

剑桥语言测试研究丛书

The Development of IELTS

A Study of the Effect of
Background Knowledge on
Reading Comprehension

雅思的发展：

背景知识对阅读理解的影响研究

Caroline Clapham 著

剑桥语言测试研究丛书

The Development of IELTS

A Study of the Effect of
Background Knowledge on
Reading Comprehension

雅思的发展：

背景知识对阅读理解的影响研究

Caroline Clapham 著

图书在版编目 (CIP) 数据

雅思的发展: 背景知识对阅读理解的影响研究 / (英) 卡洛琳·克拉彭著.

—上海: 上海外语教育出版社, 2018

(剑桥语言测试研究丛书)

ISBN 978-7-5446-5412-8

I. ①雅… II. ①卡… III. ①IELTS—阅读教学—教学研究

IV. ①H319.37

中国版本图书馆CIP数据核字 (2018) 第109040号

This is a reprint edition of the following title published by Cambridge University Press:

The Development of IELTS: A Study of the Effect of Background Knowledge on Reading Comprehension (ISBN: 9780521567084)

© Cambridge University Press 1996

This reprint edition for the People's Republic of China (excluding Hong Kong, Macau and Taiwan) is published by arrangement with the Press Syndicate of the University of Cambridge, Cambridge, United Kingdom.

© Cambridge University Press and Shanghai Foreign Language Education Press 2018

This reprint edition is authorized for sale in the People's Republic of China (excluding Hong Kong, Macau and Taiwan) only. Unauthorised export of this reprint edition is a violation of the Copyright Act. No part of this publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of Cambridge University Press and Shanghai Foreign Language Education Press.

本书版权由剑桥大学出版社和上海外语教育出版社有限公司共同所有。本书任何部分之文字及图片, 未经出版者书面许可, 不得用任何方式抄袭、节录或翻印。

此版本仅限中华人民共和国境内销售, 不包括香港、澳门特别行政区及中国台湾。不得出口。

图字: 09-2017-927号

出版发行: 上海外语教育出版社

(上海外国语大学内) 邮编: 200083

电 话: 021-65425300 (总机)

电子邮箱: bookinfo@sflep.com.cn

网 址: <http://www.sflep.com>

责任编辑: 苗 杨

印 刷: 启东市人民印刷有限公司

开 本: 635×965 1/16 印张 20 字数 419千字

版 次: 2018年12月第1版 2018年12月第1次印刷

印 数: 2 100 册

书 号: ISBN 978-7-5446-5412-8 / G

定 价: 65.00 元

本版图书如有印装质量问题, 可向本社调换

质量服务热线: 4008-213-263 电子邮箱: editorial@sflep.com

出版说明

语言测试作为语言教学的重要组成部分，一直是社会关注的焦点。近年来，我国学术界对语言测试研究的关注也在逐渐升温。国际上，学者们对语言测试的研究已经达到较高的水平，无论是从理论框架的形成到复杂数学模型的建构及其应用方面都取得了实质性的进展，而且在对有国际影响力的外语考试（如雅思、托福等）的信度、效度的研究方面也取得了不俗的成果，出版了不少优秀的著作。

英国剑桥大学出版社出版的“Studies in Language Testing”是其中较有影响力的系列之一。丛书由剑桥语言测评中心策划推出，从测试领域的各个研究视角入手，展示了该领域研究的设计、数据收集和数据分析等环节的基础理论与实践操作，反映了测试领域的最新研究动态。

外教社从中精选9种，引进出版。这些图书中，有对语言测试听说读写等方面的效度和质性研究，也有对测试反拨效应及其对教学影响的研究，还有对雅思考试和欧洲语言共同参考框架的分析思考。希望丛书的出版可以为我国语言测试研究领域的发展和创新提供具有借鉴意义的启示，打开更广阔的研究思路。

*The development of IELTS: A
study of the effect of
background knowledge on
reading comprehension*

Caroline Clapham

In memory of DCT

Series Editor's note

Background

This volume of *Studies in Language Testing*, is based on doctoral work carried out by the author, Caroline Clapham, over a number of years and supported by UCLES. Her particular interest lay in the testing of reading, which is one of the four skills-based modules that make up IELTS. She investigated the ESP beliefs underlying the design of the reading components of ELTS and an early version of IELTS. Her work shows how difficult it is both to classify students according to their background knowledge, and to select reading passages which are genuinely specific for people in any one subject area. She suggests, therefore, that it is fairer for students if they all take a single academic reading module. The conclusions are of general importance to the designers of EAP proficiency tests. The investigating methods used in the study for assessing the appropriacy of the reading passages, and the resulting demonstration that text selectors are often unable to estimate the specificity of the texts they choose, will be of interest to all those who have to select reading texts for testing purposes or for research into reading in a second language.

As a point of information, the International English Language Testing System (IELTS) as it now stands provides an assessment of whether candidates are ready to study or train through the medium of English. It is recognised widely as a language requirement for entry to courses in further and higher education, and is readily available, being offered 'on demand' at test centres around the world. IELTS is jointly managed by The University of Cambridge Local Examinations Syndicate (UCLES), the British Council and IDP Education Australia Limited.

IELTS can be taken at around 200 approved test centres in well over 100 different countries. The test is administered centrally by UCLES but the approved centres, most of which are British Council or IDP Offices, supervise the local administration of the test and ensure the provision of qualified and trained examiners. IELTS is not held on set dates during the year but is conducted according to demand. Most centres conduct a testing session at least once a month and more often at peak times.

On a historical note, the original English Language Testing Service (ELTS) was developed in the late seventies by British Council staff, headed by Brendan Carroll, and became operational in 1980. It was one of the first language tests to take into account the communicative movement and drew heavily on the work

of Munby (1978). In 1986, a decision was made to revise this test and a team at Lancaster University, headed by Charles Alderson, took on the project. The team was joined in 1987 by David Ingram, who acted as a representative of the International Development Program of Australian Universities and Colleges (IDP). At this time the test was renamed the International English Language Testing System (IELTS). The revision team drew on a variety of sources to guide them. These included: the ELTS Validation Study, carried out by Alan Davies and Clive Cripser, the views of ELTS users, EAP teachers, language testers and applied linguists.

The ELTS Validation Study reported strengths and weaknesses on a number of fronts. From the point of view of practicality, ELTS was felt to be rather cumbersome with its six academic modules. On the other hand, it was found to have high face validity for exactly the same reason. Flaws in the test design were attributed to weaknesses in the theory of English for Specific Purposes (ESP), which guided test development to some extent. Such findings along with extensive consultation guided the revision team. IELTS was released in late 1989. It had a focus on the four skills and made use of three academic reading modules and a single general training reading module. The Writing module was also subject specific and linked to the reading modules. The speaking and listening modules were general and taken by all candidates.

Consistent with UCLES policy of on-going validation, IELTS was carefully monitored in the early 1990s and by 1992, it was decided to modify the test. The monitoring suggested that the three academic modules should be reduced to one. Caroline Clapham's work reported in this volume informed this decision to some extent. The revised version of IELTS was introduced in April 1995. Materials for IELTS are now written by teams of item writers in the UK and Australia. All materials are pretested and calibrated to the IELTS scale. The test itself provides a profile of ability to use English. A score in each of the four modules or skills, and an overall score, are recorded as levels of ability, called Bands.

Assessment of performance in IELTS does not depend on reaching a fixed pass mark. It depends on how the candidate's ability in English relates to the language demands of courses of study or training. The appropriate level required for a given course of study or training is ultimately something which institutions must decide in the light of knowledge of their own course and their experience of overseas students taking them. There are six modules in IELTS. All candidates must take four modules, one in each of the four skills. All candidates take the same Listening and Speaking Modules, whereas there is a choice of Reading and Writing Modules with either a general training or academic focus.

Since its 1995 revision IELTS has adhered to a clearly stated code of practice. This has required the implementation and maintenance of systems and procedures designed specifically to validate the test, evaluate the impact of the test and provide relevant information to test users.

It is recognised that as a test provider, UCLES has an impact on educational processes and on society in general. This impact operates on at least two levels in terms of:

- i education and society in general
- ii people who are directly affected by tests and their results

We believe it to be important to be able to investigate the educational impact IELTS has within the context that it is used. As a point of principle, test developers must operate with the aim that their tests will not have a negative impact and, as far as possible, strive to make it positive. In general terms, this can be achieved through the development and presentation of test specifications and detailed syllabus designs, and provision of professional support programmes for institutions and individual teachers/students.

It is anticipated that positive educational impact in the case of IELTS can be achieved through the following practices:

- the identification of suitable experts within any given field to work on all aspects of test development;
- the training and employment of suitable experts to act as question/item writers in test production;
- the training and employment of suitable experts to act as examiners.

Procedures are required to collect information which allows impact to be estimated and attention is being focused on the following areas through routine data collection or further research:

- who is taking the test (i.e. a profile of the candidates);
- who is using the test results and for what purpose;
- who is teaching towards the test and under what circumstances;
- what kinds of courses and materials are being designed and used to prepare candidates;
- what effect the test has on public perceptions generally (e.g. regarding educational standards);
- how the test is viewed by those directly involved in educational processes (e.g. by students, test takers, teachers, parents, etc.);
- how the test is viewed by members of society outside education (e.g. by politicians, businessmen etc.).

It is hoped that aspects of this research will be reported on in this series.

Preface

The purpose of this research is firstly to investigate the ESP claim that tertiary level ESL students should be given reading proficiency tests in their own academic subject areas, and secondly to study the effect of background knowledge on reading comprehension. The study is set against a background of recent research into reading in a first and second language, and emphasises the impact that schema theory has had on this.

Students took two versions of the International English Language Testing System (IELTS) test, which has reading modules in three different academic subject areas. Analyses of variance showed that the reading subtests varied in their subject specificity: some were suitable for students in the relevant academic field, others were either too general or too specific. A Rasch analysis of the items revealed little bias against students who took an inappropriate reading module, and an investigation of the test content using a version of Bachman's Test Methods Characteristics scale showed that the test items did not seem to affect test specificity. Variation in the appropriacy of the reading passages was found to be partly due to differences in rhetorical function, partly to uncertainty among EAP teachers about concepts relating to academic and topic specificity, partly to the extent of students' familiarity with the subject areas, and partly to the fact that students vary widely in their background reading and cannot be accurately placed into three distinct subject areas.

Further analysis suggested that the relative importance of language proficiency and background knowledge in reading comprehension depended on the specificity of the reading passages, and an investigation into whether language ability affected the students' use of background knowledge supported the hypothesis that there is a threshold level below which learners have difficulty making use of this knowledge.

The book concludes by considering the implications of the findings for future test construction and research into reading.

Acknowledgements

I have to thank so many people for help with the research required for this book that I cannot do justice to everybody's contribution.

The project would not have been possible without the members of the ELTS Revision Project, Charles Alderson, Peter Hargreaves, David Ingram, John Foulkes and Gill Westaway, with whom I worked on the design and construction of the trial IELTS, and the many teachers and testers who wrote the Grammar Test and the Trial and Exemplar Versions of the Reading Modules.

The following people very kindly either administered or arranged for me to administer the tests at their language centres: Don Dunmore, Glenn Fulcher, Gail Langley, Kenneth James, Philip King, John Read, Margaret Rutherford, Chris Shaw, Philip Shaw, Anat Stavans, Susan Taylor, Alan Tonkyn, Carolyn Walker, Dianne Wall and Rosemary Zahn. I must also thank the 800 or so students who, in many cases voluntarily, spent about three hours of their time taking what proved to be very demanding tests.

I am most grateful to Carolyn Turner for providing me with the McGill system of clausal analysis, Michael Milanovic for arranging for the test papers to be printed and Lyle Bachman for allowing me to use an adaptation of his Test Methods Characteristics Rating Scale, and for answering all my queries about it. I also owe particular thanks to Barbara Adams, Joan Allwright and Nicki McLeod who spent several hours using the rating scale to assess the reading tests.

So many people have advised me about analysing my results that I cannot thank them all personally, but I would particularly like to thank Brian Francis of the Department of Applied Statistics at this university. I should also like to thank Tim McNamara, Sarah Cushing Weigle, Alastair Pollitt, Neil Jones and Mary Shegl for advice about identifying item bias, and Grant Henning, Fred Davidson and Dorry Kenyon for general advice over the years.

In particular I am indebted to Charles Alderson who encouraged me to start on the project, and inspired me to move in interesting directions.

Finally I must thank my daughter, Phoebe, for reading through the book, my son, Tom, for formatting the tables, and my husband, Christopher, for his advice and encouragement.

Caroline Clapham

Contents

Series Editor's note	VII
Preface	X
Acknowledgements	XI
Chapter One	
Introduction	1
English for Specific Purposes	2
ESP Testing	5
Research into ESP testing	6
Chapter Two	
Reading in a First Language	11
Reading Research	11
Models of Reading Comprehension	13
Schema Theory	16
Limitations of Schema Theory	27
Research into the Effect of Content and Cultural Background	
Knowledge on Reading	28
Conclusions about Reading in a First Language	31
Chapter Three	
Reading in a Second Language	33
Comparison of L1 and L2 Reading Processes	33
The Effect of Context and Background Knowledge on L2 Reading	
Comprehension	38
Overall Conclusions about Reading in a Second Language	46
Outstanding Research Issues	46
Chapter Four	
The Development of the IELTS Reading Modules	50
The Precursors of IELTS	50
IELTS - Designing the Reading Modules	57
IELTS Trials	73
Construct of the IELTS Reading Modules	76
Conclusion	78

Chapter Five	
The Pilot Study	79
Research Question	79
The Tests	79
The Examinees	80
Analysis	81
The Results	82
Discussion	85
The Main Study	87
Chapter Six	
The Main Study - Research Questions and Data Collection	88
Research Questions	88
Research Instruments	91
The Students	100
Test Administration	103
Test Results	105
Conclusion	109
Chapter Seven	
The Effect of Academic Subject Area on Test Performance	110
Section 1: Analysis of Variance - the Complete Tests	110
Section 2: The Reading Subtests	117
Section 3: Item Bias	121
Section 4: Unanswered Items	128
Conclusion	131
Chapter Eight	
Reasons for the Variation in Subtest Specificity	133
Familiarity of Subject Area and Topic	133
Subject Specialists' Views	138
Conclusion	142
Chapter Nine	
Text Source and the TMC Rating Instrument	144
Source of Reading Passages	144
Bachman's TMC Instrument	147
Conclusion	162
Chapter Ten	
Language Proficiency and Background Knowledge	164
The Revised Reading Modules	164
Effect of Background Knowledge on Reading Comprehension	166

Multiple Regression Analysis	171
Language Proficiency and the Effect of Background Knowledge	177
Conclusion	186
Chapter Eleven	
Conclusions	188
The Research Questions	188
Research Implications	194
ESAP or EGAP	199
Suggestions for Further Research	202
Summary of the Main Findings	204
References	206
Appendices	222
Subject Index	289
Author Index	297

1 Introduction

In recent years there has been increasing agreement among testers that language proficiency tests should, where possible, be related to candidates' future language needs. For example, if doctors are to be tested for their capacity to use English in an English-speaking hospital, it is considered only sensible to test them on the kinds of English that are used in the ward and the consulting room. Similarly, future air traffic controllers might be tested on the English needed in the control tower. No one is likely to question the good sense of such English for Specific Purposes (ESP) testing. The issue, however, becomes less clear cut when the proficiency test is aimed at a less sharply defined audience, such as students proposing to embark on tertiary education. Many testers consider that a language proficiency test for such students should contain samples of the kinds of language tasks required of them in their academic work, but it is not clear how much, if at all, these tasks differ from discipline to discipline, nor how much the subject matter of the test should vary according to the discipline of the examinee. The question here is whether there should be separate tests for students in the different academic disciplines, or whether all students should take a single test battery. There is some evidence to show that the language tasks in different academic disciplines are sufficiently similar for one set of test tasks to be appropriate for all (see Weir 1983 and Chapter 4 below), but it is not clear whether the subject matter of the tests should be different.

The results of research into the effects of field specific reading tests on EAP students' test performance have been somewhat contradictory, and no conclusive evidence has been produced either for or against the use of ESP tests. This book, therefore, reports on a large scale study into the effects of giving subject specific reading tests to future university students. The main aim of this study is to see whether an ESP approach to testing the reading proficiency of academic students is appropriate and feasible, and the secondary aim is to consider the effect of background knowledge on reading comprehension.

This first chapter briefly introduces ESP teaching and testing, and describes recent research into whether domain specific background knowledge affects test scores in English comprehension tests. Chapters 2 and 3 review research into the effect of background knowledge on reading in a first and a second language. Chapter 4 describes the construction of the reading component of the International English Language Testing System (IELTS) test, and Chapter 5 reports on a pilot study into the effect of subject area on test performance. Chapter 6 presents the research questions for the main study, and describes the tests, the

questionnaire and the student sample. Chapter 7 describes a replication of the pilot study using a different set of tests, and gives the results of other investigations into the effect of subject area on test performance. These studies show that the reading subtests vary in their subject specificity, and Chapters 8 and 9 discuss the reasons for this variation. Chapter 10 looks at the effects of level of language proficiency on the use of background knowledge in reading, and also compares the effects of language proficiency and background knowledge on test scores. Chapter 11 summarises the main research findings, discusses their implications, and makes proposals for further research.

English for Specific Purposes

The main drive behind the introduction of ESP was practical rather than theoretical. With the rapid increase after the war in the importance of English for education, technology and commerce, increasing numbers of people around the world needed to learn English for clearly defined reasons such as reading academic textbooks or transacting business (see Hutchinson and Waters 1987). These changes coincided with developments in communicative methods of teaching, and led ESP course designers to base their materials on texts and activities which were tailored to suit students with specific linguistic needs.

Although there had been some awareness from the 1920s of the fact that learners in different jobs needed different kinds of language (see Widdowson 1983), the ESP movement only came into existence in the 1960s. In its early stages ESP researchers focused on register analysis – the analysis of sentence-level grammatical and lexical features to see what the distinctive features were between texts in different subject areas. These analyses often took the form of frequency counts of structures or verb forms (see for example, Barber 1962/1985), or clausal analysis (see Huddleston 1971), and as Swales (1985:59) said, although such analyses had descriptive validity, they had little explanatory force. Widdowson (1979:55) pointed out that the fact that English scientific texts had a relatively high proportion of some syntactic structures and a relatively low proportion of others did not reveal anything about scientific discourse as a whole, and Hutchinson and Waters (1987:10) said that few systematic differences were found between scientific and ‘General English’ texts. As the focus of linguistic research changed from being sentence-based to include research into how sentences combine to produce meaning, and as the increasing importance of sociolinguistics led to the study of language variation in different contexts, researchers such as Widdowson (1979) and Trimble (1985) began to apply rhetorical or discourse analysis to discover the main characteristics of Scientific and Technical English (EST) texts, and to see if there were differences between these texts and non-EST ones. Trimble, for example, built his studies and teaching round three rhetorical concepts:

- a) the nature of the EST paragraph;
- b) the rhetorical techniques most commonly used in written EST discourse;
and
- c) the rhetorical functions most frequently found in written EST discourse (Trimble, 1985:14).

So much ESP research has focused on EST that it is easy to think of ESP and EST as synonymous. However, EST is an offshoot of ESP, on a par with, for example, English for Social Scientists. Since there is some disagreement among ESP teachers and researchers about the hierarchy of ESP terms, I will explain how the term 'ESP' is used in this book and how it relates to English for Academic Purposes (EAP).

There seem to be two favourite ways of treating the concepts of ESP and EAP. Some ESP teachers consider that since many EAP courses are designed to suit students in all academic disciplines, EAP is too broad in scope to be considered a branch of ESP. These teachers think of it as the superordinate category from which spring increasingly specific types of ESP (see Jordan 1989). However, this takes no account of other types of ESP such as English for Occupational Purposes (EOP). Since EAP is itself a type of ESP because it is concerned with the English required for a specific purpose, that of studying at universities and colleges, it is more usual to think of ESP as the superordinate term, with EAP and EOP branching from it (see Jordan 1989 and Robinson 1991). EAP courses can be divided into those for English for General Academic Purposes (EGAP) and those for English for Specific Academic Purposes (ESAP) (see Blue 1993). ESAP courses can range from broad groupings of subjects, such as EST or Liberal Arts, to ones which are so highly specific that they are suitable only for single students or for small groups in one narrowly defined discipline. ('Discipline', 'Field of study' and 'Subject area' are used interchangeably in this book.) Figure 1.1 shows how the terms are used here, and gives examples of possible courses at the different levels of the hierarchy.

In recent years the focus of academic textual analysis has broadened to take account of different genres of writing ranging from academic articles to article abstracts and citations. Genre analysis studies not only the composition of texts, but also the roles that those texts play. According to Swales (1990), the academic world consists of a series of 'discourse communities', each of which uses a shared set of genres in order to achieve a common set of purposes.

By 'genre' is meant a typified socially recognised form that is used in typified social circumstances. It has characteristic features of style and form that are recognised, either overtly or covertly, by those who use the genre. Thus for example, the research article has a known public purpose, and has conventions about layout, form and style that are to a large degree standardised.
(Dudley-Evans 1987)