Subject Index

access, increased of Web-based tests 25–26, 57
ACT Placement Test (American College Test) 32–33, 49–50
American Psychological Association 81
anxiety, of test-takers 17, 51, 95
applied linguistics, and technology
  comparison approach 116–17
  innovative approach 117
  language ability, and use 107–9
  language teaching, and understanding assessment 113–14
  SLA, and language assessment 109–13
  'tunnel approach' 116
artificial intelligence 112
Assessing Language for Specific Purposes (Douglas) 1
assessment, systems underlying
  activity selection process 75–76, 78
  design architecture, common framework for 74–75
  functionality requirements 74
  presentation process 75, 76, 78
  response processing 75, 76–77, 78
  summary scoring process 75, 77, 78
authenticity
  contextualization 27–28
  interactional 9–10
  situational 9–10, 28, 29
  technology, and language use options 13–14, 94–95
authoring tools
  assessment purpose 62, 79
  classroom assessment 6
  commercial companies, working with 79
  design issues 83
  examples of 72–73 (*see also* Web CT)
  'in-house' systems, and flexibility 77–78
  number and variety of 63
  professional quality, need for 114–15
  see also assessment, systems underlying
automated speaking assessment 10
automatic response scoring *see* linguistic responses

Blackboard 73
Brigham Young University 8, 88–89
Business Language Testing Service (BULATS) 26

CALICO Journal 63
CALL *see* computer-assisted language learning

CATs *see* computer-adaptive tests
CEF Construct Project (Alderson et al) 115
classroom assessment
  authoring software 6
  authoring software, and test creation 4
  online language courses 4–6
  stand-alone tests 5
cloze test responses 54
Common European Framework 104, 115
communicative competence
  computer-adaptive tests 34
  computer-mediated communication 108–9
  language and technology, joint role of 107–8
computational linguistics 112
computer-adaptive tests (CATs)
  advantages and disadvantages of 86
  construct validity 86, 91, 93–94
  input-response interaction 23, 24, 32–34
  item difficulty 7–9
  Item Response Theory 8, 32–33
  item selection 50–51
  placement test, and usefulness 88–89, 90
  quality and efficiency 8, 87
  reading assessment 1, 51
  security issues 58
computer-assisted language learning (CALL) 63, 113
computer-method effects *see* method effects
'Computer technology in test construction and processing' (Baker) 81
Computerized Oral Proficiency Instrument (COPI) 38
construct definition
  clarification, of test constructs 12–14
  detailing, of test constructs 14–15
  and new methods 10–12
construct operationalization 14–15
construct theory, and scoring methods 56– 57
construct validity 86, 91, 93–94
'Construct validity of computer-based tests' (Green) 81
contexts of language use
  authenticity 27–28
  constructed contexts, and performance 111–12
  language tests as 20–21
  method effects 15–16
convenience 22, 23
correctness, degrees of 56–57
course management software 6
Criterion (Educational Testing Service) 10, 14–15, 35–37
criterion-referenced testing 5
cross-disciplinary collaboration
  intelligent computer-assisted language learning systems 115
  professional quality authoring tools, need for 114–15

data handling, use of computers for 2
developers *see* test developers
DIALANG Web-based test 5, 25–26, 32, 57, 58, 104, 115
dichotomous *versus* polytomous s coring
  correctness, degrees of 56–57
  short constructed responses 55
  text-sequencing task 52–53
difficulty *see* item difficulty
Discovery School Quiz Center (Discovery School) 73

e-rater (Educational Testing Service) 36–37, 54
Educational Measurement 81
educational measurement, and test quality 3, 81
Educational Testing Service 10
efficiency
  classroom assessments 6
  evaluative research 87, 88–89
  test performance difference 44
error analysis 4–5
essays, scoring of 36–37, 53–54, 110
Eurocentres Vocabulary Size Test 1

evaluation
  evaluative research
    affect, performance and efficiency, impact on 88–89
    linguistic responses 87–88
  quality, promotion of
    computer-adaptive tests, advantages and disadvantages of 86
    educational measurement perspective 81
    interface issues 83–85
    positive qualities, focus on 86
    unique test factors, of CALT 81–83
  see also test usefulness; validation

fairness, issues of 3
feedback, and increased options 81, 83
'four-process architecture' *see* assessment, systems underlying

gap-filling tasks 55
'General Performance Analysis' 4
Graduated Record Examination (GRE)
  analytic writing section 52
  security issues 58
  test item comparison studies 43–44
grammar tests
  SLA grammar acquisition, knowledge about 110–11
  test item selection 50–51
  see also WebCT
Grammar/Usage Testlet *(ACT Placement Test)* 32–33, 49–50
Guidelines for Computer-Based Tests and Interpretations (APA) 81

high-stakes tests
  and fairness 3
  and security 57, 58
Hot Potatoes (Half-baked) 72
human and machine ratings, correlation between 87–88

IBM Model 805 Test Scoring Machine 34, 35f

IELTS *see* International English Language Testing System
impact, of CALT
  evolutionary *versus* revolutionary changes 103–7, 117
  language assessment, advances in 105–7
  language-related knowledge, need for increased 103–4
instructions, test 22, 23, 26, 27
intelligent computer-assisted language learning systems 115
intelligent tutoring systems 5, 117
interactional authenticity 9–10
interface issues
  planning and initial design 81, 83, 84–85
  usability testing/field trials and fine tuning 83, 85–86
International English Language Testing System (IELTS) 25
Internet, and increased access to tests 25–26, 57
item analysis 32
item difficulty
  computer-adaptive tests 7–9
  examinee anxiety 51
  true score theory 32
  and validity 49–50
Item Response Theory 8, 32–33

Journal of Educational Measurement 81

language ability
  versus computer-related ability 42–43, 45
  technology choices 107
  and use 107–9
language for specific purposes (LSP) 16
Language Learning and Technology (journal) 63
language-related knowledge, need for increased 103–4
language teachers
  assessment, need for understanding of 113–14
  classroom assessment 4–6

fairness, issues of 3
test selection 3–4
Language Testing Research Colloquium 18
large-scale assessment
ability levels, optimization of 87
administration of 25
transformative impact of CALT on 103
linguistic responses, scoring of
construct theory, and scoring methods 56– 57
dictation test task 55–56, 110
essays, scoring of 53–54, 110
evaluative research 87–88
short constructed responses 54–55
literacies, multiple technology-defined 17–18, 108–9
literacy studies, and language assessment 107–9
Longman English Assessment 9
Longman English Interactive (Rost) 2, 113
low-stakes tests *see* DIALANG Web-based test

Market Leader (Longman) 5, 113
method effects
contexts of language use 15–16
language performance 21–22, 47–48
Michigan English Language Assessment Battery (MELAB) 25

National College English Test (China) 25
natural language processing (NLP) 10, 14, 23, 24, 36, 37

open-ended responses 94, 110
Oral English Proficiency Test (Purdue University) 38
outline completion tasks 55

partial-credit scores, and reliability 90
performance analysis 4–5
PhonePass (Ordinate Corporation) 10
PLATO computer, and French curriculum 5
polytomous scoring *see* dichotomous *versus* polytomous scoring
practicality
constructed response language test tasks 38
test developer expertise 92, 96
preparation materials 27, 46
proficiency tests 5
pronunciation analysis 5
psychometric models 87
Purdue ITA speaking test 49

quality
test quality checklist 3–4
see also evaluation
Questionmark 73

rater help 38
ReCall Journal 63
reductionism, and computer-adaptive tests 50, 51
research, implications of CALT for
construct definition 12–15
SLA research, connecting to 16–17
test consequences, investigation of 17–18
validation, rethinking 15–16
validity, threats to 60–61
see also applied linguistics, and technology
rubric/instructions, test 22, 23, 26, 27

score interpretation *see* validation; validity, threats to
second language acquisition (SLA)
'assessed SLA' 113
constructed contexts, and performance 111–12
ESL writing levels, work on 110–11
grammar acquisition sequence, professional knowledge about 110–11
'instructed SLA' 113
'intelligent assessment' 112–13
processing conditions, effect of 109–13
research implications 16–17

security issues 8, 22, 23, 41, 57–58
self-access classroom activities 6, 7
self-assessment, computer-assisted 26, 113
SET-10 (Ordinate Corporation) 37–38, 52
situational authenticity 9–10, 28, 29
SLA *see* second language acquisition
speech recognition systems 10, 37–38
statistical analysis, computerized of non-computer tests 2
Surrey English Language Institute 6, 7

task types, new 28–29, 30–31
teachers *see* language teachers
'Technical guidelines for assessing computer-adaptive tests' (Green et al) 81
test developers, implications for
  authenticity 9–10
  automatic response analysis 10
  computer-adaptive tests, and item difficulty 7–9
  expertise 96
  new methods, availability of 6
  test constructs, and new methods 10–12
test method characteristics
  assessment characteristics 23, 24
  complex language performance, scoring of 34–35
  contextualization, and authenticity 27–28
  help features and tutorials 28–29, 30–31
  input-response interaction 23, 24, 31–34
  language tests, as contexts of language use 20–21
  method effects 21–22
  natural language processing technology 23, 24, 36, 37
  practicality, of test tasks 38
  rater help 38
  rating procedure 37–38
  rubric/instructions 22, 23, 26, 27
  scoring, and use of computers 34–35
  'setting' 22, 23
  test circumstances 22, 23, 25–26
  written responses, automated system for 35–37
Test of English as a Foreign Language see TOEFL
Test of Spoken English (Educational Testing Service) 87
test selection 3–4
test usefulness
  authenticity 91, 94–95
  construct validity 91, 93–94
  evidence for argument about 89
  framework for 80
  impact 92, 95
  interactiveness 91, 95
  practicality 92, 96
  reliability 90–92
Test Validity 81
'testlet approach' 32–33, 49–51
text analysis programs 53
text-sequencing task 52–53, 110
TOEFL *(Test of English as a Foreign Language)*
  comparison studies 44
  computer-based test
    help functions 28, 39
    help option 12–13
    Highlighting Task Tutorial 28–29, 31
    Instruction Screen 28, 39
    IRT/CAT techniques 32
    preparation materials 17, 27
    Reading Task 28, 39, 52
  computer delivery of 1, 2
  familiarity study 46, 88, 89
true score theory 32

UCLA's 'in-house' authoring tool 77–78
usefulness *see* test usefulness

validation
  assessment-use argument
    since-statements 100–101
    utilization argument 99–102
  interpretative argument 101
    scores, inferences based on 97
    since-statements 98–99
    and threats to validity 98

unless-statements 98–99
positive washback 17–18
research implications 15–16
validity, threats to
adaptive item selection 41
item difficulty 49–50, 51
research responses 60
test content 50–51
automatic response scoring 41
complex constructed tasks, scoring of 52
correctness, degrees of 56–57
linguistic responses 53–57
non-linguistic complex responses 52–53
research responses 60
future research, issues for 60–61
negative consequences 41, 58–59
security, compromised 41, 57–58
summary of 40–42
task types, new 41, 48–49
test performance difference 41
computer- *versus* language ability-related challenges 42–43, 45
computer-based test scores, inferences based on 40, 42–43
computer familiarity differences 45–46
differential validity 45–47
efficiency issue 44
method effects, interpretation of 47–48
research responses 60
test comparison studies 43–45
vocabulary assessment 1, 51
washback
negative 58
positive 17–18
Web-based Language Assessment System (WebLAS), University of California 9, 28, 55
WebCT
description of 63–64
examinee records, viewing
individual test taker reports 69–70
response summaries 70–71
Respondus 64, 65–66, 71
response analysis 71
student access, giving
Control Panel options 67
Designer Page, for page modification 67, 68
test preview 67–68
task types, flexibility of 71–72
test, publishing
Preview & Publish screen 66–67
test, taking
test format 68, 70
Test Homepage 68–69
test instructions 68, 69
test creation
Create button 65
Edit screen 65–66
Preview test ask 66
Retrieval and Reports screen 67
Start screen 64–65
test specification 64
Writer's Workbench 53
written responses, and automated response analysis 10, 35–37

Author index

Ackerman, T. 15
Alderson, J.C. 1, 5, 6, 10–11, 12, 16, 17, 22, 52, 54, 55, 57, 58, 104, 115, 117, 110
Almond, R.G. 74–7
Anbar, M. 55

Bachman, L.F. 2, 8, 20, 22, 80, 90, 97, 99–101, 112
Baker, E.L. 59, 81, 117
Bejar, I.I. 56, 103
Bennett, R.E. 58, 59, 103
Berbisada, N. 56, 110
Birenbaum, M. 88
Blais, J.-G. 87
Bleiler, T. 51
Bock, R.D. 8, 81
Bonk, W.J. 112
Boo, J. 32, 44
Boyle, T.A. 115
Bradlow, E.T. 32
Braun, H. 103
Bridgeman, B. 36–7
Brown, A. 8, 87
Brown, J.D. 8, 112
Brown, J.S. 116
Bruce, B.C. 108
Brutten, S. 8
Buck, G. 1, 22
Burstein, E. 54, 110
Burston, J. 8, 87
Bygate, M. 113

Campbell, J. 56, 110
Canale, M. 5, 50, 51, 117
Carr, N. 15
Carroll, J.B. 21
Carton A.S. 8, 87
Chalhoub-Deville, M. 8, 48, 87
Chanier, T. 115
Chapelle, C.A. 15, 17, 74, 90, 93, 112, 114, 116–17
Cheng, L. 114
Chodorow, M. 54, 110
Choi, I.-C. 32, 44, 98
Clark, J.L.D. 5
Cohen, A.D. 47, 97, 112
Cole, N. 103
Coniam, D. 13–14, 15, 44, 55–6, 110
Corbel, C. 11–12, 16, 117
Crookes G. 113
Crooks, T. 97
Crystal, D. 16
Curtis, A. 114

D'Arcy, S. 55
Dandonoli, P. 8, 87
Davidson, F. 2
Deville, C. 48

Dorans, N.J. 8, 42
Douglas, D. 1, 16, 22, 108, 111
Duguid, P. 116
Dunkel, P. 8, 87

Eckert, R.G. 115
Eignor, D. 42, 44, 45, 46
Embretson, S. 15

Farnum, M. 90
Figueres, N. 115
Flaugher, R. 8, 42
Fowles, M.E. 54, 90, 110
Fulcher, G. 34, 83–6, 89, 96, 101

Gass, S.M. 113
Ginther, A. 38
Green, B.F. 3, 8, 42, 81
Gross, S. 87
Gruba, P. 13

Hagen, L.K. 16
Halliday, M.A.K. 111
Hambleton, R.K. 8
Hamp-Lyons, L. 17, 58
Hasan, R. 111
Hegelheimer, V. 17, 90
Helm, C. 55
Hendrickson, A.B. 51
Henning, G. 55, 90
Hinkel, E. 111
Hogan, M.P. 108
Holland, V.M. 5, 15
Hone, A.S. 56
Hudson, T. 112
Humphreys, L.B. 8, 81

Inagaki, S. 110–11
Iwashita, N. 8, 87

Jamieson, J. 17, 44, 45, 46, 56, 90, 110

Kane, M.T. 97
Kaplan, J. 5
Kaya-Carton, E. 8, 87
Kim, H.-Y. 110–11
Kim, K.S. 32, 44
Kirsch, I. 44, 46
Kuijper, H. 115
Kukich, K. 54, 110

Larsen, J.W. 8
Laurier, M. 8, 86, 87, 96, 101
Linn, R.L. 8, 81
Liskin-Gasparro, J.E. 31
Lord, F.M. 8
Luoma, S. 26

Madsen, H.S. 8, 51, 88–9, 90, 93, 95
Malabonga, V. 38
Marty, F. 4
McKinley, R.L. 43
Messick, S. 57
Meunier, L.E. 86, 96, 101
Mills, C.N. 43
Mislevy, R.J. 8, 42, 74–7, 88
Molholt, G. 5, 87–8
Monaghan, W. 36–7
Monville-Burston, M. 8

Noijons, J. 3, 81–3, 96, 101
Nold, G. 115
Norfleet, L. 56, 110
Norris, J. 110, 112

Oller, J. 54
Ortega, L. 112
Otto 4

Palmer, A.S. 22, 90
Paltridge, B. 111
Pengelly, M. 115
Percsich, R. 52, 57
Perkins, K. 8
Powers, D.E. 54, 90, 110
Presler, A.M. 5

Ramsey, P. 90
Rassool, N. 108
Read, J. 1
Reckase, M.D. 8, 81
Reese, C.M. 43

Reid, J. 53
Robinson, P. 111
Roever, C. 25, 57
Rogers, H.J. 8
Rost, M. 2

Sams, M. 5
Sawaki, Y. 47, 48
Schaeffer, G.A. 43
Self, J. 115
Shermis, M.D. 8
Skehan, P. 111, 113
Smith, W.F. 115
Stansfield, C. 18
Steffen, M. 43
Steinberg, L. 8, 42, 45, 74–7
Stevenson J. 87
Stoynoff, S. 114
Swain, M. 113
Swaminathan, H. 8
Szabo, G. 52, 57

Takala, S. 115
Tardieu, C. 115
Tarnanen, M. 26
Tatsuoka, K.K. 88
Taylor, C. 44, 45, 46, 90
Thissen, D. 8, 42, 45
Tung, P. 8
Twidale, M. 115
Tyner, K. 108

Vispoel, W.P. 51

Wainer, H. 8, 42, 45, 47, 50
Warschauer, M. 18, 108–9
Watanabe, Y. 114
Weigle, S.C. 1
Williamson, D.M. 56
Wolfe-Quintero, K. 110–11
Wresch, W. 53

Young, R. 8, 87