

Contents

Acknowledgements	<i>page</i> ix
Preface	xi
1 Teaching and testing	i
2 Testing as problem solving: an overview of the book	8
3 Kinds of tests and testing	11
4 Validity	26
5 Reliability	36
6 Achieving beneficial backwash	53
7 Stages of test development	58
8 Common test techniques	75
9 Testing writing	83
10 Testing oral ability	113
11 Testing reading	136
12 Testing listening	160
13 Testing grammar and vocabulary	172
14 Testing overall ability	186
15 Tests for young learners	199
	vii

Contents

16 Test administration	215
Appendix 1 The statistical analysis of test data	218
Appendix 2 Item banking	234
Appendix 3 Questions on the New Zealand youth hostels passage	236
Bibliography	237
Subject Index	246
Author Index	250

Subject Index

- accent, 61, 116, 163, 171
- acceptable responses, agreeing, 49-50
- accuracy, 62, 127
- achievement tests, 12-15, 66-69, 173, 179
- ACTFL (American Council for the Teaching of Foreign Languages), 24, 98-100, 111, 135
- ALTE (Association of Language Testers in Europe), 12, 24-25, 73
- addressees, 60, 84, 85, 114, 116
- administration of tests, 48, 215-217
- alternate forms estimate of reliability, 39-40
- ambiguous items, 46-47
- anagrams, 208-209
- analytic scoring, 100-106, 130-133
- anchor items, 73-74, 232-233, 234, 235
- ANOVA (Analysis of Variance), 222
- aptitude, language, 10
- ARELS (Association of Recognised English Language Services), 198
- ARELS Examinations Trust, 122-123, 124, 166-167
- articles, 16, 175
- assessment, 2-5, 7
 - self, 5, 200-201
- authenticity, 17, 24, 93, 163
- backwash, 1-2, 12, 27, 53-57, 78, 86, 106, 135, 138, 141, 160, 172, 173, 179
- benchmark scripts, 107
- Berkshire Certificate of Proficiency in German, 20-21
- BNC (British National Corpus), 73, 174
- calibration, 65, 105, 106, 133
- CAEL (Canadian Academic English Language) Assessment, 24
- CCSE (Cambridge Certificate in Communication Skills in English), 62, 73, 84, 86-89, 113-114, 116-118, 118, 127, 135, 159
- change of topic, 119
- channel, 61
- cheating, 78, 79
- checklists for moderation, 64, 154
- classical item analysis, 225-228
- cloze, 153, 187-195, 197-198
 - conversational, 191-193
 - mini-, 194
 - summary, 149-151
- COBUILD, 73, 174, 185
- communicative language testing, 23, 25, 35
- compensation, 100
- completion items, 177
- computer adaptive testing, 22-23
- content, test, 59-61
- correlation coefficient, 28, 35
- Council of Europe, 73
- CPE (Cambridge Proficiency Examination), 12, 198
- criterial levels of performance, 61-62, 141-142, 163
- criterion referenced testing, 19-22, 42, 52, 55
- C-Test, 153, 194-195, 197
- decision consistency, 42
- definition items, 181, 183, 208
- diagnostic tests, 15-16, 19, 160, 162, 179

- DIALANG, 16, 24
 dictation, 195-196, 198
 partial, 168-169, 196
 mini-partial, 196
 direct testing, 17-19, 34, 54
 discontinuous grammatical elements,
 80-81
 discrete point testing, 19
 discrimination indices, 226-228
 discrimination, lack of, 45
 discussion in oral testing, 121
 distractors, 75, 77, 78, 180-182, 228
- ELTS (English Language Testing Service),
 35
 error analysis, 103
 error gravity, 103, 112
 ETA (Educational Test Analysis), 218, 233
 ethics, 2, 6
 EVA project, 200, 207, 214
- facility value, 225-226
 FBI listening summary translation exam,
 24
 FCE (First Certificate Examination), 12,
 176, 177
 feedback to candidates, 5, 108, 200
 fit of items and people, 229
 Flesch-Kincaid Grade Level Score, 73
 Flesch Reading Ease Score, 73
 flexibility, 62, 127
 formats (*see also* techniques), 119-124
 freedom, restricting candidates', 45-46,
 93
 frequency table, 219
 FSI (Foreign Service Institute), 130
- gap filling items, 80-81, 148-151, 166,
 183-184, 198, 210
 GCSE (General Certificate of Secondary
 Education), 55
 general knowledge, 142-143, 155
 grammar, testing of, 172-179
 graphic features of texts, 140
 group oral, 134
 guessing, 76-77, 79, 208
- halo effect, 103
 handbooks, test, 66
- histogram, 220
 holistic scoring, 94-100, 111, 128-130
- identification of candidates, 50
 ILR (Interagency Language Roundtable),
 20, 24, 100, 111, 127-130
 impact, 2, 6
 independence of items, 44, 155
 indirect testing, 17-19, 31
 inferences, 139-140
 information transfer, 91-93, 151-152,
 166-167, 206
 innovation theory, 57
 instructions, 47
 integrative testing, 19, 25
 interaction with fellow candidates,
 121-122
 internal consistency, coefficient of, 40
 interpreting, 120-121
 interruption, 119
 interview, 119-121
 introspection (*see also* retrospection), 171
 invitation to ask questions, 119
 item analysis, 225-231
 item banking, 234-235
 item response theory (IRT), 23, 41-42,
 228-233
 item-test correlation, 226
 items, number of, 61, 73, 76, 79
- Japan, 57
 Japanese, 213
 JMB (Joint Matriculation Board), 155
- language of items and responses, 153, 159
 layout of tests, 47
 listening, testing of, 160-171, 203-206
- mean, 220, 221
 median, 220, 221
 medium, 61
 Michigan Test, 21
 modal verbs, 16, 81
 mode, 221
 moderation, 63-64, 153-154, 168, 169
 monologue, 121
 multiple choice, 3, 48, 75-77, 79,
 143-144, 165-166, 204, 207-208,
 228

Testing for language teachers

- needs analysis, 85
- normal distribution, 52
- norm referenced testing, 19-22
- Norway, 199
- notes as basis of writing task, 93
- note taking, 167-168
- objectives, 13-15, 55, 200
- objective testing, 22, 48
- open-ended items, 48
- operations, 60, 84, 85, 89, 113-114, 136-140, 161-162
- options, 75, 76-77
- oral ability, testing, 53-54, 113-135
- overall ability, 186-198
- Oxford Preliminary, 156
- paraphrase items, 176-177
- passages, number of, 61
- past continuous, 80-81
- pictures, use of, 120, 182-183, 203-205, 207, 208-209, 210, 211
- placement tests, 16-17, 28-29, 70-72, 179, 186
- pop quizzes, 15
- portfolio assessment, 111
- practicality, 56, 89, 142
- practice materials, 48, 168
- proficiency tests, 11-12, 172, 179
- questions, 119
- range of scores, 222
- Rasch analysis, 228-233
- rating (*see also* scoring), 134
- readability, 61, 73, 140
- readership, intended, 140
- reading aloud, testing of, 121
- reading, testing, 136-159, 207-208
 - expeditious, 138
 - slow and careful, 138, 139-140
 - speed of, 141
- recordings as test stimuli, 122-124, 166, 170
- reduced redundancy, 70, 187
- referents, identifying, 144-145
- reliability, 3-4, 36-52, 53, 76, 94, 142, 222-223
 - and validity, 50
 - coefficient, 38-40, 51-52, 222-223
 - scorer, 43-44, 48-50, 52
 - inter-scorer, 52
 - intra-scorer, 52
- requests for information and elaboration, 119
- residuals, 230-231, 233
- retrospection (*see also* introspection), 32, 35
- role play, 120, 122, 135
- RSA (Royal Society of Arts), 159
- sample tests, 48
- sampling behaviour, 44-48, 54, 94
- sampling specifications, 62-63, 85-89, 116-117, 180
- scales, 94-106, 111, 127-133, 134-135
- scanning, 138, 142, 154
- scoring, 50, 62, 76, 79, 94-108, 127-134, 155, 170, 178-179
- search reading, 138, 142
- second language acquisition, 35
- self assessment, *see* assessment, self
- semi-direct testing, 19, 122, 135
- short-answer items, 79-80, 144-148, 166, 208
- sequencing items, 146-148, 159
- simulated conversation, 122-123
- size of response, 62, 127
- skills,
 - informational, 114-115, 161
 - interactional, 115, 161-162
 - in managing interactions, 115
 - sub-skills, 31-32, 60, 159
- skimming, 138
- Spearman-Brown formula, 44, 51, 223
- specifications, 26-27, 33, 59-61, 83-85, 113-116, 136-142, 159, 160-163, 173, 180
- speech, rate of, 116, 162
- speed of processing, 61
- split half estimate of reliability, 40, 223
- Sri Lanka, 57
- stages of test development, 58-74
- stakeholders, 5-6
- standard deviation, 221
- standard error of individual's score, 42, 232

- standard error of measurement (SEM),
40-41, 51, 223-224
- statistical analysis, 9-10, 218-233
- stem, 75
- structural range, 61
- structure of tests, 61
- style, 61, 85, 116, 140
- subjective testing, 22, 53
- syllabus content approach, 13
- synonym items, 180-181

- techniques, test (*see also* formats), 48, 61,
75-82, 119-124, 143-153,
165-169, 174-178, 180-184,
201-212
- TEEP (Test of English for Educational
Purposes), 73
- test-retest estimate of reliability, 39
- test wiseness, 159
- text form, 140, 162
 - length, 60, 85, 140, 142, 162
 - types, 60, 84, 85, 114, 115, 140, 162
- texts, selection of, 142-143, 163-164
- think-aloud, 32, 35
- timing, 61
- TOEFL (Test of English as a Foreign
Language), 6, 57, 80, 95, 171
- topics, 60, 84, 85, 89, 111, 116, 140
- training of staff, 49, 66, 106-107,
126-127, 133
- transcription, 169
- trialling, 63-64, 234, 235
- true/false items, 79, 144, 200
- true score, 40-41

- TSE (Test of Spoken English), 123-124,
135
- TWE (Test of Written English), 6, 95,
96-98, 111

- UCLA ESLP Test, 189
- UCLES (University of Cambridge Local
Examinations Syndicate), 24, 73
- unfamiliar words, predicting meaning of,
145
- unique answer items, 79-80, 144-145

- validation (*see also* validity), 66
- validity, 26-35
 - and reliability, 50
 - coefficient, 28-29
 - concurrent, 27-29
 - construct, 26, 30-32, 35, 76
 - content, 26-27, 85, 141, 142, 172, 174
 - criterion-related, 27-30
 - face, 33, 35
 - in scoring (*see also* scoring), 32-33
 - predictive, 29-30
 - and reliability, 50
- vocabulary range, 61, 116
 - testing of, 179-184

- washback, 57
- weighting, 103
- writing, testing of, 83-112

- yes/no items, 79, 214
- young learners, 199-214

Author Index

- Abraham, R. G., 197
AERA, 73
Alderson, J. C., 7, 10, 25, 35, 57, 73, 159
Allan, A., 159
Allison, D., 162
Anastasi, A. 34, 51
Anderson, J., 100, 103
Arnold, J., 171

Bachman, L. F., 35, 73, 135
Bailey, K. M., 57
Baker, R., 233
Berwick, R., 35
Bradshaw, J., 35
Brennan, R. L., 51
Brown, J. D., 7, 24, 51, 57, 197
Buck, G., 10, 35, 73, 171
Bygate, M., 114

Cameron, L., 213
Canale, M., 25
Carpenter, K., 213
Carroll, J. B., 10, 25
Chalhoub-Deville, M., 25, 134
Chapelle, C., 185, 197
Clapham, C., 7
Cohen, A. D., 35
Collins Cobuild 16, 185
Conrad, C. A., 188, 197
Council of Europe, 24, 111
Courtney, R., 171
Criper, C., 35
Crystal, D., 171
Cumming, A., 35

Davidson, F., 73
Davies, A., 2, 25, 35
Davy, D., 171
DeVicenzi, F., 7
Deville, C., 25
Dexter, Colin, 175
Douglas, D., 135

Edenelbos, P. 214
Ebel, R. L., 24

Farhady, H., 197
Feldt, L. S., 51
Fletcher, P., 52, 233
Freedle, R., 159, 171
Fulcher, G., 24, 35, 134

Gardner, S., 7
Garman, M., 198
Gipps, C., 7
Godshalk, F. I., 24, 83, 112
Greenberg, K., 111

Hale, G. A., 171
Hamp-Lyons, L., 6, 57, 111
Hannah, J., 149
Harris, D. P., 95, 100
Hasselgren, A., 214
Heaton, J. B., 82
Hicks, D., 212
Hilke, R., 57
Hudson, T., 7, 24, 51, 57
Hughes, A., 7, 24, 25, 30, 35, 52, 57,
111, 171, 198, 233

- Inbar, O., 171
- Jacobs, H. L., 103, 111
- Jafarpur, A., 198
- James, C., 112
- Jennings, M., 24, 111
- Keramati, M. N., 197
- Krahnke, K. J., 10
- Klein-Braley, C., 194, 198
- Kormos, J., 135
- Kostin, I., 159, 171
- Kunnan, A. J., 6
- Lazaraton, A., 135
- Lado, R., 18, 39, 51, 198
- Lee, J. F., 159
- Leech, G., 185
- Leung, C., 214
- Lewkowicz, J. A., 24
- Limb, S., 146-148
- Littlejohn, A., 212
- Lowe, P., 135
- Lumley, T., 135, 159
- Luoma, S., 134, 135
- Lynch, B., 24
- MacWhinney, B., 10
- McLaughlin, B., 10
- McNamara, T., 135, 233
- Messick, S., 34-35, 57
- Mislevy, R. J., 10
- Morrow, K., 25, 35
- Nitko, A. J., 7, 34, 51, 52
- North, B., 111
- Oller, J. W., 25, 188, 195, 196, 197
- Palmer, A. S., 35, 73
- Pilliner, A., 24
- Pimsleur, P., 10
- Popham, W. J., 24
- Porter, D., 25, 35, 159
- Powers, D. E., 135
- Raatz, U., 194, 198
- Raven, J., 7
- Read, J., 185
- Rea-Dickens, P., 6, 7, 213, 214
- Riley, G. L., 159
- Rixon, S., 213
- Ross, S., 7
- Salaberry, R., 135
- Savignon, S. J., 135
- Schneider, G., 111
- Scott, M. L., 24
- Sherman, J., 171
- Shohamy, E., 57, 135, 159, 171
- Skehan, P., 10, 24
- Spolsky, B., 6, 10
- Stansfield, C. W., 10
- Sternberg, R. J., 10
- Storey, P., 35, 198
- Swain, M., 25
- Swan, M., 156, 158
- Tatsuoka, K., 171
- Tauroza, S., 162
- Teasdale, A., 214
- Trim, J. L. M., 73, 185
- Trudgill, P., 149, 171
- Underhill, N., 134
- Urbina, S., 34, 51
- Urquhart, A., 159
- van Ek, J. A., 73, 185
- Vinje, M. P., 214
- Wadden, P., 57
- Wall, D., 24, 35, 57
- Walter, C., 158
- Watanabe, Y., 57
- Weigle, S. C., 111
- Weir, C. J., 25, 35, 73, 159
- West, M., 185
- Wigglesworth, G., 135
- Woods, A., 52, 233
- Wu, Y., 35