaufman Functional Academic Skills Test (K-FAST) 657
Summary 660

APPENDIX A Alternative Approach to Interpreting the WAIS-III: Applying the Theory-Based Flanagan-Kaufman Interpretive Method for the WISC-IV 663

APPENDIX B WAIS-III Interpretive Worksheet 690
CONTENTS

APPENDIX C  Norm Tables for Computing Standard Scores on the General Ability Index (GAI) and the Clinical Clusters  698

APPENDIX D  Watkins and Canivez's Critique of the Kaufman-Lichtenberger Interpretive System and Articulation of a New Theory-Based Approach to Profile Interpretation  708

References  718

Name Index  776

Subject Index  787
The Second Edition of Assessing Adolescent and Adult Intelligence, published in 2002, was a thorough revision of the original 1990 text. A comprehensive revision of the 1990 text was necessary because the Wechsler scale featured in the First Edition (the WAIS-R) was replaced on the clinical scene by the WAIS-III, and because much pertinent research had been conducted in the 12 years that elapsed between the two editions. In contrast, this Third Edition, published only four years after the previous edition, represents only a slight revision of the text. The WAIS-III is still the most recent version of Wechsler’s adult scales and continues to be featured in this edition. Indeed, the 15 chapters that composed the Second Edition are retained intact in the Third Edition. However, four Appendixes have been added to the book, all of which concern the interpretation of the WAIS-III.

Flanagan and Kaufman (2004) developed a new theory-based approach to the interpretation of the WISC-IV that they featured in their book Essentials of WISC-IV Assessment. Several of our colleagues inquired whether this new method of profile interpretation, developed from Cattell-Horn-Carroll (CHC) theory, could be applied to the WAIS-III. The answer was “Yes,” but a translation of the new system to the WAIS-III required reconceptualization (because the WISC-IV and WAIS-III contain different sets of subtests), review of the recent WAIS-III literature on test interpretation, and the development of new interpretive and norms tables. We decided to do the work necessary to translate the WISC-IV system to the WAIS-III, and to append the new interpretive system to the existing set of chapters.

Appendixes A, B, C, and D are included in the Third Edition to provide examiners with an alternative system for interpreting the WAIS-III—specifically an analog of the CHC approach that Flanagan and Kaufman developed for the WISC-IV. The “sequential” and “simultaneous” interpretive approaches that we developed for the WAIS-III in the Second Edition of this book are still featured in Chapters 11 and 12 and remain viable systems for interpreting the WAIS-III. However, the new theory-based method that is presented in depth in the Appendixes provides a useful alternative approach, one that examiners may use instead of—or as a supplement to—the more traditional methods described in Chapters 11 and 12. The Appendixes in the Third Edition of this book incorporate important research conducted by Flanagan and Kaufman, Tulsky and his colleagues, and Longman, and deals with a vital aspect of Wechsler assessment—namely theory-based profile interpretation. The Third Edition, therefore, reflects an important enhancement of the previous edition of this text.
ACKNOWLEDGMENTS

We are grateful to Dr. Dawn Flanagan for her important contributions to the Appendixes, to Dr. Jack Naglieri for conducting important WAIS-III data analyses, to Dr. Laney Margolis for urging us to apply the new WISC-IV interpretive system to the WAIS-III, to Dr. James Kaufman for his research assistance, and to Ms. Tracey Belmont for her important role in arranging for John Wiley to publish the Third Edition of Assessing Adolescent and Adult Intelligence.

ALAN S. KAUFMAN
ELIZABETH O. LICHTENBERGER
IQ Tests: Their History, Use, Validity, and Intelligent Interpretation

The field of intelligence, particularly of adolescent and adult mental development, has dominated the psychological literature for decades, and now encompasses a diversity of domains within cognitive psychology, clinical psychology, psychobiology, behavioral genetics, education, school psychology, sociology, neuropsychology, and everyday life. Excellent handbooks are available with chapters written by experts in many aspects of intellectual theory, measurement, and development (e.g., Flanagan, Genshaft, & Harrison, 1997; Groth-Marnat, 2000), and even these texts cover only a portion of the territory and quickly become outdated. Consequently, in writing this text on the assessment of adolescent and adult intelligence, we have had to make several decisions about which areas to include and how thoroughly to cover each topic.

First, this book focuses on the clinical assessment of intelligence, and every topic must bear, either directly or indirectly, on the clinical aspect of mental measurement. Because clinical assessment within the fields of neuropsychology, special education, and clinical, school, and counseling psychology involves individual evaluations, research on group-administered tests is subordinated to the more pertinent research on individual intelligence tests. The 1990 version of this text covered group-administered intelligence tests to some extent. However, the adolescent and adult assessment scene has changed during this past decade, with clinicians having options beyond Wechsler’s tests. Whereas the Wechsler Adult Intelligence Scale—Third Edition (WAIS-III; Psychological Corporation, 1997; Wechsler, 1997) is still the most used test, and is clearly the featured instrument in this revised text, the availability of a variety of new in-depth and brief intelligence tests, and a proliferation of research on these instruments, has impelled us to focus on individually administered intelligence tests.

For example, the monumental efforts of Schaie (1958, 1983b, 1994) and his colleagues (Hertzog & Schaie, 1988; Schaie & Labouvie-Vief, 1974; Schaie & Strother, 1968; Schaie & Willis, 1993) to understand the development of
adult intelligence have been based on the group-administered Primary Mental Abilities Test. The key findings from these innovative cross-sequential studies are of interest to psychology in general, but have limited applicability to the work of clinical and neuropsychological practitioners. Consequently, investigations by Schaie will only be discussed in the context of aging studies on clinical instruments (e.g., Kaufman, 2000b, 2001; Kaufman & Horn, 1996), especially the WAIS-III, WAIS-R (Wechsler, 1981) and Kaufman Adolescent and Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993).

Consistent with the focus on clinical tests of intelligence, we have also eliminated sections and chapters from the first edition on clinical tools that are only tangentially related to IQ assessment, most notably neuropsychological instruments, adaptive behavior surveys, and individual achievement tests.

**OUTLINE OF THE BOOK**

*Assessing Adolescent and Adult Intelligence* (2nd ed.) has five parts:

I. Introduction to the Assessment of Adolescent and Adult Intelligence (Chapters 1–3)

II. Individual Differences on Age, Socioeconomic Status, and Other Key Variables (Chapters 4–5)

III. Integration and Application of WAIS-III Research (Chapters 6–9)

IV. Interpretation of the WAIS-III Profile: IQs, Factor Indexes, and Subtest Scaled Scores (Chapters 10–12)

V. Additional Measures of Adolescent and Adult IQ (Chapters 13–15)

Part I includes: Chapter 1, which discusses pertinent historical information, issues regarding validation of the IQ construct, and our philosophy of intelligent testing; Chapter 2, which discusses pressing issues and challenges to the IQ concept (e.g., heritability and malleability of the IQ); and Chapter 3, which provides the rationale for the WAIS-III subtests for adolescents and adults and traces the empirical and logical continuity from the Wechsler-Bellevue to the WAIS to the WAIS-R and to the WAIS-III.

Part II presents research on individual differences in intelligence associated with pertinent background variables on the WAIS-III and other instruments, notably gender, ethnicity, socioeconomic status, and urban–rural residence (all treated in Chapter 4), and aging across the adult lifespan (Chapter 5).

Parts III and IV are devoted to the WAIS-III and, occasionally, its predecessors (e.g., WAIS-R) or “alternate-form” at age 16 (WISC-III). In Part III, the focus is on research, delving into topics such as administration and scoring (Chapter 6), factor analysis (Chapter 7), and Verbal Performance (V-P) IQ differences, especially as they pertain to lateralized brain lesions (Chapter 8) and other clinical disorders (Chapter 9). The three chapters of Part IV (Chapters 10, 11, and 12) are all devoted to an empirical and clinical approach to interpretation of the WAIS-III multiscore profile.

Part V is composed of three chapters; each focuses exclusively on additional (non-Wechsler) measures for adolescent and adult assessment and integrates them with the WAIS-III: the KAIT (Chapter 13), the Woodcock-Johnson—Third Edition or WJ III (Chapter 14, authored by McGrew, Woodcock, and Ford), and a variety of brief tests of intelligence (Chapter 15). The tests discussed in the latter chapter, for example, the Peabody Picture Vocabulary Test—Third Edition (PPVT-III), the Kaufman Brief Intelligence Test (K-BIT), and the Wechsler Abbreviated Scale of Intelligence (WASI), may be used as supplements to the WAIS-III, KAIT, or WJ III, or may be used instead of comprehensive intelligence tests in certain circumstances (e.g., screening or research purposes).

The discussion of non-Wechsler tests in Part IV is essential to round out the cognitive assessment scene, but the WAIS-III, like the WAIS-R, WAIS, and Wechsler-Bellevue before it, remains the key tool for clinical and neuropsychological evaluation of adolescents and adults and, hence,
the focus of all sections of the book. The chapters on clinical applications of intelligence tests, along with the previous parts of the book, place the focus of this text squarely on the WAIS-III.

**Wechsler’s Scales**

Even a casual observer of the clinical or neuropsychological assessment scene is aware that Wechsler’s scales are uncontested as the primary cognitive measures of adolescent and adult intelligence. Individuals in their teens and adults of all ages are invariably administered the Wechsler Intelligence Scale for Children—Third Edition (WISC-III; Wechsler, 1991) or the WAIS-III when they are referred to a competent professional for a thorough assessment of their intellectual abilities, usually as part of a clinical, vocational, neuropsychological, or psychoeducational evaluation. The WISC-III is used for adolescents as old as 16 years, while the WAIS-III is used for individuals aged 16 to 89. Therefore, they overlap at age 16, giving clinicians a choice of Wechsler test for that age group.

Using the WISC-III as a clinical and psychometric tool has been discussed elsewhere in a comprehensive text (Kaufman, 1994a). For practical purposes, then, this book is primarily devoted to the WAIS-III, child of the WAIS-R (Wechsler, 1981), grandchild of the WAIS (Wechsler, 1955), and great-grandchild of the Wechsler-Bellevue Form I (Wechsler, 1939).

**Clinical Relevance of Theory**

To be included in this book in any depth, a topic needs to contribute to a psychologist’s understanding of intelligence in the clinical arena, not in the laboratory. For example, the Cattell-Horn-Carroll (CHC; McGrew & Flanagan, 1998) theory—an amalgam of Horn’s (1989) expansion of Horn-Cattell Gf-Gc theory and Carroll’s (1993, 1997) model of intelligence—is treated throughout the book because it is instrumental in explaining changes in verbal and nonverbal abilities with advancing age, and it (or Horn-Cattell theory) underlies three tests of adolescent and adult intelligence: the Woodcock Johnson Psycho-Educational Battery—Third Edition (WJ III; Woodcock, McGrew, & Mather, 2000), the Stanford-Binet Intelligence Scale, Form IV (Thorndike, Hagen, & Sattler, 1986a), and the KAIT (Kaufman & Kaufman, 1993). In contrast, Sternberg’s (1985) three-pronged triarchic theory of intelligence, though popular and widely discussed, is not emphasized because of its limited application to clinical assessment and the interpretation of the WAIS-III and other individual intelligence tests. Currently the Sternberg Triarchic Abilities Test (Sternberg, 1993), a group-administered measure, is available as an unpublished research instrument available from its author. However, if it is ever adapted as an individually administered, commercially published, standardized measure that translates laboratory principles to the domain of the clinical psychologist, neuroclinician, and psychoeducational diagnostician, the theory may become even more popular.

In addition, other theories of intelligence such as Gardner’s (1993a, 1993b) theory of multiple intelligences—which defines intelligence as the ability to solve problems, or to create products, that are valued within one or more cultural settings—is also not emphasized in this book. The theory of multiple intelligences calls for measuring intelligences by asking individuals to solve problems in the contexts in which they naturally occur. Although the multiple intelligences theory has attracted much attention in the fields of cognition and education (Kornhaber & Kreechesky, 1995), thus far its practical application to clinical assessment and the interpretation of the WAIS-III and other major standardized individual intelligence tests is limited.

**A Short History of IQ Tests**

The history of intellectual assessment is largely a history of the measurement of the intelligence of children or retarded adults. Sir Francis Galton (1869, 1883) studied adults and was interested in
giftedness when he developed what is often considered the first comprehensive individual test of intelligence (Kaufman, 2000a). But despite Galton’s role as the father of the testing movement (Shouksmith, 1970), he did not succeed in constructing a true intelligence test. His measures of simple reaction time, strength of squeeze, or keenness of sight proved to assess sensory and motor abilities, skills that relate poorly to mental ability, and that are far removed from the type of tasks that constitute contemporary intelligence tests.

The Binet-Simon Scales

Alfred Binet and his colleagues (Binet & Henri, 1895; Binet & Simon, 1905, 1908) developed the tasks that survive to the present day in most tests of intelligence for children and adults. Binet (1890a, 1890b) mainly studied children; beginning with systematic developmental observations of his two young daughters, Madeleine and Alice, he concluded that simple tasks like those used by Galton did not discriminate between children and adults. In 1904, the Minister of Public Instruction in Paris appointed Binet to a committee to find a way to distinguish normal from retarded children. But 15 years of qualitative and quantitative investigation of individual differences in children—along with considerable theorizing about mental organization and the development of a specific set of complex, high-level tests to investigate these differences—preceded the “sudden” emergence of the landmark 1905 Binet-Simon intelligence scale (Murphy, 1968).

The 1908 scale was the first to include age levels, spanning the range from III to XIII. This important modification stemmed from Binet and Simon’s unexpected discovery that their 1905 scale was useful for much more than classifying a child at one of the three levels of retardation: moron, imbecile, idiot (Matarazzo, 1972). Assessment of older adolescents and adults, however, was not built into the Binet-Simon system until the 1911 revision. That scale was extended to age level XV and included five ungraded adult tests (Kite, 1916). This extension was not conducted with the rigor that characterized the construction of tests for children, and the primary applications of the scale were for use with school-age children (Binet, 1911).

Measuring the intelligence of adults, except those known to be mentally retarded, was almost an afterthought. But the increased applicability of the Binet-Simon tests for various child-assessment purposes dawned on Binet just prior to his untimely death in 1911: “By 1911 Binet began to foresee numerous uses for his method in child development, in education, in medicine, and in longitudinal studies predicting different occupational histories for children of different intellectual potential” (Matarazzo, 1972, p. 42).

Terman’s Stanford-Binet

Lewis Terman was one of several people in the United States who translated and adapted the Binet-Simon scale for use in the United States, publishing a “tentative” revision (Terman & Childs, 1912) 4 years before releasing his painstakingly developed and carefully standardized Stanford Revision and Extension of the Binet-Simon Intelligence Scale (Terman, 1916). This landmark test, soon known simply as the Stanford-Binet, squashed competing tests developed earlier by Goddard, Kuhlmann, Wallin, and Yerkes. Terman’s success was undoubtedly due in part to heeding the advice of practitioners whose demand “for more and more accurate diagnoses ... raised the whole question of the accurate placing of tests in the scale and the accurate evaluation of the responses made by the child” (Pintner & Patterson, 1925, p. 11).

But, like Binet, Terman (1916) saw intelligence tests useful primarily for the detection of mental deficiency or superiority in children and for the identification of “feeblemindedness” in adults. He cited numerous studies of delinquent adolescents and adult criminals, all of which pointed to the high percentage of mentally deficient juvenile delinquents, prisoners, or prostitutes, and concluded that “there is no investigator who denies
the fearful role played by mental deficiency in the 
production of vice, crime, and delinquency” (p. 9).
Terman also saw the potential for using intelli-
gence tests with adults for determining “vocational 
fitness,” but, again, he emphasized employing “a 
psychologist…to weed out the unfit” or to “deter-
mine the minimum ‘intelligence quotient’ neces-
sary for success in each leading occupation” 
(p. 17).
Perhaps because of this emphasis on the as-
essment of children or concern with the lower 
end of the intelligence distribution, Terman 
(1916) did not use a rigorous methodology for 
constructing his adult-level tasks. Tests below 
the 14-year level were administered to a fairly 
representative sample of about 1,000 children 
and early adolescents. To extend the scale above 
that level, data were obtained from 30 business-
men, 50 high school students, 150 adolescent delinquents, and 150 migrating unemployed 
men. Based on a frequency distribution of the 
mental ages of a mere 62 adults (the 30 business-
men and 32 of the high school students above 
age 16), Terman partitioned the graph into the 
following MA categories: 13–15 (inferior adults), 
15–17 (average adults), and above 17 (superior 
adults).

**The World War I Tests**

The infant field of adult assessment grew rapidly 
with the onset of World War I, particularly after 
U.S. entry into the war in 1917 (Anastasi & Ur-
bina, 1997; Vane & Motta, 1984). Psychologists 
saw with increasing clarity the applications of in-
telligence tests for selecting officers and placing 
enlisted men in different types of service, apart 
from their generation-old use for identifying the 
mentally unfit. Under the leadership of Robert 
Yerkes and the American Psychological Associa-
tion, the most innovative psychologists of the 
day helped translate Binet’s tests to a group for-
mat. Arthur Otis, Terman’s student, was instru-
mental in leading the creative team that 
developed the Army Alpha, essentially a group-
administered Stanford-Binet, and the Army 
Beta, a novel group test composed of nonverbal 
tasks.

Yerkes (1917) opposed Binet’s age-scale ap-
proach and favored a point-scale methodology, 
one that advocates selection of tests of specified, 
important functions rather than a set of tasks 
that fluctuates greatly with age level and devel-
opmental stage. The Army group tests reflect a 
blend of Yerkes’s point-scale approach and Bi-
et’s notions of the kind of skills that should be 
measured when assessing mental ability. The 
Army Alpha included the Binet-like tests of 
Directions or Commands, Practical Judgment, 
Arithmetical Problems, Synonym-Antonym, Dis-
sarranged Sentences, Analogies, and Information. 
Even the Army Beta had subtests resembling 
Stanford-Binet tasks: Maze, Cube Analysis, Picto-
rial Completion, and Geometrical Construction. 
The Beta also included novel measures like Digit 
Symbol, Number Checking, and X-O Series 
(Yoakum & Yerkes, 1920).

Never before or since have tests been normed 
and validated on samples so large; 1,726,966 men 
were tested (Vane & Motta, 1984)!

Point-scores on the Army Alpha or Army Beta were converted 
to letter grades ranging from A to D- (the Beta 
was given only to illiterate and non-English-
speaking candidates). Validity was demonstrated 
by examining the percent of As obtained by a va-
riety of Army ranks, for example, recruits (7.4%), 
corporals (16.1%), sergeants (24.0%), and majors 
(64.4%). In perhaps the first empirical demon-
stration of the Peter Principle in action, second 
lieutenants (59.4% As) outperformed their direct 
superiors—first lieutenants (51.7%) and captains 
(53.4%)—while those with ranks above major 
performed slightly worse than majors (Yoakum 
& Yerkes, 1920, Table 1). Can there be any more 
compelling affirmation of the validity of the Army 
intelligence tests? Another intelligence scale was 
developed during the war, one that became an 
alternative for those who could not be tested val-
idly by either the Alpha or Beta. This was the 
Army Performance Scale Examination, composed 
of tasks that would become the tools-of-trade for 
clinical psychologists, school psychologists, and
neuropsychologists into the twenty-first century: Picture Completion, Picture Arrangement, Digit Symbol, and Manikin and Feature Profile (Object Assembly). Except for Block Design (developed by Kohs in 1923), Wechsler’s influential Performance Scale was added to the Army battery, “[t]o prove conclusively that a man was weakminded and not merely indifferent or malingering” (Yoakum & Yerkes, 1920, p. 10).

Wechsler’s Creativity

David Wechsler assembled a test battery in the mid-1930s that comprised subtests developed primarily by Binet and World War I psychologists. His Verbal Scale was essentially a Yerkes point-scale adaptation of Stanford-Binet tasks; his Performance Scale, like other similar nonverbal batteries of the 1920s and 1930s (Cornell & Coxe, 1934; Pintner & Patterson, 1925), was a near replica of the tasks and items making up the individually administered Army Performance Scale Examination.

In essence, Wechsler took advantage of tasks developed by others for nonclinical purposes to develop a clinical test battery. He paired verbal tests that were fine-tuned to discriminate among children of different ages with nonverbal tests that were created for adult males who had flunked both the Alpha and Beta exams—nonverbal tests that were intended to distinguish between the nonmotivated and the hopelessly deficient. Like Terman, Wechsler had the same access to the available tests as did other psychologists; like Terman and Binet before him, Wechsler succeeded because he was a visionary, a man able to anticipate the needs of practitioners in the field.

While others hoped intelligence tests would be psychometric tools to subdivide retarded individuals into whatever number of categories was currently in vogue, Wechsler saw the tests as dynamic clinical instruments. While others looked concretely at intelligence tests as predictors of school success or guides to occupational choice, Wechsler looked abstractly at the tests as a mirror to the hidden personality. With the Great War over, many psychologists returned to a focus on IQ testing as a means of childhood assessment; Wechsler (1939), however, developed the first form of the Wechsler-Bellevue exclusively for adolescents and adults.

Most psychologists saw little need for nonverbal tests when assessing English-speaking individuals other than illiterates. How could it be worth 2 or 3 minutes to administer a single puzzle or block-design item when 10 or 15 verbal items can be given in the same time? Some test developers (e.g., Cornell & Coxe, 1934) felt that Performance scales might be useful for normal, English-speaking people to provide “more varied situations than are provided by verbal tests” (p. 9), and to “test the hypothesis that there is a group factor underlying general concrete ability, which is of importance in the concept of general intelligence” (p. 10).

Wechsler was less inclined to wait a generation for data to accumulate. He followed his clinical instincts and not only advocated the administration of a standard battery of nonverbal tests to everyone but placed the Performance Scale on an equal footing with the more respected Verbal Scale. Both scales would constitute a complete Wechsler-Bellevue battery, and each would contribute equally to the overall intelligence score.

Wechsler also had the courage to challenge the Stanford-Binet monopoly, a boldness not unlike Binet’s when the French scientist created his own forum (the journal L’Année Psychologique) to challenge the preferred but simplistic Galton sensorimotor approach to intelligence (Kaufman, 2000a). Wechsler met the same type of resistance as Binet, who had had to wait until the French Ministry of Public Instruction “published” his Binet-Simon Scale. When Wechsler’s initial efforts to find a publisher for his two-pronged intelligence test met failure, he had no cabinet minister to turn to, so he took matters into his own hands. With a small team of colleagues, he standardized Form I of the Wechsler-Bellevue by himself. Realizing that stratification
on socioeconomic background was more crucial than obtaining regional representation, he managed to secure a well-stratified sample from Brooklyn, New York.

The Psychological Corporation agreed to publish Wechsler's battery once it had been standardized, and the rest is history. Although an alternative form of the Wechsler-Bellevue (Wechsler, 1946) was no more successful than Terman and Merrill's (1937) ill-fated Form M, a subsequent downward extension of Form II of the Wechsler-Bellevue (to cover the age range 5 to 15 instead of 10 to 59) produced the wildly successful WISC (Wechsler, 1949). Although the Wechsler scales did not initially surpass the Stanford-Binet in popularity, serving an apprenticeship to the master in the 1940s and 1950s, the WISC and the subsequent revision of the Wechsler-Bellevue, Form I (WAIS; Wechsler, 1955) triumphed in the 1960s. "With the increasing stress on the psychoeducational assessment of learning disabilities in the 1960s, and on neuropsychological evaluation in the 1970s, the Verbal-Performance (V-P) IQ discrepancies and subtest profiles yielded by Wechsler's scales were waiting and ready to overtake the one-score Binet" (Kaufman, 1983b, p. 107).

Irony runs throughout the history of testing. Galton developed statistics to study relationships between variables—statistics that proved to be forerunners of the coefficient of correlation, later perfected by his friend Karl Pearson (DuBois, 1970). The ultimate downfall of Galton's system of testing can be traced directly to coefficients of correlation, which were too low in some crucial (but, ironically, poorly designed) studies of the relationships among intellectual variables (Sharp, 1898–99; Wissler, 1901). Similarly, Terman succeeded with the Stanford-Binet while the Goddard-Binet (Goddard, 1911), the Herrings-Binet (Herring, 1922), and other Binet-Simon adaptations failed because he was sensitive to practitioners' needs. He patiently withheld a final version of his Stanford revision until he was certain that each task was appropriately placed at an age level consistent with the typical functioning of representative samples of U.S. children.

Terman continued his careful test development and standardization techniques with the first revised version of the Stanford-Binet (Terman & Merrill, 1937). But 4 years after his death in 1956, his legacy was devalued when the next revision of the Stanford-Binet comprised a merger of Forms L and M, without a standardization of the newly formed battery (Terman & Merrill, 1960). The following version saw a restandardization of the instrument, but without a revision of the placement of tasks at each age level (Terman & Merrill, 1973). Unfortunately for the Binet, the abilities of children and adolescents had changed fairly dramatically in the course of a generation, so the 5-year level of tasks (for example) was now passed by the average 4½-year-old!

Terman's methods had been ignored by his successors. The ironic outcome was that Wechsler's approach to assessment triumphed, at least in part because the editions of the Stanford-Binet in the 1960s and 1970s were beset by the same type of flaws as Terman's competitors in the 1910s. The newest Stanford-Binet (Thorndike, Hagen, & Sattler, 1986a, 1986b) attempted to correct these problems and even adopted Wechsler's multisubtest, multiscale format. However, these changes in the Fourth Edition of the Binet were too little and too late to be much threat to the popularity of the Wechsler scales, to offer much contribution to the field of intelligence testing, or to merit the linkage with the Binet tradition.

**Surveys of Test Usage for Adults**

Surveys of test use in the United States have appeared increasingly in the literature in the past decade. These surveys are usually based on data from clinical agencies and hospitals (Lubin, Larsen, & Matarazzo, 1984; Petrowski & Keller, 1989), school systems (Goh, Teslow, & Fuller,
PART I  INTRODUCTION TO THE ASSESSMENT OF ADOLESCENT AND ADULT INTELLIGENCE

1981; Hutton, Dubes, & Muir, 1992; Wilson & Reschly, 1996), industry (Swenson & Lindgren, 1952), military settings (Lubin, Larsen, Marrazzo, & Seever, 1986), forensic settings (Lees-Hayley, Smith, Williams, & Dunn, 1996), or private practitioners (Archer, Maruish, Imhof, & Piotrowski, 1991; Camara, Nathan, & Puente, 2000; Harrison et al., 1988; Lubin et al., 1986; Watkins, Campbell, Nieberding, & Hallmark, 1995). Data from such studies of test use are becoming increasingly important in light of the role that managed-care companies play in reimbursement for assessment services. Data from surveys that help determine which are the typical instruments used for various types of assessment and the amount of time practitioners usually spend on an assessment may serve a function in setting standard approved rates for practitioner compensation by managed-care companies. Thus, we reviewed the recent literature to attempt to discover which instruments are most commonly used by practitioners with a variety of backgrounds and find out how much time is typically spent on assessments.

Has Test Use Changed over the Years?

Overall, little substantive change has occurred in the most popular instruments used in the last several decades (Camara et al., 2000). Test usage was first documented by Louttit and Brown (1947), with data collected spanning the mid-1930s to the mid-1940s. Since that early survey, subsequent surveys have shown that the most commonly used tests have not changed much over the years. The Wechsler family of tests has remained on the top of the assessment list for most psychologists, across a variety of settings (Ball, Archer, & Imhof, 1994; Brown & McGuire, 1976; Camara et al., 2000; Harrison et al., 1988; Lubin et al., 1971). The WAIS and WAIS-R have consistently been mentioned in surveys as the most often used adult intelligence tests by clinical psychologists, school psychologists, neuropsychologists, and forensic psychologists, and the WAIS-III will surely follow suit in future surveys.

Many studies of test usage lump together tests from all areas of assessment, including intellectual assessment, personality assessment, adaptive functioning assessment, achievement assessment, and neuropsychological assessment. Nonetheless, even when considering all these different types of assessment, the Wechsler tests remain ranked in the top 10.

Because the WAIS-III is fairly new, we were unable to find any published surveys that reported on the latest adult Wechsler test. The most recent survey at the time that this book went to press had a 2000 publication date, but the authors collected their data in late 1994, before the WAIS-R was revised (Camara et al., 2000). However, it is safe to assume that the WAIS-III will maintain the high ranking enjoyed by the WAIS-R.

Test Usage of 1,500 Psychologists and Neuropsychologists

Camara et al.'s (2000) collected survey data on test usage and assessment from 933 clinical psychologists and 567 neuropsychologists who were randomly selected from the American Psychological Association (APA) and the National Association of Neuropsychology (NAN). The authors were interested in data from practitioners who conducted assessments on a regular basis, so they ultimately conducted their analyses on data from respondents who engaged in 5 or more hours per week of assessment-related services. Thus, the final sample used for ranking test usage comprised 179 clinical psychologists (19% of the clinical psychologist respondents) and 447 neuropsychologists (79% of the neuropsychologist respondents). Table 1.1 displays the hours spent administering, scoring, and interpreting psychological tests during a typical week, for the total number of respondents to the survey (N = 1,500).