

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second
Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Examining Reading

Research and practice in assessing
second language reading

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Also in this series:

Dictionary of Language Testing

Alan Davies, Annie Brown, Cathie Elder, Kathryn Hill, Tom Lumley, Tim McNamara

Learner Strategy Use and Performance on Language Tests: A structural equation modeling approach

James E. Purpura

Fairness and Validation in Language Assessment: Selected papers from the 19th Language Testing Research Colloquium, Orlando, Florida

Antony John Kunnan

Issues in Computer-adaptive Testing of Reading Proficiency

Micheline Chalhoub-Deville

Experimenting with Uncertainty: Essays in honour of Alan Davies

Edited by A. Brown, C. Elder, N. Iwashita, E. Grove, K. Hill, T. Lumley, K. O'Loughlin, T. McNamara

An Empirical Investigation of the Componentiality of L2 Reading in English for Academic Purposes

Edited by Cyril J. Weir, Yang Huizhong, Jin Yan

The Equivalence of Direct and Semi-direct Speaking Tests

Kieran O'Loughlin

A Qualitative Approach to the Validation of Oral Language Tests

Anne Lazaraton

Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913–2002

Edited by Cyril J. Weir and Michael Milanovic

A Modular Approach to Testing English Language Skills: The development of the Certificates in English Language Skills (CELS) examination

Roger Hawkey

Issues in Testing Business English: The revision of the Cambridge Business English Certificates

Barry O'Sullivan

European Language Testing in a Global Context: Proceedings of the ALTE Barcelona Conference July 2001

Edited by Cyril J. Weir and Michael Milanovic

IELTS Collected Papers: Research in speaking and writing assessment

Edited by Lynda Taylor and Peter Falvey

Testing the Spoken English of Young Norwegians: A study of testing validity and the role of 'smallwords' in contributing to pupils' fluency

Angela Hasselgreen

Changing Language Teaching through Language Testing: A washback study

Liyong Cheng

The Impact of High-stakes Examinations on Classroom Teaching: A case study using insights from testing and innovation theory

Dianne Wall

Assessing Academic English: Testing English proficiency 1950–1989 – the IELTS solution

Alan Davies

Impact Theory and Practice: Studies of the IELTS test and *Progetto Lingue 2000*

Roger Hawkey

IELTS Washback in Context: Preparation for academic writing in higher education

Anthony Green

Examining Writing: Research and practice in assessing second language writing

Stuart D. Shaw and Cyril J. Weir

Multilingualism and Assessment: Achieving transparency, assuring quality, sustaining diversity.

Proceedings of the ALTE Berlin Conference, May 2005

Edited by Lynda Taylor and Cyril J. Weir

Examining FCE and CAE: Key issues and recurring themes in developing the First Certificate in English and Certificate in Advanced English exams

Roger Hawkey

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Examining Reading

Research and practice in assessing
second language reading

Hanan Khalifa

Assistant Director

University of Cambridge ESOL Examinations

and

Cyril J Weir

Powdrill Professor in English Language Acquisition

University of Bedfordshire



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press

The Edinburgh Building, Cambridge CB2 8RU, UK

www.cambridge.org

Information on this title: www.cambridge.org/9780521736718

© UCLES 2009

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2009

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data

Khalifa, Hanan.

Examining reading : research and practice in assessing second language reading /

Hanan Khalifa and Cyril J. Weir.

p. cm.

Includes bibliographical references and indexes.

ISBN 978-0-521-73671-8

1. English language--Study and teaching--Foreign students--Evaluation.

2. Reading (Higher education)--Ability testing. 3. Second language acquisition--Research. I. Weir, Cyril J. II. Title.

PE1128.A2K418 2009

428.0076--dc22

ISBN 9780521736718

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Contents

List of abbreviations	vi
Series Editors' note	ix
Acknowledgements	xiii
1 Introduction	1
2 Test-taker characteristics	18
3 Cognitive validity	34
4 Context validity	81
5 Scoring validity	143
6 Consequential validity	169
7 Criterion-related validity	190
8 Conclusions and recommendations	217
Appendices	
Appendix A: Sample Reading tasks at five levels and answer keys	233
Appendix B: Candidate Information Sheet	261
Appendix C: Administrative setting	262
Appendix D: Standard procedures for the production of Reading examination materials	269
Appendix E: General marking: performance management	276
Appendix F: Computer-based (CB) reading tests versus paper-based (PB) versions: some issues and directions	281
References	307
Author index	332
Subject index	340

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Abbreviations

Adj	Adjective
AERA	American Educational Research Association
ALTE	Association of Language Testers in Europe
AOG	Assessment & Operations Group
APA	American Psychological Association
AWL	Academic Word List
BEC	Business English Certificates
BNC	British National Corpus
BULATS	Business Language Testing Service
CAE	Certificate in Advanced English
CB	Computer Based
CBT	Computer Based Testing
CEF	Common European Framework
CEFR	Common European Framework of Reference
CELS	Certificates in English Language Skills
CELTA	Certificate in English Language Teaching to Adults
CET	College English Test
CFA	Confirmatory Factor Analysis
CIS	Candidate Information Sheet
CLA	Communicative Language Ability
CLB	Canadian Language Benchmarks
CPE	Certificate of Proficiency in English
CRELLA	Centre for Research in English Language Learning and Assessment
CTT	Classical Test Theory
CUP	Cambridge University Press
DELTA	Diploma in English Language Teaching to Adults
DIALANG	Diagnostic Language (Assessment)
DIF	Differential Item Functioning
EAL	English as an Additional Language
EAP	English for Academic Purposes
EBAFLS	European Bank of Anchor Items for Foreign Language Skills
EFL	English as a Foreign Language
ELT	English Language Teaching
EPS	Examination Processing System
ESM	Electronic Script Management

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Abbreviations

ESOL	English for Speakers of Other Languages
ESP	English for Specific Purposes
ETS	Educational Testing Service
FCE	First Certificate in English
GEPT	General English Proficiency Test
GM	General marking
GPC	Grapheme–Phoneme Correspondence
GRE	Graduate Record Examination
IATM	Instrument for the Analysis of Textbook Materials
iBT TOEFL	Internet Based Test of English as a Foreign Language
ICFE	International Certificate in Financial English
ICR	Intelligent Character Recognition
IDP	International Development Program
IELTS	International English Language Testing System
IF	Item Facility
IIS	IELTS Impact Study
ILEC	International Legal English Certificate
ILTA	International Language Testing Association
IRT	Item Response Theory
ITC	International Test Commission
IW	Item Writer
IWG _s	Item Writer Guidelines
KET	Key English Test
KR-20	Kuder-Richardson 20
L1	First Language
L2	Second Language
LCE	Lower Certificate in English
LIBS	Local Item Banking System
MCQ	Multiple Choice Questions
MQC	Marking-Quality Coordinator
N	Noun
NCME	National Council on Measurement in Education
OMR	Optical Mark Reader
PARA	Partnership for Accessible Reading Assessment
PB	Paper Based
PET	Preliminary English Test
PMS	Performance Management System
PRD	Project Research and Development
QMS	Quality Management System
QPP	Question Paper Production
QPT	Quick Placement Test
RITCME	Recruitment, Induction, Training, Co-ordination, Monitoring, Evaluation

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Abbreviations

RNIB	Royal National Institute for the Blind
SAQ	Short Answer Questions
SD	Standard Deviation
SE	Standard Error
SEM	Standard Error of Measurement
SILT	Studies in Language Testing
SO	Subject Officer
SVO	Subject Verb Object order
TEEP	Test in English for Educational Purposes
TEFL	Teaching English as a Foreign Language
TESOL	Teaching English to Speakers of Other Languages
TKT	Teaching Knowledge Test
TLX	Task Load Index
TM	Test Method characteristics
TOEFL	Test of English as a Foreign Language
TOEIC	Test of English for International Communication
TTR	Type-Token Ratio
UAT	User Acceptance Test
UCLES	University of Cambridge Local Examinations Syndicate
UETESOL	University Entrance Test in English for Speakers of Other Languages
UN	United Nations
VDT	Visual Display Text
VDU	Visual Display Unit
VP	Vocabulary Profiler
VRIP	Validity, Reliability, Impact, Practicality
VSO	Verb Subject Object order
WPM	Words per minute
YLE	Young Learners English Tests

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Series Editors' note

Examining Reading is the second volume in the Studies in Language Testing (SiLT) series that addresses the approach used by Cambridge ESOL in the assessment of language skills, the first being SiLT 26, *Examining Writing* by Shaw and Weir (2007). This volume sets out to describe and evaluate how Cambridge ESOL tests different levels of reading in English as a second language across the range of examinations it offers spanning the Reference Levels of the Common European Framework of Reference (CEFR) from A2 to C2, through focusing largely on the five examinations in the Cambridge ESOL Main Suite (KET, PET, FCE, CAE, CPE). As with *Examining Writing* (2007), it does so by presenting an explicit framework that structures the approach to validation according to a number of dimensions or parameters. It utilises the same theoretical framework as *Examining Writing* which was originally proposed by Weir (2005) and which seeks to take account of both the aspects of cognition, related to the mental processes the individual needs to engage in order to address a task, and the features of language use in context that affect the ways in which a task is addressed. The authors also look at the practical assessment issues related to the marking and scoring of reading tests. As with *Examining Writing* therefore, this volume explores the triangular relationship between three critical internal dimensions of language testing tasks – the test takers' *cognitive abilities*, the *context* in which the task is performed and the *scoring process*. Set alongside these are the twin external dimensions of consequential validity and criterion-related validity.

Cambridge ESOL has been involved in the assessment of reading skills ever since it launched its first English language examination in 1913. Since that time we have seen a significant development in our understanding of the reading construct from its early conceptualisation as an integrated skill, assessed largely by translation and even reading aloud tasks, through to the approaches documented in this volume.

When the CPE was introduced in 1913, reading as a skill in its own right did not feature explicitly among those to be examined. The history and most recent revision of CPE is well documented in SiLT 15, *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913–2002* (Weir and Milanovic 2003). The volume explains how the approach to the design of CPE in 1913 was based on *The Practical Study of Languages* (Sweet 1899) and candidates spent over twelve hours on a demanding set of activities that included translation to and from English, an essay on

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Examining Reading

a topic such as Elizabethan travel and discovery, an English literature paper, English phonetics, dictation, reading aloud and conversation. While various reading skills were required throughout the examination, the skill itself was not singled out for particular attention as it would be today nor in the way that phonetics, grammar, reading aloud and conversation were singled out in the 1913 examination.

Little changed in relation to the assessment of reading for the 1938 version of CPE other than the introduction of an alternative to the English Literature paper focussing on General Economic and Commercial Knowledge. Candidates had to read extensively in order to engage with these papers but did so outside the context of the examination room, clearly extremely powerful with regard to the impact of the test on learning/teaching. Summarisation and explanation were required of candidates when tackling some of the questions in both of these papers. Typically, candidates had to choose between summarising a passage, which included defining the meaning of words and phrases in the text, and explaining a poem in detail including a focus on style and diction.

The Lower Certificate in English, introduced in 1939 and later known as the First Certificate in English (1975), tackled reading in a slightly different way. In its paper entitled Prescribed Texts, candidates had to read up to four prescribed texts typically taken from what we would today refer to as the classics of English literature but what would then have been closer to contemporary fiction. In the examination candidates had to answer questions on two of the texts. A short excerpt from each was presented and candidates were asked a number of open-ended questions, sometimes requiring close inspection of the text itself and sometimes a broader interpretation. In the Composition paper Part C was mandatory and required candidates to answer some short questions focusing on the meaning of vocabulary in a text, propose a title for it and summarise a part of it. Some of these activities would be familiar in reading tests today.

The 1955 variant of CPE introduced a Use of English paper and reading was tested both in this paper and in the English Language paper (introduced in the place of the Composition paper in 1945), though it was not referred to as reading in either. Short answer comprehension questions, summarisation and vocabulary questions were all used. From the 1940s to the mid 1970s relatively little changed on the surface but behind the scenes a revolution was taking place that would mark a radical shift in an approach that had been changing gradually since 1913.

There was much talk internally in the 1960s that the Cambridge approach lacked objectivity, that is, that it was in some way behind the times, that it needed to focus more explicitly on the four skills and that it was time to start making use of more 'scientific' methods of assessment, i.e. multiple choice questions, analysis, and, importantly, the demonstration of test reliability.

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Series Editors' note

This affected CPE gradually at first in that multiple choice questions were introduced into the 1967 variant of the Use of English Paper. The focus was largely on vocabulary and the extent to which the questions were subjected to analysis was limited, but an important statement had been made. The 1975 release of both FCE and CPE made a much bolder statement. Cambridge exams now explicitly tested the four skills of Reading, Listening, Speaking and Writing. Both Listening and Reading were introduced in their own right. At the same time, the Use of English paper continued with its micro focus on grammatical and lexical knowledge. The reading construct, though not defined as such at this stage, involved reading a number of short passages and answering multiple-choice questions on them. The focus was on explicit and implicit meaning at both levels. There were also 25 discrete-point multiple-choice vocabulary questions. The reading construct appeared to reflect a combination of reading skills along with a knowledge of vocabulary, sometimes of a somewhat esoteric nature.

The FCE and CPE examinations changed relatively little in the next 21 years; however, as before, there was significant activity behind the scenes. The work of the Council of Europe in its Modern Languages programme, the emergence of the Threshold level, and the rise of the communicative language teaching movement all happened in the 1970s and 1980s and impacted on the Cambridge approach to language testing. In addition, an important study carried out in the late 1980s was to have quite a powerful influence on the shape of things to come. Bachman et al (1995) carried out a Cambridge-sponsored study entitled *An Investigation into the Comparability of Two Tests of English as a Foreign Language* (SiLT 1). While ostensibly looking at the comparison between FCE and TOEFL in order to establish an empirical link between the level systems of each examination, this study actually ended up providing an in-depth critique of the Cambridge approach with specific reference to the then well developed and documented psychometrically-oriented approach as demonstrated by the TOEFL. Significant issues in relation to reliability and validity emerged which were addressed vigorously with the 1996 release of the FCE and subsequent release of CPE in 2002. Apart from a range of measurement issues not particularly relevant to this discussion, a much sharper focus on test construct definition and validation emerged. Where test construct had had to be pieced together post hoc from test specifications in earlier releases of FCE and CPE, there were now explicit statements on test construct. The Reading paper in particular, underwent fairly radical revision and its construct was defined as 'Understanding the propositional, functional and sociolinguistic meanings at word, phrase, sentence or discourse levels and of reading outcomes relevant to FCE takers (i.e. gist, specific information, detail, main idea, deduced information).' Measures were put in place not only to develop test content with systematic reference to the underlying construct but also to validate the nature of that construct. It

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Examining Reading

is this important step that has taken us to the work described in this volume and the others in this series. The volume on Writing assessment has already been published as SiLT 26, and parallel volumes on Speaking and Listening are currently in preparation for publication over the next two years.

More recently, we have also seen the emergence of the Common European Framework of Reference for Learning, Teaching and Assessment (Council of Europe 2001) which encourages examination providers to map their certification to the Framework. Khalifa and Weir examine how Cambridge has approached this task in significant depth when exploring criterion-related validity. The approach taken by Cambridge seeks not only to establish the relationship with the Framework as a one-off study, but to deploy a methodology that ensures a long-term and continually verifiable relationship which is surely in the overall best interests of test users.

Skills assessment at Cambridge is now underpinned more formally than ever by a validation framework based on Weir (2005) and building on the work of Bachman (1989) which informed validation activities in the 1990s, as well as the VRIP approach developed by Cambridge in the 1990s. The approach outlined in this volume not only allows Cambridge to determine where current examinations are performing satisfactorily in relation to a range of relevant validity parameters, it also provides the basis for improvement and the construction of an ongoing research agenda. It provides an important benchmark against which test developers can evaluate the effectiveness of their respective approaches and it offers test users a model of what to expect from responsible examination providers.

Michael Milanovic and Cyril J Weir
Cambridge – April 2009

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Acknowledgements

In bringing this volume to fruition, we are deeply indebted to a great number of colleagues. Their expertise, sound advice and collaboration have undoubtedly contributed to the quality of this volume.

We would like to thank Professor Eddie Williams (Bangor University) and Dr Roger Hawkey (University of Bedfordshire) for their many insightful comments when reviewing the whole manuscript. We wish to express special thanks to Dr Lynda Taylor for her painstaking reviews of, and thought-provoking comments on, the various drafts of the manuscript.

We are also grateful to a number of experts within the field of applied linguistics and language testing who provided valuable comments on individual chapters: Professor Antony Kunnan (UCLA) for his reflections on the test-taker characteristics chapter; Dr John Field (Reading University) whose views have been instrumental in shaping the chapter on cognitive validity; Dr Felicity O'Dell (Testing Consultant) and Dr Norbert Schmitt (Nottingham University) for their analysis of lexical, functional and structural resources in Main Suite reading papers; Dr Nick Saville and Dr Roger Hawkey for their valuable input to the chapter on consequential validity. Particular thanks go to the chairs of Main Suite Reading papers for their diligent and attentive reviews of text relating to their respective areas of interest: Sharon Ashton, Annette Capel, Anthea Bazin, Wendy Sharp and Judith Wilson.

Sincere thanks are due to three Cambridge ESOL colleagues who have read through the different manuscripts, judiciously reviewed them, provided insightful comments and sometimes heated discussions: Angela Ffrench (FCE/CAE/CPE Subject Manager), Mick Ashton (PET Subject Officer) and Glyn Hughes (FCE Subject Officer). We deeply appreciate their expertise and the time they have given us from their busy schedules.

This volume could not have been completed without the additional co-operation of numerous Cambridge ESOL personnel, many of whom enabled us to represent fully both the practitioner voice and the researcher voice. We would like to acknowledge here contributions made by members of the Assessment and Operations Group: Anne Gutch (Assistant Director), Edward Hackett (KET/PET Subject Manager), Margaret Cooze (KET Subject Officer), Helen Coward (CAE Subject Officer), Andrew Balch (CPE Subject Officer), Sharon Jordan (Computer Based Testing Subject Manager), Cris Betts (Senior Operations Manager), David Corkill (Use of English Subject Officer), Dr Rod Boroughs (IELTS Subject Officer) and Hugh

Cambridge University Press

978-0-521-73671-8 - Examining Reading: Research and Practice in Assessing Second Language Reading

Hanan Khalifa and Cyril J. Weir

Frontmatter

[More information](#)

Acknowledgements

Bateman (BEC Subject Manager). We would also like to acknowledge contributions made by staff from the Research and Validation Group: Dr Neil Jones (Assistant Director), Dr Ardeshir Geranpayeh (Assistant Director), Dr Andrew Blackhurst (Senior Coordinator), Dr Andrew Somers (Main Suite Validation Officer), Dr Fiona Barker (Corpora Validation Officer), and Nick Beresford-Knox (Research Projects Assistant). Last but not least, we would like to thank Sonia Liddiard (Centre Manager) and Juliet Wilson (Assistant Director) for their contributions to Appendix C.

Finally, we would like to recognise the contribution of Dr Michael Milanovic (Cambridge ESOL Chief Executive) whose vision made the publication of the construct series a reality and whose willingness to support future research into issues raised in this volume makes this publication worthwhile.

To all of the above, and to others we have failed to mention, we offer our most sincere thanks and appreciation.

The publishers are grateful to the copyright holders for permission to use the copyright material reproduced in this book. Council of Europe and Cambridge University Press for Table 1.1, Table 1.2 and Table 4.7 from *Common European Framework of Reference for Languages: Learning, teaching, assessment*, 2001.