INDEX

A
AASA (American Association of School Administrators), 51
ABCs of School Testing, The, 33
Academic aptitude testing, 381–384
group-administered, 384
history of, 381–382
individually administered, 419–420
IQ scores, stability of, 382
IQ tests, predictions of, 383–384
Accuracy:
defined, 293
error and, see Error; Standard error of measurement \(S_m\)
Achievement:
with effort, 218–219
with improvement, 219
motivation, 194
process of evaluating, 209–210
Achievement tests, 378–381
diagnostic, 381
test batteries, 379–380
Adequate Yearly Progress (AYP), 58
Adjective Checklist, 386
Administrating the test, 193–195
Administrative policy decisions, 62
Advantages, equalizing, 223–224
Affective domain, 97–99
classification level, 99
organization level, 98
performance objectives in, 162
receiving (attending) level, 98
responding level, 98
valuing level, 98
Age, gender, and development, 355
Age equivalent scores, 344–345
Age-related factors in interpreting standardized tests, 355
Algebra, 362–363
Alternate forms reliability, 306
methods of, 325–327
Ambiguity, 201
American Association of School Administrators (AASA), 51
Analysis level of cognitive domain, 96
Analyzing the test, 196–207
criterion-referenced test, modifications for, 203–207
qualitative item analysis, 202–203
quantitative item analysis, 196–201
A Nation at Risk: The Imperative for Educational Reform, 21
Answers, recording, 192
Answers for practice questions, 402–407
Answer key, 192, 195
Application level of cognitive domain, 96
Aptitude, 218, 331, 356–358
Aptitude-achievement discrepancies, 358–361
Assembling the test, 191–193
checklist, 193
computers used for, 191–192
reproducing, 193
Assessment. See also Testing and assessment:
defined, 10
formative and summative, 4–5
performance-based, see Performance-based assessment portfolio, see Portfolio assessment tests vs., 7–8
Assessment process, 2, 7–10
Asymmetrical distributions, 244–245
Attitude, of test administrators, 193–194, 325
Aurora Public Schools (Colorado), 159–160
Austin American-Statesman, 8
Average, arithmetic, 246–247
AYP (Adequate Yearly Progress), 58

B
Back-end equating, 226–229
Band interpretation, 321, 328–332
add and subtract standard error of measurement, 329
defined, 320, 328
determine standard error of measurement, 329
graphical results, 329
interpret the bands, 330–332
listing data, 329
methods, comparison of, 330–331
Bar graph, 240, 241
Basic-fundamental movements, 100
Behavioral (specific) objectives, 81–82
Bimodal distribution, 253
Binet, Alfred, 381–382
Blind scoring, 195
Bush, Jeb, 8

C
California Achievement Test, 359
CASE (Council for the Advancement and Support of Education), 51
Cattell-Horn-Carroll (CHC) theory, 384
INDEX

Causality, 286–287
CCSS (Common Core State Standards), 26, 81, 336
CCSSO (Council of Chief State School Officers), 31
Characterization level of affective domain, 99
Checking test copies, 195
Checklists:
  in scoring rubrics, 167
  systems of, 220–221
Classroom dialogue, a summary, 390–395
Cloud computing, 84
Coefficient Alpha, 308
Coefficient of determination, 287–288
Coefficients, validity, 294
Cognitive Abilities Test (CogAT), 359, 382, 384
Cognitive domain, 94–97
  analysis level, 96
  application level, 96
  comprehension level, 95–96
  evaluation level, 97
  knowledge level, 95
  performance objectives in, 160–162
  synthesis level, 97
Collecting tests, 195
Common Core State Standards (CCSS), 26, 81, 336
Competency Testing For Teachers, 31–32
Completion tests, 123–126
  advantages/disadvantages of, 129
  good and bad items in, identifying, 124–125
  writing, suggestions for, 125–126
Comprehension level of cognitive domain, 95–96
Comprehensive Tests of Basic Skills, 346, 348, 349
Computers:
  objective test items and, 128
  quantitative item analysis and, 202
  test blueprints and, 100–103
  used for writing instructional objectives, 83–84
Concurrent validity evidence, 294–295
Construct validity evidence, 295–296
Criterion-referenced tests (CRT), 12–13
  defining, 69–72
  linguistic and cultural diversity in, 74–75
  norm-referenced vs., 72–73
  validity evidence and, 75
Criterion-related validity evidence, 294–296
  concurrent, 294–295
  predictive, 295
CTB-McGraw-Hill, 379
Curriculum-based measurements (CBM), 11, 13, 24
  formative assessments, 45
  progress monitoring, 45–46, 62
Curriculum-based measurements of reading fluency (CBM-R), 46–47
Curvilinearity, 288

D
Darwin School (Winnipeg, Manitoba), 157
Data:
  continuous, 240
  frequency, tabulating, 233–239
  graphing, see Graphing data
Data-based decision making in RTI, 46–48, 53
Debriefing, 207–209
  changes, being equitable with, 209
  double-checking work, 209
  on-the-spot decisions, avoiding, 206–207
  problem items, discussing, 208
  problem items, identified by students, 207
  student reactions, listening to, 208
Decision-making (types of educational decisions), 57–62
  administrative policy, 62
  counseling and guidance, 61
  diagnostic, 60–61
  grading, 60
  instructional, 59–60
  placement, 61
  program or curriculum, 61–62
  selection, 61
Defensibility, 169
Development-related factors in interpreting standardized tests, 355
Deviation, see Standard deviation (SD)
Deviation score method, 264
Diagnostic achievement tests, 380–381
Diagnostic decisions, 60–61
Dialectical journal format, 150–151
Difficulty index, 196–198
Directions for tests, checking, 192
Directly measurable learning outcomes, identifying, 89
Disability-related factors in interpreting standardized tests, 356
Discrepancy approach for SLD determination, 41, 358–359
Discrimination index, 198–200
Distractors, 196
Distractions, minimizing, 195
Distributions:
  bimodal, 253
  grouped frequency, 235–239
  heterogeneous, 267
  hypothetical scores of, see Standard error of measurement ($S_m$)
  measures of central tendency in, 254–255
  negatively skewed, 245–246
  normal, 267–270
  rectangular, 252
  simple frequency, 234–236
  symmetrical and asymmetrical, 244–245
  unimodal, 252

E
Ease of construction, 169
Educational Testing Service (ETS), 32
Education of All Handicapped Children Act (EAHC), 23
Education reform:
  Competency Testing For Teachers, 31–32
  Globalization and International Competitiveness, 30–31
  Merging Regular and Special Education Reform: IDEIA and NCLB, 24–25
Other Trends: Computer-Adaptive Achievement Testing, 29–30
Race to the Top (RTT), 27–28
Regular Education Reform, 21–23
Special Education Reform, 23–24
Standards-Based Reform, 25–27
Teacher Evaluation Based on Student Test Scores:
  Value-Added Models (VAM), 32
  The Future: NCLB, RTT, CCSS, and the CCSS-Aligned Tests, 28–29
Edwards Personal Preference Schedule (EPPS), 387
Elementary and Secondary Education Act (ESEA), 21
Emotional state on test day, 355–356
Equating before weighting, 225–229
  back-end equating, 226–229
  front-end equating, 226
Equivalence reliability, 305
Error. See also Standard error of measurement
  defined, 316–317
  in reliability coefficients, see Reliability coefficients, error sources in
  in scoring, 325
  in test administration, 325
  within test-takers (intra-individual error), 324
  Error score distribution, see Standard error of measurement ($S_m$)
  Essay cumulative portfolio rating form, 181
  Essay portfolio rating form, 279
  Essay tests, scoring, 140–146
    content, 144
    of extended response and higher level questions, 142–144
    of general essays, 146
    organization, 143–145
    process, 145–146
    rating method, 146
    rating scale for, 144
  Essay tests, writing, 132–154. See also Open-book questions and exams
    advantages and disadvantages of, 138
    defined, 134
    extended response items, see Extended response essay items
    goals of, 133
    planning guidelines, 153–154
    potential of, 134
    restricted response items, see Restricted response essay items
    suggestions for, 139–140
  ETS (Educational Testing Service), 32
  Evaluation level of cognitive domain, 97
  Expressive (general) objectives, 81–82
  Extended response essay items, 135
    defined, 135
    scoring, 142–144

F
50th percentile, 248–252
Formats in objective test items, 107–108
  advantages and disadvantages of, 128
Formative assessments, 4–5, 10, 46
Formative progress monitoring assessments, 337
Free appropriate public education (FAPE), 23
Frequency data, tabulating, 233–234
  grouped frequency distribution, 235–236
  lists, 234
  simple frequency distribution, 234–235
Frequency polygon, 240–243
Front-end equating, 226

G
Gender and racial bias:
  in interpreting standardized tests, 355
  in objective test items, 126
Globalization, 30
Globalization and International Competitiveness, 30–31
Global Scholar/Spectrum K12 School Solutions, 38
GPA, 299–300
Grade equivalent scores, 343–344

Grades/grading:
- on the curve, 216. See also Norm-referenced tests (NRT)
- decisions, 59–60
- review of, 395

Graphing data, 239–246
- bar graph, or histogram, 240
- frequency polygon, 240–243
- negatively skewed distributions, 245–246
- smooth curve, 243–245
- symmetrical and asymmetrical distributions, 244–245

Group-administered academic aptitude testing, 384–385

Grouped frequency distribution, 235–239
- constructing, 237–239
- range of scores in, 237

Grouping test items, 191

Guessing, 201

H

Habits of mind, 160

Hans Huber Medical Publisher, 387

Heterogeneous distribution, 267

Higher Education Act, 32

High Stakes Testing (HST), 20–33, 336–340

Histogram, 240

Holistic scoring, 168–169

HST, see High-stakes testing (HST)

Hypothetical scores, 317, 323. See also Standard error of measurement ($S_m$)

I

Individual Educational Plan (IEP), 58, 339

Individually administered academic aptitude testing, 384–385

Individuals with Disabilities Education Act (IDEA), 24

Individuals with Disabilities Education Improvement Act (IDEIA), 24–25, 33
- implications for general education teachers, testing, and assessment, 339
- response to intervention and, 38, 42
- Specific Learning Disability determination and, 44, 359
- students with disabilities participating in standardized tests, 339–340
- teacher requirements in decision-making, 57, 81

Instructional objectives. See also Learning outcomes; Taxonomy of educational objectives

 behavioral and expressive, 81–82
- checklist for writing, 92
- components of, 88
- computers used for writing, 83–84
- matching test items to, 92–94
- origin of, 83
- reasons for, 78–83
- taxonomy of, 94–100
- test blueprint and, 100–103
- as time-saver for teachers, 83–84

Instructions and explanations, 325

Internal consistency methods of estimating reliability:
- Kuder-Richardson methods, 308–309
- problems with, 309
- split-half methods, 307–308

Internal consistency methods of score reliability, 326–327

International competitiveness, 30

Interstate New Teacher Assessment and Support Consortium (INTASC), 31

Intra-individual error, 324

Intra-test error, 324

Iowa Tests of Basic Skills (ITBS), 346, 364, 379

IQ scores, stability of, 382

IQ tests, 234, 356–358

IQ tests, predictions of, 383–384
- emotional adjustment, 383
- happiness, 483–484
- job success, 483

Item analysis, 196

Item Analysis Terminology, 196–197

Item-total correlations, 308

J

Jargon checklist, 246

John Wiley & Sons, 33

Joint Committee on Testing Practices (JCTP), 33

K

Kaufman Assessment Battery for Children-II, 382

KeyMath 3 Diagnostic Assessment, 380

Knowledge level of cognitive domain, 95

Knowledge organization in:
- assessing, 147–149
- planning, 153–154
- webbing, 149, 155

Kuder-Richardson methods of reliability, 308–309
- formula 20 (KR20), 308
- formula 21 (KR21), 308

Kuhlmann-Anderson Intelligence Tests, 382

K

Language Arts Plan, 227

Learning activities, 89

Learning outcomes:
- conditions, stating, 90
- criterion levels, stating, 90–91
- identifying, 89–90
- exercises in, 89–90
- observable and directly measurable, 89–90
- keeping it simple, 91–92
- learning activities vs., 88

Least restrictive environment (LRE), 23

Letter grades, 219–220

Licensing tests, 32
Linguistic and cultural factors in interpreting standardized tests, 354
Local norms, 354
establishing, 354
interpreting, 355
national vs., 349
norms tables, specialized, 349

M
Machine-scored answer sheets, 195
Marks and marking systems, 213–229
combining, 221–225
comparisons in, see Marking systems, comparisons
in concerns over, 214
symbols in, see Symbols in marking systems
weighting, 221–225
Marking systems, comparisons in, 216–219
of achievement with effort, 218–219
of achievement with improvement, 219
with aptitude, 218
choosing a system, 219
with established standards, 217
with other students, 216–217
Massachusetts Comprehensive Assessment System (MCAS), 337
Matching tests, 111–114
advantages and disadvantages of, 128–130
ambiguous lists in, 112
guessing in, 112
lack of homogeneity in, 111–112
poor directions in, 112
revising, 113–114
suggestions for writing items in, 114
too many correct responses in, 112
wrong order of lists in, 112
Math problem-solving cumulative portfolio rating form, 183
Math problem-solving portfolio rating form, 182
Math Skills review, 16
Mean (M), 246–248
Measurements, educational, 62–66. See also Instructional
objectives
of achievement in the classroom, 209
of central tendency, see Measures of central tendency
defining what to measure, 64–65
how to measure, 65–66
pinch of salt scenario, 62–66
three-stage model of, 76–78
Measures of central tendency, 246–255
jargon checklist, 246
mean, 246–248
median, 248–252
mode, 252–255
in various distributions, 54–255
Median (MDN), 248–252
Merging Regular and Special Education Reform: IDEIA and
NCLB, 24–25
Middle score (Median), 248–252
Minnesota Multiphasic Personality Inventory-2 (MMPI-2), 387
Minnesota Multiphasic Personality Inventory-2-RF
(MMPI-2-RF), 387
Miskeying, 200–201
Mode, 252–255
Monitoring students, 177
Motivation-related factors in interpreting standardized tests, 355
Multiple-choice tests, 114–119
advantages/disadvantages of, 128–19
analogies demonstrating relationships among terms, using,
121–122
construction applied to higher-level questions, 114–119
construction applied to knowledge-level questions, 113–114
good and bad items in, identifying, 108–111
grammatical clues in, 108–109
learned principles or procedures used to novel situations, 122
multiple defensible answers in, 116
opinionated items in, 118
pictorial, graphical, or tabular stimuli, using, 121
redundancies in, eliminating, 119
stem and option in, 117
stem clues in, 117
using “all of the above” or “none of the above”
as options, 119
writing, suggestions for, 123
Myers-Briggs Type Indicator (MBTI), 387

N
NASDSE (National Association of Departments of Special
Education), 51
National Association of Departments of Special Education
(NASDSE), 51
National Board for Professional Teaching Standards (NBPTS),
31
National Center for Research on Learning Disabilities (NCRLD),
42
National Center for Student Progress Monitoring, 45
National Commission on Excellence in Education, 21
National Council on Measurement in Education (NCME), 33
National Joint Committee on Learning Disabilities (NJCLD), 52
National Research Council, on limitations in reliance on tests, 32
Negative correlation, 281
perfect, 283
Negative discrimination index, 197
Negatively skewed distributions, 245–246
Nelson-Denny Reading Test, 380
No Child Left Behind (NCLB), 20, 21–22, 24–25
students with disabilities participating in standardized tests,
339–340
Nondiscursive communication, 100
Nonverbal tests, 66
Normal distribution, 267–270
defined, 268
properties of, 268–270
Normative sample, 352
Norming group, 352
Norm-referenced tests (NRT), 12–13
criterion-referenced vs., 72–73
defining, 69–72
linguistic and cultural diversity in, 74–75
validity evidence and, 75–76
Norms tables, specialized, 349
Numerical grades, 220
Objective test items, 106–131
completion items, 124–126
defined, 107
formats, 107–108
formats:
advantages and disadvantages of, 128–130
gender and racial bias in, 126
matching items, 111–114
multiple-choice items, 114–123
true-false items, 108–111
writing, guidelines for, 127–128
Objective tests, 11, 66
personality, 386–387
Observable learning outcomes, identifying, 89–91
Odd-even reliability, 307–308
One, Two, Three Times Plan, 227
On-the-spot decisions, avoiding, 208–209
Open-book questions and exams, 149–153
dialectical journal format, 151–152
example of, 151
planning, 153–154
quotations, paraphrases, or summaries, 151–152
techniques, 149–153
Organization level of affective domain, 98
Otis-Lennon Mental Ability Tests, 382
Otis-Lennon School Abilities Tests:
Eighth Edition (OLSAT-8), 380, 385
Seventh Edition (OLSAT-7), 380
Performance and portfolio assessment, 9
standardized tests vs., 341
Performance-based assessment:
of affective and social skills, 158–160
developing, see Performance-based assessment, developing
as direct measures of learning, 157
embedding in lessons, 158
habits of mind in, 160, 162
of processes and products, 157–158
Performance-based assessment, developing, 160
designing assessment context, 162–165
scoring rubrics, see Rubrics in performance-based assessment
testing constraints, specifying, 170–171
Performance objectives:
in affective and social domain, 162
in cognitive domain, 160–162
Personality:
defined, 385
objective personality assessment, 385–386
projective personality assessment, 386
Physical abilities, 100
Physical comfort, 325
Placement decisions, 61
Portfolio assessment, 174–188
defined, 174–175
developing, see Portfolio assessment, developing
ownership of, 178
purpose of, 174–175
rationale for, 175–176
relevance of, 176
representativeness in, 176
rubrics in, 176
validity of, 176
Portfolio assessment, developing, 176–188
access to portfolios, 186
checklist, 286–288
cognitive skills and disposition, identifying, 177
conferences, planning, 186
how products are kept, 186
how products are turned in and returned, 186
products to include, deciding on, 175
purpose for, deciding on, 175
rubrics, building, 179–183
scoring criteria for, traits to consider, 181
timelines, 186
who will plan the portfolio, deciding, 177
Portfolio rating forms:
aggregating all, 184–185
eSSay, 179–181
eSSay cumulative, 181
Math problem-solving, 182–183
Math problem-solving cumulative, 183
weighting, 183
Positive correlation, 284
Positive discrimination index, 197
Power tests, 66, 309
Pre- and posttests, 203–204
Predictive validity evidence, 295–302
Primary Test of Cognitive Skills (PTCS), 385
Primary trait scoring, 167–168
standard protocol approach vs., 48–50
Program for International Student Assessment (PISA), 31
Program or curriculum decisions, 61–62
Progress Monitoring (PM), 9, 13, 33, 52
Projective personality tests, 387
Proofreading, 191
Psychomotor domain, 99–100
basic-fundamental movements, 100
nondiscursive communication, 100
perceptual abilities, 100
physical abilities, 100
reflex movements, 99
skilled movements, 100
Q
Qualitative item analysis, 202–203
Quality of feedback, 170
Quantitative item analysis, 196–202
ambiguity, 201
customers used for, 202
discrimination index, 198–200
guessing, 201
item analysis terminology, 197
miskeying, 200–201
R
Race to the Top (RTT), 27–28, 336
Range of scores:
in grouped frequency distribution, 237–239
inclusive, 259
semi-interquartile, 260–261
in variability, determining, 258–259
Rank difference correlation \((r_p)\), 285–286
Rating method of scoring, 146
Rating scales in scoring rubrics, 167–168
Raw scores, 270–271. See also Standard deviation (SD)
Raw score method, 265–267
RC (Restructured Clinical) scales, 387
Receiving (attending) level of affective domain, 98
Recognition items, 123
Rectangular distribution, 252
Reflex movements, 99
Reform initiatives:
merging regular and special education, 24–25
regular education reform, 24–25
special education reform, 24–25
trends in, 29–30
Regression line, 284
Regression toward the mean effect, 218
Regular Education Reform, 21–23. See also
Education reform
Reliability:
defined, 170, 293
of interpreting standardized tests, test-related factors, 350–351
review of, 393
Reliability, methods of estimating:
alternate forms or equivalence, 306
internal consistency, see Internal consistency methods of estimating reliability
test-retest or stability, 305–306
Reliability coefficients, 305–312
defined, 305
error sources in, see Reliability coefficients, error sources in estimating, see Reliability, methods of estimating interpreting, 309–312
Reliability coefficients, error sources in alternate-forms, 325–327
internal consistency, 326–327
test-retest, 325–326
Report cards, 219. See also Symbols in marking systems
Responding level of affective domain, 98
Response to intervention (RTI), 24
benefits of, 73
components of, 43–48
definitions, 42
educational outcomes for students and, improved, 41
educational reform and, 40
explained, 38
impact of, on regular education teachers, 38
implementation approaches to importance of, to regular education teachers, 40–41
origin of, 37
promises and controversies of, 52–53
scenario, 38–40
severe discrepancy model and, 331
Restricted response essay items, 135–137
defined, 136
examples of, 136
when to use, 136–137
Restructured Clinical (RC) scales, 387
Riverside Publishing Company, 379, 384
Rorschach Inkblot Technique, 387
RTI Action Network, 42, 50–52
Rubrics in performance-based assessment, 165–170
for accomplishments, 165
checklists, 166
combining scoring systems, 169
comparing scoring systems, 169
holistic scoring, 168–169
point limitation in, 169
rating scales, 169–170
Rubrics in portfolio assessments, 176–188. See also Portfolio rating forms
Rules, clarifying, 194
INDEX

S

Scaled score (SS), 349

Scores:
- converted, see Converted scores hypothetical, 317, 323
- IQ, 356–357
- obtained, 316
- overinterpreting small differences in, 321. See also Band interpretation

range of, in grouped frequency distribution, 2–239
- raw, 270–271
- scale, 347
- stanines, 346–347

Scores offered for standardized tests, 343–347
- age equivalents, 344–345
- grade equivalents, 343–344
- percentile ranks, 345–346
- standard scores, 346–347

Score reports from publishers, 366–373
- Home Report, 366–367
- Individual Profile Report, 366–368
- Group List Report, 366, 369
- Your Child’s Test Scores, 366, 370

Scoring efficiency, 170

Scoring rubrics, see Rubrics in performance-based assessment

Scoring the test, 194

Selection decisions, 61

Self-check test, 16

Semi-interquartile range (SIQR), 260–261

Severe discrepancy model, 331

Simple frequency distribution, 234–235

Skilled movements, 100

Slosson Intelligence Test, 382

Smarter Balanced Assessment Consortium (Smarter Balanced), 27–30, 336–341

Smooth curve, 243–245

Social domain, performance objectives in, 162

Spacing test items, 191

Spearman-Brown prophecy formula, 311

Special Education Reform, 23–24. See also Education reform

Specific Learning Disability (SLD), 24, 44, 359

Speed tests, 66

Speeded tests, 309

Split-half reliability, 307–308

Stability reliability coefficients, 304–312

Standards-Based Reform, 25–27

Standard deviation (SD), 261–267
- defined, 223, 260
- deviation score method for computing, 264
- of error score distribution, see Standard error of measurement \( (S_{\text{e}}) \)
- raw score method for computing, 265–267
- standard error of measurement vs., 323

Standard error of measurement (\( S_{\text{e}} \)), See also Band interpretation; Error

Standardized tests, interpreting:
- parent-teacher conferences and educational decision making, 350–358
- score reports from publishers, 366–370
- student-related factors, 354–358
- test-related factors, 351–354

Standardized tests, types of, 377–378
- academic aptitude, 381–385
- achievement tests, 378–381
- standardized personality assessment instruments, 385–388

Standard protocol (SP) approach to RTI implementation, 48

Stanford Achievement Test 10th edition (SAT-10), 380

Stanford-Binet Intelligence Scale, 381, 384

Stanford-Binet V, 384

Stanford Diagnostic Mathematics Test, Fourth Edition (SDMT4), 381

Stanford Diagnostic Reading Test, Fourth Edition (SDRT4), 381

Statistics. See also Correlation; Data; Measures of central tendency
- defined, 232–233
- using, reasons for, 233

Student-related factors in interpreting standardized tests, 349–350

Standardized criterion-referenced tests:
- defined, 337

Standardized norm-referenced tests, 338

Standardized personality assessment instruments, 385–388
- objective personality tests, 386–388
- personality, defined, 385
- projective personality tests, 386
age, gender, and development, 355
aptitude, 356–358
disabilities, 356
emotional state on test day, 355–356
linguistic and cultural, 354–355
motivational, 355
Subjective tests, 66
Summary, classroom dialogue of, 390–395
Summative assessment, 4–5, 9, 366–338
Survey batteries, 379–380
Symbols in marking systems, 219–221
checklists, 220–221
letter grades, 219–220
numerical grades, 220
pass-fail (P-F), 220
Symmetrical distributions, 244–245, 267
Synthesis level of cognitive domain, 97

T
table of specifications, see Test blueprint
Taxonomy of educational objectives, 94–100
affective domain, 97–99
cognitive domain, 94–97
psychomotor domain, 99–100
Teachers:
accountability and the classroom
teacher, 57–58
competency testing for, 31–32
effective/successful, factors associated with, 84
impact of IDEIA and NCLB on, 24–25
objectives, see Instructional objectives
response to intervention and, 36–53
training in testing and assessment, 13
Teacher Evaluation Based on Student Test Scores: Value-Added
Models (VAM), 32
Teacher licensing tests, 31–32
Teacher-made tests, 12
Terman, Louis, 381
Terra-Nova, Third Edition, 337
Tests, types of, 10–13
criterion-referenced, 12–13
curriculum-based measurements, 13
essay, 11
high-stakes, see High-stakes testing (HST) norm-referenced,
12–13
objective, 11
performance and portfolio, 11
power, 309
speeded, 309
standardized, 11
teacher-made, 11
written, 65–66
Test batteries, 379–380
Test blueprint, 100–103
categories, 102
True-false tests, 108–109
  advantages and disadvantages of, 128–130
  double-barreled items in, 110
  good and bad items in, identifying, 108–111
  guessing factor in, 112
  writing, suggestions for, 110–111
Truncated range, 288–289
T-scores, 277
Twin Peaks Middle School (Poway, California), 158
U
  Unimodal distribution, 253
  Universal Screening, 23, 44–45
V
  Validity, defined, 292–294
  Validity coefficients:
    concurrent and predictive validity evidence, 298–302
    content validity evidence, 298
  Validity evidence, 292–296
    concurrent and predictive, 295
    construct, 295–296
    content, 293–294
    criterion-related, 294–295
    defined, 292–293
    review of, 296–298
  Valuing level of affective domain, 98
Variability:
  estimates of, 259, 263
  range in, determining, 259
  standard deviation, 261–267
Variance, 263
Verbal tests, 66
W
  Waco, Texas, public schools, 7, 8
  Webbing, 146
  Wechsler Adult Intelligence Scale-IV (WAIS-IV), 384
  Wechsler Intelligence Scale for Children V (WISC-V), 384
  Wechsler Preschool and Primary Scale of Intelligence-IV
    (WPPSI-IV), 384
  Weighting, 183, 221–225
    differential, 221
    equating before, 225–229
  West Orient School (Gresham, Oregon), 158
  Within-test error, 324–325
  Words read correctly per minute (WRCPM), 47
  Written tests, 65–66
    how to measure, 65–66
    types of, 66
Z
  Zero discrimination index, 197
  z-scores, 273–276