

For Jenny and Robin

سرشناسه:	فالتجر، گلن
عنوان و نام پدیدآور:	Fulcher, Glenn Language Testing and Assessment: An Advanced Resource Book / Glenn Fulcher and Fred Davidson.
مشخصات نشر:	تهران: رهنما، ۱۳۹۱ = ۲۰۱۲ م.
مشخصات ظاهری:	۱۲۴ ص: مصور، جدول، نمودار.
وضعیت فهرست نویسی:	فیب، انگلیسی.
یادداشت:	قیمت از روی چاپ ۲۰۰۷: لندن.
یادداشت:	آوانویس عنوان: لنگواج تستینگ اند...
موضوع:	زبان - آزمون های توانایی
شناسه افزوده:	دیویدسن، فرد، ۱۹۵۴ - م.
شناسه افزوده:	Davidson, Fred
ردمبندی کنگره:	P۴۱/۳۷:۵۹۱ ۱۳۹۱
ردمبندی دیویی:	۴۱۸/۰۰۷۶
شماره کتابشناسی ملی:	۲۷۷۲۴۷۳

Language Testing and Assessment، مؤلفین: Glenn Fulcher & Fred Davidson، لیوگرافی: رهنما، چاپ: چاپخانه سلمان فارسی، چاپ اول: ۱۳۹۱، تیراژ: ۱۵۰۰ نسخه، نشر: انتشارات رهنما، مقابل دانشگاه تهران، خیابان فروردین، نیش خیابان شهیدای ژاندارمری، پلاک ۱۱۲، تلفن: ۶۶۴۰۰۹۲۷ - ۶۶۴۱۶۶۰۴، ۶۶۴۸۱۶۶۲، فاکس: ۶۶۴۶۷۴۳۴، فروشگاه رهنما، سادات آباد، خیابان علامه طباطبائی جنوبی، بین ۴۰ و ۴۲ شرقی، پلاک ۲۹، تلفن: ۸۸۶۹۴۱۰۲، آدرس فروشگاه شماره ۴، خیابان پیروزی نیش خیابان سوم نیروی هوایی، تلفن: ۷۷۴۸۲۵۰۵، نمایشگاه کتاب رهنما، مقابل دانشگاه تهران پلاز فرورنده، تلفن: ۶۶۹۵۰۹۵۷

قیمت: ۸۰۰۰۰ ریال

Series editors' preface

This series provides a comprehensive guide to a number of key areas in the field of applied linguistics. Applied linguistics is a rich, vibrant, diverse and essentially interdisciplinary field. It is now more important than ever that books in the field provide up-to-date maps of ever-changing territory.

The books in this series are designed to give key insights into core areas. The design of the books ensures, through key readings, that the history and development of a subject is recognized while, through key questions and tasks, integrating understandings of the topics, concepts and practices that make up its essentially interdisciplinary fabric. The pedagogic structure of each book ensures that readers are given opportunities to think, discuss, engage in tasks, draw on their own experience, reflect, research and to read and critically re-read key documents.

Each book has three main sections, each made up of approximately ten units:

A: An Introduction section: in which the key terms and concepts are introduced, including introductory activities and reflective tasks, designed to establish key understandings, terminology, techniques of analysis and the skills appropriate to the theme and the discipline.

B: An Extension section: in which selected core readings are introduced (usually edited from the original) from existing books and articles, together with annotations and commentary, where appropriate. Each reading is introduced, annotated and commented on in the context of the whole book, and research/follow-up questions and tasks are added to enable fuller understanding of both theory and practice. In some cases, readings are short and synoptic and incorporated within a more general exposition.

C: An Exploration section: in which further samples and illustrative materials are provided with an emphasis, where appropriate, on more open-ended, student-centred activities and tasks, designed to support readers and users in undertaking their own locally relevant research projects. Tasks are designed for work in groups or for individuals working on their own.

The books also contain a glossary or glossarial index and a detailed, thematically organized A-Z guide to the main terms used in the book, which lays the ground for

further work in the discipline. There are also annotated guides to further reading and extensive bibliographies.

The target audience for the series is upper undergraduates and postgraduates on language, applied linguistics and communication studies programmes as well as teachers and researchers in professional development and distance learning programmes. High-quality applied research resources are also much needed for teachers of EFL/ESL and foreign language students at higher education colleges and universities worldwide. The books in the Routledge Applied Linguistics series are aimed at the individual reader, the student in a group and at teachers building courses and seminar programmes.

We hope that the books in this series meet these needs and continue to provide support over many years.

The Editors

Professor Christopher N. Candlin and Professor Ronald Carter are the series editors. Both have extensive experience of publishing titles in the fields relevant to this series. Between them they have written and edited over one hundred books and two hundred academic papers in the broad field of applied linguistics. Chris Candlin was president of AILA (International Association for Applied Linguistics) from 1997 to 2003 and Ron Carter is Chair of BAAL (British Association for Applied Linguistics) from 2003 to 2006.

Professor Christopher N. Candlin
Senior Research Professor
Department of Linguistics
Division of Linguistics and Psychology
Macquarie University
Sydney NSW 2109
Australia

and

Professor of Applied Linguistics
Faculty of Education and Language Studies
The Open University
Walton Hall
Milton Keynes MK7 6AA
UK

Professor Ronald Carter
School of English Studies
University of Nottingham
Nottingham NG7 2RD
UK

A4.9 Towards spec-driven theory	60
Summary	61
Unit A5 Writing items and tasks	62
A5.1 Introduction	62
A5.2 Evidence-centred design (ECD)	63
A5.3 Describing items and tasks	69
A5.4 Tasks and teaching	73
Summary	75
Unit A6 Prototypes, prototyping and field tests	76
A6.1 Introduction	76
A6.2 Prototypes	76
A6.3 Prototyping	79
A6.4 Field testing	85
A6.5 The iterative nature of the process	89
Summary	89
Unit A7 Scoring language tests and assessments	91
A7.1 Introduction	91
A7.2 Defining the quality of language	93
A7.3 Developing scoring systems	96
A7.4 Intuition and data	98
A7.5 Problems with scales	98
A7.6 Scoring in classical test theory	101
A7.7 Reliability	104
A7.8 Score transformations	108
A7.9 Item response theory	109
A7.10 Endowing a score with special meaning	111
Summary	114
Unit A8 Administration and training	115
A8.1 Introduction	115
A8.2 Getting things done	117
A8.3 Quality management systems	127
A8.4 Constraints	128
A8.5 Test administration within the ECD delivery model	129
A8.6 Rater and interlocutor training	131
A8.7 Security	132
A8.8 Test administration for disabled people	135
Summary	137
Unit A9 Fairness, ethics and standards	138
A9.1 Introduction	138
A9.2 Professionalism as a community of practitioners	138
A9.3 Professionalism and democracy	141
A9.4 Consequentialism	142
A9.5 On power and pessimism	144
A9.6 Professional conduct: standards for practice	155
A9.7 Responsibilities of language testers and their limitations	156
A9.8 Accountability	157
Summary	158

Contents

List of figures and tables	xiv
Series editors' preface	xv
Acknowledgments	xvii
How to use this book	xix

SECTION A: INTRODUCTION

Unit A1 Introducing validity	3
A1.1 Introduction	3
A1.2 Three 'types' of validity in early theory	4
A1.3 Cutting the validity cake	12
Summary	21
Unit A2 Classroom assessment	23
A2.1 Introduction	23
A2.2 Pedagogy and the measurement paradigm	25
Summary	35
Unit A3 Constructs and models	36
A3.1 Introduction	36
A3.2 The nature of models	37
A3.3 Canale and Swain's model of communicative competence	38
A3.4 Canale's adaptations	39
A3.5 Bachman's model of communicative language ability (CLA)	42
A3.6 Celce-Murcia, Dörnyei and Thurrell's model of communicative competence	47
A3.7 Interactional competence	49
A3.8 From models to frameworks: validity models and performance conditions	50
Summary	51
Unit A4 Test specifications and designs	52
A4.1 Introduction	52
A4.2 Planning in test authoring	53
A4.3 Guiding language versus samples	54
A4.4 Congruence (or fit-to-spec)	55
A4.5 How do test questions originate? Reverse engineering and archetypes	56
A4.6 Reverse engineering	57
A4.7 Where do test items come from? What is the true genesis of a test question?	58
A4.8 Spec-driven test assembly, operation and maintenance	59

LANGUAGE TESTING AND ASSESSMENT

Routledge Applied Linguistics is a series of comprehensive resource books, providing students and researchers with the support they need for advanced study in the core areas of English language and Applied Linguistics.

Each book in the series guides readers through three main sections, enabling them to explore and develop major themes within the discipline.

- Section A, Introduction, establishes the key terms and concepts and extends readers' techniques of analysis through practical application.
- Section B, Extension, brings together influential articles, sets them in context and discusses their contribution to the field.
- Section C, Exploration, builds on knowledge gained in the first two sections, setting thoughtful tasks around further illustrative material. This enables readers to engage more actively with the subject matter and encourages them to develop their own research responses.

Throughout the book, topics are revisited, extended, interwoven and deconstructed, with the reader's understanding strengthened by tasks and follow-up questions.

Language Testing and Assessment:

- provides an innovative and thorough review of a wide variety of issues from practical details of test development to matters of controversy and ethical practice
- investigates the importance of the philosophy of pragmatism in assessment, and coins the term 'effect-driven testing'
- explores test development, data analysis, validity and their relation to test effects
- illustrates its thematic breadth in a series of exercises and tasks, such as analysis of test results, study of test revision and change, design of arguments for test validation and exploration of influences on test creation
- presents influential and seminal readings in testing and assessment by names such as Michael Canale and Merrill Swain, Michael Kane, Alan Davies, Lee Cronbach and Paul Meehl and Pamela Moss.

Written by experienced teachers and researchers in the field, *Language Testing and Assessment* is an essential resource for students and researchers of Applied Linguistics.

Glenn Fulcher is Senior Lecturer in the School of Education at the University of Leicester, UK.

Fred Davidson is Associate Professor in the Division of English as an International Language at the University of Illinois at Urbana-Champaign, USA.

ROUTLEDGE APPLIED LINGUISTICS

SERIES EDITORS

Christopher N. Candlin is Senior Research Professor in the Department of Linguistics at Macquarie University, Australia, and Professor of Applied Linguistics at the Open University, UK. At Macquarie, he has been Chair of the Department of Linguistics; he established and was Executive Director of the National Centre for English Language Teaching and Research (NCELTR) and foundational Director of the Centre for Language in Social Life (CLSL). He has written or edited over 150 publications and co-edits the *Journal of Applied Linguistics*. From 1996 to 2002 he was President of the International Association of Applied Linguistics (AILA). He has acted as a consultant in more than thirty-five countries and as external faculty assessor in thirty-six universities worldwide.

Ronald Carter is Professor of Modern English Language in the School of English Studies at the University of Nottingham. He has published extensively in applied linguistics, literary studies and language in education, and has written or edited over forty books and a hundred articles in these fields. He has given consultancies in the field of English language education, mainly in conjunction with the British Council, in over thirty countries worldwide, and is editor of the Routledge Interface series and advisory editor to the Routledge English Language Introduction series. He was recently elected a fellow of the British Academy of Social Sciences and is currently UK Government Advisor for ESOL and Chair of the British Association of Applied Linguistics (BAAL).

TITLES IN THE SERIES

Intercultural Communication: An advanced resource book
Adrian Holliday, Martin Hyde and John Kullman

Translation: An advanced resource book
Basil Hatim and Jeremy Munday

Grammar and Context: An advanced resource book
Ann Hewings and Martin Hewings

Second Language Acquisition: An advanced resource book
Kees de Bot, Wander Lowie and Marjolijn Verspoor

Corpus-based Language Studies: An advanced resource book
Anthony McNery, Richard Xiao and Yukio Tono

Language and Gender: An advanced resource book
Jane Sunderland

English for Academic Purposes: An advanced resource book
Ken Hyland

Language Testing and Assessment: An advanced resource book
Glenn Fulcher and Fred Davidson

Unit A10 Arguments and evidence in test validation and use	159
A10.1 Introduction	159
A10.2 Argumentation as solution	162
A10.3 The form of an argument	164
A10.4 Argument in evidence-centred design	167
A10.5 Arguments in language testing	168
A10.6 Arguments and feasibility	176
A10.7 Argument, evidence and ethics	176
Summary	178
 SECTION B: EXTENSION	 179
Unit B1 Construct validity	181
Cronbach, L. J. and Meehl, P. E. 'Construct validity in psychological tests'	182
Unit B2 Pedagogic assessment	192
Moss, P. 'Reconceptualizing validity for classroom assessment'	193
Unit B3 Investigating communicative competence	203
Canale, M. and Swain, M. 'Theoretical bases of communicative approaches to second language teaching and testing'	203
Unit B4 Optimal specification design	212
Davidson, F. and Lynch, B. K. <i>Testcraft: A Teacher's Guide to Writing and Using Language Test Specifications</i>	212
Unit B5 Washback	221
Alderson, J. C. and Wall, D. 'Does washback exist?'	222
Unit B6 Researching prototype tasks	230
Cumming, A., Grant, L., Mulcahy-Ernt, P. and Powers, D. <i>A Teacher-Verification Study of Speaking and Writing Prototype Tasks for a New TOEFL</i>	230
Unit B7 Scoring performance tests	249
Hamp-Lyons, L. 'Scoring procedures for ESL contexts'	250
Unit B8 Interlocutor training and behaviour	258
Brown, A. 'Interviewer variation and the co-construction of speaking proficiency'	260
Unit B9 Ethics and professionalism	270
Davies, A. 'Demands of being professional in language testing'	270
Unit B10 Validity as argument	278
Kane, M. T. (1992). 'An argument-based approach to validity'	278

SECTION C: EXPLORATION	291
Unit C1 Validity – an exploration	293
Unit C2 Assessment in school systems	298
Unit C3 What do items really test?	304
Unit C4 Evolution in action	312
Unit C5 To see a test in a grain of sand . . .	320
Unit C6 Analysing items and tasks	326
Unit C7 Designing an alternative matrix	333
Unit C8 Administration and alignment	343
Unit C9 In a time far, far away . . .	352
Unit C10 To boldly go	361
Glossary	369
Notes	378
References	379
Index	396

www.ketab.ir

Contents cross-referenced

Section A: Introduction			
Unit A1	Introducing validity	3	A6.3 Prototyping 79
A1.1	Introduction	3	A6.4 Field testing 85
A1.2	Three 'types' of validity in early theory	4	A6.5 The iterative nature of the process 89
A1.3	Cutting the validity cake	12	Summary 89
	Summary	21	
Unit A2	Classroom assessment	23	Unit A7 Scoring language tests and assessments 91
A2.1	Introduction	23	A7.1 Introduction 91
A2.2	Pedagogy and the measurement paradigm	25	A7.2 Defining the quality of language 93
	Summary	35	A7.3 Developing scoring systems 96
Unit A3	Constructs and models	36	A7.4 Intuition and data 98
A3.1	Introduction	36	A7.5 Problems with scales 98
A3.2	The nature of models	37	A7.6 Scoring in classical test theory 101
A3.3	Canale and Swain's model of communicative competence	38	A7.7 Reliability 104
A3.4	Canale's adaptations	39	A7.8 Score transformations 108
A3.5	Bachman's model of communicative language ability (CLA)	42	A7.9 Item response theory 109
A3.6	Celce-Murcia, Dörnyei and Thurrell's model of communicative competence	47	A7.10 Endowing a score with special meaning 111
A3.7	Interactional competence	49	Summary 114
A3.8	From models to frameworks: validity models and performance conditions	50	Unit A8 Administration and training 115
	Summary	51	A8.1 Introduction 115
Unit A4	Test specifications and designs	52	A8.2 Getting things done 117
A4.1	Introduction	52	A8.3 Quality management systems 127
A4.2	Planning in test authoring	53	A8.4 Constraints 128
A4.3	Guiding language versus samples	54	A8.5 Test administration within the ECD delivery model 129
A4.4	Congruence (or fit-to-spec)	55	A8.6 Rater and interlocutor training 131
A4.5	How do test questions originate? Reverse engineering and archetypes	56	A8.7 Security 132
A4.6	Reverse engineering	57	A8.8 Test administration for disabled people 135
A4.7	Where do test items come from? What is the true genesis of a test question?	58	Summary 137
A4.8	Spec-driven test assembly, operation and maintenance	59	Unit A9 Fairness, ethics and standards 138
A4.9	Towards spec-driven theory	60	A9.1 Introduction 138
	Summary	61	A9.2 Professionalism as a community of practitioners 138
Unit A5	Writing items and tasks	62	A9.3 Professionalism and democracy 141
A5.1	Introduction	62	A9.4 Consequentialism 142
A5.2	Evidence-centred design (ECD)	63	A9.5 On power and pessimism 144
A5.3	Describing items and tasks	69	A9.6 Professional conduct: standards for practice 155
A5.4	Tasks and teaching	73	A9.7 Responsibilities of language testers and their limitations 156
	Summary	75	A9.8 Accountability 157
Unit A6	Prototypes, prototyping and field tests	76	Summary 158
A6.1	Introduction	76	Unit A10 Arguments and evidence in test validation and use 159
A6.2	Prototypes	76	A10.1 Introduction 159
			A10.2 Argumentation as solution 162
			A10.3 The form of an argument 164
			A10.4 Argument in evidence-centred design 167
			A10.5 Arguments in language testing 168
			A10.6 Arguments and feasibility 176
			A10.7 Argument, evidence and ethics 176
			Summary 178

Section B: Extension

Unit B1	Construct validity	181
	Cronbach, L. J. and Meehl, P. E. 'Construct validity in psychological tests'	182
Unit B2	Pedagogic assessment	192
	Moss, P. 'Reconceptualizing validity for classroom assessment'	193
Unit B3	Investigating communicative competence	203
	Canale, M. and Swain, M. 'Theoretical bases of communicative approaches to second language teaching and testing'	203
Unit B4	Optimal specification design	212
	Davidson, F. and Lynch, B. K. <i>Testcraft: A Teacher's Guide to Writing and Using Language Test Specifications</i>	212
Unit B5	Washback	221
	Alderson, J. C. and Wall, D. 'Does Washback Exist?'	222
Unit B6	Researching prototype tasks	230
	Cumming, A., Grant, L., Mulcahy-Ernt, P. and Powers, D. <i>A Teacher Verification Study of Speaking and Writing Prototype Tasks for a New TOEFL</i>	230
Unit B7	Scoring performance tests	249
	Hamp-Lyons, L. 'Scoring procedures for ESL contexts'	250
Unit B8	Interlocutor training and behaviour	258
	Brown, A. 'Interviewer variation and the co-construction of speaking proficiency'	260
Unit B9	Ethics and professionalism	270
	Davies, A. 'Demands of being professional in language testing'	270
Unit B10	Validity as argument	278
	Kane, M. T. (1992). 'An argument-based approach to validity'	278

Section C: Exploration

Unit C1	Validity – an exploration	293
Unit C2	Assessment in school systems	298
Unit C3	What do items really test?	304
Unit C4	Evolution in action	312
Unit C5	To see a test in a grain of sand ...	320
Unit C6	Analysing items and tasks	326
Unit C7	Designing an alternative matrix	333
Unit C8	Administration and alignment	343
Unit C9	In a time far, far away ...	352
Unit C10	To boldly go	361

Figures and tables

FIGURES

A3.1	Models, frameworks and specifications	37
A3.2	Canale's adaptation of the Canale and Swain model	41
A3.3	Components of communicative language ability in language use	42
A3.4	Components of language competence	43
A3.5	Some components of language use and language test performance	46
A3.6	The Celce-Murcia et al. model of communicative competence	47
A5.1	Models in the conceptual assessment framework of ECD	68
A7.1	The <i>what</i> and <i>how much/how good</i> of language testing	91
A7.2	The structure of the CEF	99
A7.3	Three possible cut scores	113
A8.1	Four-process delivery model	130
A10.1	The basic form of an argument	165
A10.2	The argument of Harry's nationality	166
A10.3	Mislevy's language testing example	167
A10.4	An argument structure for textual competence items	169
A10.5	The cloze argument	170
A10.6	A test use argument structure	173

TABLES

A1.1	Facets of validity	13
A6.1	Basic item statistics for five items included on a prototype test of textual competence	87
A7.1	Responses of twenty test takers to ten test items	101
A7.2	Item facility values	102
A7.3	Point biserial correlations	104
A7.4	Rasch difficulty estimates for ten items	110
A7.5	Ability estimates for twenty test takers	110
C6.1	Item statistics	326
C7.1	Ethelynn's term-final results	334
C7.2	Data from the alternative universe	336
C7.3	Data mining structure	339

Acknowledgments

We would like to thank the series editors, Chris Candlin and Ron Carter, for their timely feedback on drafts of this manuscript, and their constant encouragement and help. At every stage in the writing process their wisdom and insight have been an inspiration, and we could not have asked for better editors with whom to work. At Routledge we would also like to thank Louisa Semlyen and Nadia Seemungal for their help, advice and efficiency.

The picture of the Golden Gate Bridge in Unit A4 is taken from <http://www.sonic.net/~playland/ggsag.html>. This picture is by an unknown photographer and is copyright-free, but we wish to acknowledge the source.

The examples of rapid prototyping in manufacturing in Unit A6 are reproduced by kind permission of Laser Prototypes Europe Ltd. The authors wish to thank Mr Tom Walls for his cooperation and advice while writing this part of Unit A6 (Tom@laserproto.com).

We would like to thank Mr R. J. Hann, erstwhile Staff Sergeant of Delta Company Fourth Battalion (4 RAR/NZ (ANZAC)), for permission to publish the picture in Unit C3, taken circa 1970 and published at http://www.diggerhistory.info/pages-uniforms/slouch_hat-b.htm.

The PERT and Gantt Charts used in Unit C6 are reproduced with permission of Alan Levine of Maricopa Community College.

The extracts in Section B are reproduced by kind permission of the publishers, as follows.

Unit B2

Moss, P. A. (2003) 'Reconceptualizing validity for classroom assessment.' *Educational Measurement: Issues and Practice* 22, 4, 13–25. Reproduced by permission of Blackwell Publishing.

Unit B3

Canale, M. and Swain, M. (1980) 'Theoretical bases of communicative approaches to second language teaching and testing.' *Applied Linguistics* 1, 1, 1–47 by permission of Oxford University Press.

Unit B4

Davidson, F. and Lynch, B. K. (2002) Chapter 3: 'Problems and issues in specification writing.' In *Testcraft: A Teacher's Guide to Writing and Using Language Test Specifications*. Copyright © 2002 by Yale University. Reproduced by permission of Yale University Press.

Unit B5

Alderson, J. C. and Wall, D. (1993) 'Does washback exist?' *Applied Linguistics* 14, 2, 115–129 by permission of Oxford University Press.

Unit B6

TOEFL materials selected from TOEFL Monograph Series MS-26 – *A Teacher-Verification Study of Speaking and Writing Prototype Tasks for a New TOEFL*, Educational Testing Service 2005. Reprinted by permission of Educational Testing Service, the copyright owner. However, the test questions and any other testing information are provided in their entirety by Routledge. No endorsement of this publication by Educational Testing Service should be inferred.

Unit B7

Hamp-Lyons, L. (1993) 'Scoring procedures for ESL contexts.' In Hamp-Lyons, L. (ed.) *Assessing Second Language writing in Academic Contexts*. Norwood, NJ: Ablex, 241–276. Copyright © 1991 by Ablex Publishing Corporation. Reproduced with permission of Greenwood Publishing Group, Inc., Westport, CT.

Unit B8

Reproduced with permission from Brown, A. (2003) 'Interviewer variation and the co-construction of speaking proficiency.' *Language Testing* 20, 1, 1–25. Copyright © Sage Publications 2003 by permission of Sage Publications Ltd and the author.

Unit B9

Reproduced with permission from Davies, A. (1997) 'Demands of being professional in language testing.' *Language Testing* 14, 3, 328–339. Copyright © Sage Publications Ltd 1997, by permission of Sage Publications Ltd.

Unit B10

Kane, M. T. (1992) 'An argument-based approach to validity.' *Psychological Bulletin* 112, 527–535. Copyright © 1992 by the American Psychological Association. Reprinted with permission.

How to use this book

Testing and assessment are part of modern life. Schoolchildren around the world are constantly assessed, whether to monitor their educational progress, or for governments to evaluate the quality of school systems. Adults are tested to see if they are suitable for a job they have applied for, or if they have the skills necessary for promotion. Entrance to educational establishments, to professions and even to entire countries is sometimes controlled by tests. Tests play a fundamental and controversial role in allowing access to the limited resources and opportunities that our world provides. The importance of understanding *what we test, how we test* and the *impact* that the use of tests has on individuals and societies cannot be overstated. Testing is more than a technical activity; it is also an ethical enterprise.

The practice of language testing draws upon, and also contributes to, all disciplines within applied linguistics. However, there is something fundamentally different about language testing. Language testing is all about building better tests, researching how to build better tests and, in so doing, understanding better the things that we test.

Sociolinguists do not create 'sociolinguistic things'. Discourse analysts do not create discourses. Phonologists do not create spoken utterances. Language testing, in contrast, is about *doing*. It is about *creating tests*.

In a sense, therefore, each section of this book is about the practical aspects of *doing* and of *creating*. And so each section has a research implication; no section is concerned purely with exposition. Research ideas may be made explicit in the third section, *Exploration*, but they are implicit throughout the book; put another way, the creative drive of language testing makes it a research enterprise, we think, at all times.

In the text we do not merely reflect the state of the art in language testing and assessment; nor do we simply introduce existing research. Our discussion is set within a new approach that we believe brings together testing practice, theory, ethics and philosophy. At the heart of our new approach is the concept of *effect-driven testing*. This is a view of test validity that is highly pragmatic. Our emphasis is on the outcome of testing activities. Our concern with test effect informs the order and structure of chapters, and it defines our approach to test design and development.

As test design and development is about *doing, creating* and *researching*, we have taken special care over the activities. With Dewey, we believe that through *doing* we grow as language testers, as applied linguists and as language teachers.

The book is divided into three sections. *A: Introduction* consists of ten units dealing with the central concepts of language testing and assessment. It contains activities for you to carry out alone, or with others if you are studying this book as part of a course. *B: Extension* provides extracts from articles or books relating to language testing and assessment which give you further insights into the concepts introduced in Section A. Each extract in *B: Extension* is accompanied by activities to focus your reading and help you to evaluate critically what you have read and understand how it links to a wider discussion of language testing and assessment. *C: Exploration* builds on the material you will already have found in the book. In this section we provide extended activities that help you to work through practical and theoretical problems that have been posed in the other sections. We also present ideas for individual and group project work, as well as suggestions for research projects.

The organization of this book allows you to concentrate on particular themes, such as *classroom assessment* or *writing items and tasks*, by reading the relevant units from *A: Introduction*, *B: Extension* and *C: Exploration* consecutively. Alternatively, you may wish to read the whole of *A: Introduction* before embarking on Sections B and C. In fact, you may decide to read the sections in any sequence, just as you would read Julio Cortázar's novel *Hopscotch*: there is no one right place to start, and each path through the text provides a different experience. Whichever choice you make, the book is extensively cross-referenced and carefully indexed so that you can easily find your way around the material.

At the end of the book we provide a glossary of key terms that are not explained within the text itself. If you come across a term about which you feel uncertain, simply turn to the glossary for an explanation. We also provide an extensive list of references for additional reading.

In addition to the book itself, there is also a website <http://www.routledge.com/textbooks/9780415339476> in which we provide very extensive additional reading, activities, links to relevant websites and further ideas for projects that you might like to undertake on your own or with colleagues.