LANGUAGE ASSESSMENT

Principles and Classroom Practices

H. DOUGLAS BROWN
San Francisco State University
CONTENTS

Preface ix
Text Credits xii

1 Testing, Assessing, and Teaching 1

What Is a Test?, 3
Assessment and Teaching, 4
Informal and Formal Assessment, 5
Formative and Summative Assessment, 6
Norm-Referenced and Criterion-Referenced Tests, 7

Approaches to Language Testing: A Brief History, 7
Discrete-Point and Integrative Testing, 8
Communicative Language Testing, 10
Performance-Based Assessment, 10

Current Issues in Classroom Testing, 11
New Views on Intelligence, 11
Traditional and “Alternative” Assessment, 13
Computer-Based Testing, 14

Exercises, 16

For Your Further Reading, 18

2 Principles of Language Assessment 19

Practicality, 19
Reliability, 20
Student-Related Reliability, 21
Rater Reliability, 21
Test Administration Reliability, 21
Test Reliability, 22

Validity, 22
Content-Related Evidence, 22
Criterion-Related Evidence, 24
Construct-Related Evidence, 25
Consequential Validity, 26
Face Validity, 26
Authenticity, 28
Washback, 28
Applying Principles to the Evaluation of Classroom Tests, 30
1. Are the test procedures practical? 31
2. Is the test reliable? 31
3. Does the procedure demonstrate content validity? 32
4. Is the procedure face valid and "biased for best"? 33
5. Are the test tasks as authentic as possible? 35
6. Does the test offer beneficial washback to the learner? 37

Exercises, 38
For Your Further Reading, 41

3 Designing Classroom Language Tests

Test Types, 43
Language Aptitude Tests, 43
Proficiency Tests, 44
Placement Tests, 45
Diagnostic Tests, 46
Achievement Tests, 47

Some Practical Steps to Test Construction, 48
Assessing Clear, Unambiguous Objectives, 49
Drawing Up Test Specifications, 50
Devising Test Tasks, 52
Designing Multiple-Choice Test Items, 55
1. Design each item to measure a specific objective, 56
2. State both stem and options as simply and directly as possible, 57
3. Make certain that the intended answer is clearly the only correct one, 58
4. Use item indices to accept, discard, or revise items, 58

Scoring, Grading, and Giving Feedback, 61
Scoring, 61
Grading, 62
Giving Feedback, 62
Exercises, 64
For Your Further Reading, 65

4 Standardized Testing

What Is Standardization?, 67
Advantages and Disadvantages of Standardized Tests, 68
Developing a Standardized Test, 69
1. Determine the purpose and objectives of the test, 70
2. Design test specifications, 70
3. Design, select, and arrange test tasks/items, 74
4. Make appropriate evaluations of different kinds of items, 78
5. Specify scoring procedures and reporting formats, 79
6. Perform ongoing construct validation studies, 81
Standardized Language Proficiency Testing, 82
Four Standardized Language Proficiency Tests, 83
   Test of English as a Foreign Language (TOEFL®), 84
   Michigan English Language Assessment Battery (MELAB), 85
   International English Language Testing System (IELTS), 85
   Test of English for International Communication (TOEIC®), 86
Exercises, 87
For Your Further Reading, 87
Appendix to Chapter 4:
Commercial Proficiency Tests: Sample Items and Tasks, 88
   Test of English as a Foreign Language (TOEFL®), 88
   Michigan English Language Assessment Battery (MELAB), 93
   International English Language Testing System (IELTS), 96
   Test of English for International Communication (TOEIC®), 100

5 Standards-Based Assessment 104
ELD Standards, 105
ELD Assessment, 106
CASAS and SCANS, 108
Teacher Standards, 109
The Consequences of Standards-Based and Standardized Testing, 110
   Test Bias, 111
   Test-Driven Learning and Teaching, 112
Ethical Issues: Critical Language Testing, 113
Exercises, 115
For Your Further Reading, 115

6 Assessing Listening 116
Observing the Performance of the Four Skills, 117
The Importance of Listening, 119
Basic Types of Listening, 119
Micro- and Macroskills of Listening, 121
Designing Assessment Tasks: Intensive Listening, 122
   Recognizing Phonological and Morphological Elements, 123
   Paraphrase Recognition, 124
Designing Assessment Tasks: Responsive Listening, 125
Designing Assessment Tasks: Selective Listening, 125
   Listening Cloze, 125
   Information Transfer, 127
   Sentence Repetition, 130
Designing Assessment Tasks: Extensive Listening, 130
CONTENTS

Dictation, 131
Communicative Stimulus-Response Tasks, 132
Authentic Listening Tasks, 135

Exercises, 138
For Your Further Reading, 139

7 Assessing Speaking 140

Basic Types of Speaking, 141
Micro- and Macroskills of Speaking, 142
Designing Assessment Tasks: Imitative Speaking, 144
PhonePass® Test, 145
Designing Assessment Tasks: Intensive Speaking, 147
Directed Response Tasks, 147
Read-Aloud Tasks, 147
Sentence/Dialogue Completion Tasks and Oral Questionnaires, 149
Picture-Cued Tasks, 151
Translation (of Limited Stretches of Discourse), 159
Designing Assessment Tasks: Responsive Speaking, 159
Question and Answer, 159
Giving Instructions and Directions, 161
Paraphrasing, 161
Test of Spoken English (TSE®), 162
Designing Assessment Tasks: Interactive Speaking, 167
Interview, 167
Role Play, 174
Discussions and Conversations, 175
Games, 175
Oral Proficiency Interview (OPI), 176
Designing Assessment: Extensive Speaking, 179
Oral Presentations, 179
Picture-Cued Story-Telling, 180
Retelling a Story, News Event, 182
Translation (of Extended Prose), 182

Exercises, 183
For Your Further Reading, 184

8 Assessing Reading 185

Types (Genres) of Reading, 186
Microskills, Macroskills, and Strategies for Reading, 187
Types of Reading, 189
Designing Assessment Tasks: Perceptive Reading, 190
Reading Aloud, 190
Written Response, 191
CONTENTS

Multiple-Choice, 191
Picture-Cued Items, 191
Designing Assessment Tasks: Selective Reading, 194
  Multiple-Choice (for Form-Focused Criteria), 194
  Matching Tasks, 197
  Editing Tasks, 198
  Picture-Cued Tasks, 199
  Gap-Filling Tasks, 200
Designing Assessment Tasks: Interactive Reading, 201
  Cloze Tasks, 201
  Impromptu Reading Plus Comprehension Questions, 204
  Short-Answer Tasks, 206
  Editing (Longer Texts), 207
  Scanning, 209
  Ordering Tasks, 209
  Information Transfer: Reading Charts, Maps, Graphs, Diagrams, 210
Designing Assessment Tasks: Extensive Reading, 212
  Skimming Tasks, 213
  Summarizing and Responding, 213
  Note-Taking and Outlining, 215

Exercises, 216
For Your Further Reading, 217

9 Assessing Writing 218

Genres of Written Language, 219
Types of Writing Performance, 220
Micro- and Macroskills of Writing, 220
Designing Assessment Tasks: Imitative Writing, 221
  Tasks in [Hand] Writing Letters, Words, and Punctuation, 221
  Spelling Tasks and Detecting Phoneme-Grapheme Correspondences, 223
Designing Assessment Tasks: Intensive (Controlled) Writing, 225
  Dictation and Dicto-Comp, 225
  Grammatical Transformation Tasks, 226
  Picture-Cued Tasks, 226
  Vocabulary Assessment Tasks, 229
  Ordering Tasks, 230
  Short-Answer and Sentence Completion Tasks, 230
Issues in Assessing Responsive and Extensive Writing, 231
Designing Assessment Tasks: Responsive and Extensive Writing, 233
  Paraphrasing, 234
  Guided Question and Answer, 234
  Paragraph Construction Tasks, 235
  Strategic Options, 236
Test of Written English (TWE®), 237
Scoring Methods for Responsive and Extensive Writing, 241
  Holistic Scoring, 242
  Primary Trait Scoring, 242
  Analytic Scoring, 243
Beyond Scoring: Responding to Extensive Writing, 246
  Assessing Initial Stages of the Process of Composing, 247
  Assessing Later Stages of the Process of Composing, 247
Exercises, 249
For Your Further Reading, 250

10 Beyond Tests: Alternatives in Assessment 251
The Dilemma of Maximizing Both Practicality and Washback, 252
Performance-Based Assessment, 254
Portfolios, 256
Journals, 260
Conferences and Interviews, 264
Observations, 266
Self- and Peer-Assessments, 270
  Types of Self- and Peer-Assessment, 271
  Guidelines for Self- and Peer-Assessment, 276
  A Taxonomy of Self- and Peer-Assessment Tasks, 277
Exercises, 279
For Your Further Reading, 280

11 Grading and Student Evaluation 281
Philosophy of Grading: What Should Grades Reflect? 282
  Guidelines for Selecting Grading Criteria, 284
  Calculating Grades: Absolute and Relative Grading, 285
  Teachers' Perceptions of Appropriate Grade Distributions, 289
Institutional Expectations and Constraints, 291
  Cross-Cultural Factors and the Question of Difficulty, 292
  What Do Letter Grades “Mean”? , 293
Alternatives to Letter Grading, 294
Some Principles and Guidelines for Grading and Evaluation, 299
Exercises, 300
For Your Further Reading, 302

Bibliography 303
Name Index 313
Subject Index 315