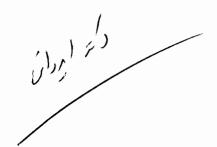
Fundamental Considerations in Language Testing

Lyle F. Bachman



Acknowledgments

The author and publishers would like to thank the following for permission to reproduce the material below that falls within their copyright:

The American Council on Education for the extracts from R. L. Thorndike (ed.): *Educational Measurement* (Second edition) and R. L. Linn (ed.): *Educational Measurement* (Thirdedition)

The American Psychological Association for the extract from the paper by S. A. Messick in *American Psychologist* 30.

Brooks/Cole Publishing Company for the table on page 30, adapted from M. J. Allen and W. M. Yen: *Introduction to Measurement Theory*

The Center for Applied Linguistics for the extract from the pager by J. O. Carroll in *Testing the English Proficiency of Foreign Students*

Educational and Psychological Measurement for the extracts from the paper by C. I. Mosier in Volume 7

Jay Haley for the extract from Strategies of Psychotherapy

The John Hopkins University Press for the figure on page 341, adapted from A. J. Nitko, 'Defining the criterion-referenced test' in R. A. Berk (ed.): A *Guide to Criterion-Referenced Test Construction*, and for the extract from the same paper

Language Learning, and the authors, for the table on page 196, from the paper by J. van Weeren and T. J. J. Theunissen in Volume 37

Newbury House Publishers for three figures (from J. A. Upshur, 'Context for language testing') in J. W. Oller and J. C. Richards: *Focus* on *the Learner*

Pergamon Press PLC for the extract from K. Johnson (ed.): Communicative Syllabus Design and Methodology

Psychometrika, and the authors, for the extract from the paper by K. K. Tatsuoka and M. M. Tatsuoka in Volume 52

Contents

Preface		VIII
1	Introduction The aims of the book The climate for language testing Research and development: needs and problems Research and development: an agenda Overview of the book Notes	1 2 8 12 13 15
2	Measurement inProduction Definition of terms: measurement, test, evaluation Essential measurement qualities Properties of measurement scales Characteristics that limit measurement Steps in measurement Summary Notes Further reading Discussion questions	18 18 24 26 30 40 49 50 52
3	Uses of Language Tests Introduction Uses of language tests in educational programs Research uses of language tests Features for classifying different types of language test Summary Further reading Discussion questions	53 53 67 70 78 79
4	Communicative Language Ability Introduction	81

vi Contents

	A theoretical framework of communicative language	04
	ability	84 107
	Summary Notes	107
	Further reading	109
	Discussion questions	109
5	Test Methods	
	Introduction	111
	A framework of test method facets	116
	Applications of this framework to language resting	152
	Summary	156
	Notes	157
	Further reading	158
	Discussion questions	159
6	Reliability	
	Introduction	160
	Factors that affect language test scores	163
	Classical true score measurement theory	166
	Generalizability theory	187
	Standard error of measurement: interpreting individual test scores within classical true score and generalizability	197
	theory Item response theory	202
	Reliability of criterion-referenced test scores	202
	Factors that affect reliability estimates	220
	Systematic measurement error	222
	Summary	226
	Notes	227
	Further reading	232
	Discussion questions	233
7		
	Introduction	236
	Reliability and validity revisited	238
	Validity as a unitary concept	24 1
	The evidential basis of validity	243
	Test bias	271
	The consequential or ethicai basis of validity	279
	Postmortem: face validity	285

Contents vii

	Summary	289
	Notes	<i>29</i> 1
	Further reading	294
	Discussion questions	294
8	Some Persistent Problems and Future Directions	20.6
	Introduction	296
	Authentic language tests	3Q0
	Some future directions	333
	A general model for explaining performance on language tests	348
	Apologia et prolegomenon	351
	Summary	<i>356</i>
	Notes	358
	Further reading	358
	Discussion questions	358
B	Bibliography	
Δ.	uthor index	395
Subject index		397