

*Studies in
Language
Testing* **24**

Impact Theory
and Practice

Studies of the IELTS test
and *Progetto Lingue 2000*

Roger Hawkey

Series Editors
Michael Milanovic
and Cyril Weir



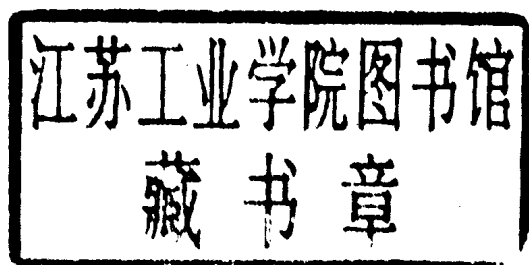
UNIVERSITY of CAMBRIDGE
ESOL Examinations

CAMBRIDGE
UNIVERSITY PRESS

Impact Theory and Practice

Studies of the IELTS test
and *Progetto Lingue 2000*

Roger Hawkey
ESOL Consultant



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

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

Cambridge University Press

The Edinburgh Building, Cambridge CB2 2RU, UK

www.cambridge.org

Information on this title: www.cambridge.org/9780521680974

© UCLES 2006

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 2006

Printed in the United Kingdom at the University Press, Cambridge

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

ISBN -13 978-0-521-68097-4

ISBN -10 0-521-68097-2

The author is grateful to the copyright holders for permission to use the copyright material reproduced in this book. Every effort has been made to trace the copyright holders. Cambridge University Press apologises for any unintentional omissions and would be pleased, in such cases, to add an acknowledgement in further editions.

Acknowledgements

I am sincerely grateful to the many people who helped me, in various ways, with the impact studies described here, and with the writing and editing of the book.

I was fortunate enough to be able to meet, at Lancaster, members of the Lancaster University team who played such a key part in the early phases of the study of IELTS impact. From Charles Alderson, Professor of Linguistics and English Language Education in the Department of Linguistics and English Language, Dr Jayanti Banerjee, researcher in language testing and assessment and English for Academic Purposes (EAP), and Tania Horak, Research Associate currently working on the New TOEFL (Test of English as Foreign Language), I received valuable insights and advice on both historical and current aspects of the study of impact. My thanks, too, to Dr Dianne Wall, also at my Lancaster meetings, for sharing her considerable and continuing experience in the study of test impact.

The second group of individuals whose help and support I should like to acknowledge played key roles during the implementation stages of the studies. Dr Raffaele Sanzo, Co-ordinator of the *PL2000* in the Ministry of Education, Italy, not only helped to facilitate the Cambridge ESOL Impact Study, but offered encouragement and wise insights throughout its implementation. Richard Turner, of Cambridge ESOL Research and Validation, who held an expert camcorder during the video-ing of IELTS impact study preparation classes, also played a major role in the data management and analysis for both IELTS and *PL2000* studies. Liam Vint, Cambridge ESOL Country Manager for Italy, turned *PL2000* Impact Study camera person, too, in addition to providing vital contacts and background knowledge throughout planning and implementation stages. Lee Knapp, Cambridge Development Manager, UK, accompanied me on some of my IELTS impact study visits to UK centres and provided me with information and views from his many other IELTS-related contacts. My thanks also to Tony Green, Validation Officer at Cambridge ESOL, who shared data and validating operations with me for the IELTS study. I should like in addition, to express my gratitude to Susan Chapman, of Cambridge ESOL Research and Validation, who gave invaluable administrative support to both studies, and to Rowena Akinyemi and Sally Downes, who co-ordinated the final editing and production of the book.

I am grateful to those commissioned to read and make suggestions on the manuscript, namely Alan Davies, Emeritus Professor, Department of Theoretical and Applied Linguistics, University of Edinburgh and Peter Falvey, Consultant to Cambridge ESOL. To Nick Charge, Subject Manager for IELTS and Liam Vint, my thanks for reading and advising on parts of the book of particular relevance to their work. May I acknowledge, too, Dr Mike Milanovic, CEO of Cambridge ESOL, and Cyril Weir, Professor of English Language Acquisition at the University of Luton, for guidance, beyond their duties as SiLT series editors, on content and emphasis.

Finally I should like to acknowledge the constant interest in the project taken by Nick Saville and Lynda Taylor, Director and Assistant Director respectively, of the Cambridge ESOL Research and Validation Group. Their wise and helpful advice was based both on their interest in and readings of this book.

Last, and in some ways most, I should like to recognise all the many participants in the studies described in this book. To the students and their parents, the test takers, the teachers, the school and centre heads, the parents, the administrators and managers, including Cambridge ESOL Local Secretaries, involved in IELTS and the *Progetto Lingue 2000*, thank you for your co-operation and helpfulness, for your contribution to the good cause of impact study.

Series Editors' note

Cambridge has always taken a great interest in the educational and social impact of its tests and assessments. For many years while testing experts around the world were preoccupied with the quantitative aspects of assessment worrying about impact was not considered to be particularly relevant. The Cambridge attention to this area was probably even thought of as slightly quirky and old fashioned. However, in recent years, the concept of identifying and measuring how tests impact on the environment in which they operate has been recognised as a very relevant concern. Indeed, three further volumes in this series by Liying Cheng, who looks at washback in Hong Kong, Diane Wall, who documents an impact study in Sri Lanka, and Tony Green, who focuses on IELTS, demonstrate the growing importance of impact research as an aspect of test validation. Languages in general and English in particular, are of ever growing importance, not only for economic reasons, but also for social and political ones. Stakeholders in the language assessment process increasingly require evidence on the interactions between examinations, the stakeholders involved and the outcomes expected. The effective conceptualisation of the dimensions of test impact and its systematic study – within the context of test validation research – is one of the ways that will help us to address this requirement better.

This volume is written from the perspective of an international language testing agency although the issues discussed are of relevance in national and local assessment situations. Roger Hawkey, who has now conducted extensive work in the area of test impact, considers its dimensions and why understanding test impact is important. After some discussion of the concepts of impact and washback and how they fit into a broader educational, research and social context, he looks at the role of impact studies in the Cambridge ESOL test development, validation and revision systems, with particular reference to the *Progetto Lingue 2000* in Italy and the study of IELTS impact.

In the fields of language teaching and testing, the concepts of washback and impact, as Hawkey explores in some depth, are a matter of both theoretical and practical differentiation and concern. Through the 1980s and into the early 1990s attention focused on the concept of test washback and as such took a relatively narrow view, focusing largely on the teaching–learning relationship with some attention paid to the role of publishers and course materials. But beyond the learners and teachers affected by the washback of a language test

are a range of other stakeholders on whom an examination has impact even though they do not take the test or teach it. These stakeholders, for example, parents, employers, university admissions officers and others, form what we might refer to as the language testing constituency. Cambridge ESOL has defined this constituency particularly in relation to candidates taking its own examinations but the definition applies in other contexts too. Different tests will have different constituencies and an examination board like Cambridge ESOL will be dealing with numerous and varied constituencies, quite possibly for the same test and at the same time. The stakeholders interact with the test construct, format, conditions and assessment criteria in various ways.

Cambridge ESOL routinely conducts impact studies as part of the test validation process on an ongoing basis. It is our view that an examination board must be prepared to review and revise what it does in the light of how its stakeholders use and feel about its examinations. As educational processes and social needs change it is vital that examinations adapt to meet the requirements of these changes and the study of test impact facilitates this process even if the interrelationships involved are complex and highly context-dependent. Hawkey rightly points out that impact research is an exemplification of the growing importance of evidence-based approaches to education and assessment. Evidence-based education requires policy and practice capable of being justified in terms of sound evidence about their likely effects. Given that education, or indeed assessment, is not an exact science, it is too important to allow it to be determined by unfounded opinion, whether of politicians, teachers, researchers or anyone else. Validation research, including research into test impact, aims to seek out the evidence necessary to develop, redevelop or indeed operate a testing system in an appropriate and ethical manner.

The role of ethics in language testing has risen to the fore much more significantly in the last decade or so. The intention of those concerned with ethical language testing is to implement codes of professionally and socially responsible practice. These codes should provide tighter yet feasible guarantees of test development rigour and probity, with properly-defined targets, appropriate and reliable evaluation criteria, comprehensive, transparent and fair test interpretation and reporting systems, continuous validation processes, and a keener regard for the rights of candidates and other stakeholders (for example, see the ALTE Code of Practice, the ALTE quality assurance work, and the IELTS Handbook and Annual Review).

An ethical approach to language testing is a must in the modern world and test impact studies play an important role in demonstrating that language tests are used ethically. However, impact studies can also help address some of the concerns raised by the critical language testing lobby. The critical language

testing movement characterises tests as, intentionally or not, biased, undemocratic, and unfair means of selecting or policy-changing. It is argued that the main actual impact of language tests is the imposition of constraints, the restriction of curricula, and the possible encouragement of boring, mechanical teaching approaches.

Whether this is the case or not needs to be a matter of research rather than opinion and in such a general context a focus on test impact is an important area of study. It is driven by considerations in the field of language testing of wanting to do the job right and providing the appropriate evidence to back any claims. It is also driven by a broader social, political, educational and even cognitive impetus, and we see again the growing movement in education to develop the notion of basing what we do on sound evidence. Indeed, the evidence-based education manifesto argues that we need a culture in which evidence is valued over opinion, and where appropriate action (or inaction) is valued over action for the sake of being seen to do something. This applies just as much to what critical language testers have to say as it does to the claims of examination boards, education departments, schools and so on.

This volume is intended to provide the reader with an approach to the study of test impact which allows evidence to be gathered and displayed. It documents in some detail aspects of two impact studies that have been conducted in the Cambridge context and as such, we believe it makes a unique and much needed addition to the field. Its focus on the use of international assessments in state systems in the *Progetto Lingue 2000* is relevant as English becomes a core subject in many countries around the world, and it is vital that there is a good understanding of what impact international assessment may have. The focus on IELTS is no less significant as international mobility continues to increase. The extensive IELTS research takes us beyond a narrow focus on the test itself to the broader impact that it has and demonstrates very clearly that IELTS impacts positively on language learning and teaching in addition to its well known measurement attributes.

Two further volumes on IELTS will be published soon after this volume. The first, entitled *IELTS Collected Papers: Research in speaking and writing assessment* and edited by Lynda Taylor and Peter Falvey, documents a range of research studies with a particular focus on speaking and writing. The second, written by Alan Davies and entitled *Assessing Academic English: Testing English proficiency 1950–2005 – the IELTS solution*, documents the development of the testing of academic English from the 1950s to the present day.

Michael Milanovic
Cyril Weir
2005

Contents

| | |
|--|------|
| Acknowledgements | vi |
| Series Editors' note | viii |
| | |
| Chapter 1 | |
| Impact, washback, evaluation and related concepts: definitions and examples | 1 |
| | |
| Chapter 2 | |
| Impact study approaches | 29 |
| | |
| Chapter 3 | |
| Impact study objectives, designs and research questions | 41 |
| | |
| Chapter 4 | |
| Impact study instrumentation | 54 |
| | |
| Chapter 5 | |
| Collecting and managing the data | 85 |
| | |
| Chapter 6 | |
| The IELTS impact study: main messages | 99 |
| | |
| Chapter 7 | |
| The <i>PL2000</i> Impact Study: main messages | 133 |
| | |
| Chapter 8 | |
| Related developments, further research and lessons learned | 162 |
| | |
| Appendices | 179 |
| | |
| References | 215 |
| | |
| Index | 225 |

1 Impact, washback, evaluation and related concepts: definitions and examples

Impact is the main topic of this book, so this opening chapter will attempt to clarify the concept and its implications. It will also consider related terms, for example *evaluation*, *monitoring* and *washback*.

Context for the discussion throughout will be via reference to the role of impact studies in the test development and validation systems of the University of Cambridge English for Speakers of Other Languages (ESOL) Examinations.

This book will refer to two Cambridge ESOL impact study projects in particular.

One is the study of the impact of the International English Language Testing System (IELTS), an examination ‘designed to assess the English language ability of people whose first language is not English and who need to study, work or live where English is used as the language of communication’ (www.ielts.org). The second is a study of the impact of the *Progetto Lingue 2000* (Year 2000 Languages Project), a Ministry of Education, Italy, state school foreign language education improvement programme.

This should set the scene for Chapter 2, which considers different approaches to the collection and analysis of impact data, and Chapter 3, on the definition of research objectives and questions. Chapter 4 then traces the development of impact study instrumentation, and Chapter 5 the collection, management and analysis of data. In Chapters 6 and 7, some of the main findings of the studies into IELTS and the *Progetto Lingue 2000* impacts are presented, in their own right and as examples of the outcomes that may be expected from research into the foreign language learning and testing aspects of educational impact. Chapter 8 traces research and other developments related to the two studies, considers lessons to be learned, and suggests approaches for the continuing study of educational impact.

But first some key *terms* need to be defined.

Impact in educational research

Impact of process and product

Taking an *educational evaluation* viewpoint, Weiss defines *impact* as ‘the net effects of a programme (i.e. the gain in outcomes for program participants minus the gain for an equivalent group of non-participants)’ (1998: 331). She then broadens this somewhat narrow definition by adding that ‘impact may also refer to program effects for the larger community’, and admitting that ‘more generally it is a synonym for *outcome*’ [all italics mine]. This wider view of the impact construct is reflected in a definition from *developmental education* studies:

Impacts (also referred to as effects) may be planned or unplanned; positive or negative; achieved immediately or only after some time; and sustainable or unsustainable ... Impacts may be observable/measurable during implementation, at project completion, or only some time after the project has ended. Different impacts may be experienced by different stakeholders (Department for International Development (DFID) Glossary of terms 1998).

Note that this definition of impact appears to include a focus on *processes* as well as *outcomes* or product, a distinction often at issue in impact and evaluation studies. Roy defines process and product studies as follows:

A study of the product is expected to indicate the pay-off value while a study of the process is expected to indicate the intrinsic values of the programme. Both are needed, however, to find the worth of the programme (1998:71).

Weiss defines a process focus more straightforwardly as the study of ‘what goes on while a program is in progress’, whereas outcome studies measure and describe the ‘end results of the program’ (1998:334–335).

In the field of education, *impact studies* most commonly focus on the effects of interventions, including both teaching programmes and tests, on the people participating in them in various ways. Given the formative nature of education and learning, such studies seek to measure and analyse both outcomes, for example test results or subsequent performance on the criteria the test is measuring, and processes, for example the learning and teaching approaches and activities of programmes preparing candidates for a test.

The study of the impacts of the IELTS test is, by definition, a form of summative evaluation, concerned with outcomes such as candidate test performances. But a test such as IELTS, used as an English language qualification for academic studies in English-speaking countries and for immigration, training and employment purposes, also has significant potential impact on processes such as the ways candidates learn and prepare for the test

itself and for their English language activities beyond it (there is further discussion on this below). There are thus formative aspects (intended to provide information to improve programmes or tests) as well as summative aspects to impact studies. As for the *Progetto Lingue 2000 (PL2000)* Impact Study, of a Ministry of Education project for the improvement of language learning in the state sector, there is a focus on developments in areas where the Project is intended to have influence. These include, of course, teaching/learning processes and foreign language performance outcomes. Impact studies of tests, like impact studies of learning programmes, are likely to be process- as well as product- or outcome-oriented.

Impact studies and evaluation studies

Varghese contrasts *impact* studies with *evaluation* studies, the latter, he feels, tending to focus more closely on the immediate objectives of projects rather than their longer-term development.

... An evaluation of adult literacy programmes may indicate the total number of persons made literate by the programme. An impact study of the programme will focus on the social implications of the outcomes ... It will also ask, for example, whether the reading habits of the community improved (1998:49).

Varghese (49–50) reminds us that impacts are changes (or effects) rather than the achievement of project targets, which are often seen as the focus of evaluation studies.

Weiss defines evaluation as the ‘systematic assessment of the operation and/or outcomes of a program or policy, compared to explicit or implicit standards, in order to contribute to the improvement of the program or policy’ (1998:330). The term ‘evaluation’ refers to an overall process; an evaluation study is, after all, an exercise to appraise (that is, measure the value of) an educational programme. Impact may well be *one of the areas* of the programme covered by an evaluation. So, evaluation and impact are linked, with evaluation in some cases tending to include impact, in the sense of programme effects which evaluators want to find out about as part of their evaluation. Why? In order to make proposals to adjust these effects to ‘contribute to the improvement of the program or policy’ (see Weiss above).

But the evaluation study literature (e.g. Agar 1986, Connell et al 1995, Cronbach 1982, MacDonald 1974, Parlett and Hamilton 1972, Weiss 1998) warns us regularly that the nature of evaluation should not be over-simplified as it was following Scriven’s 1967 contrast between summative and formative evaluation. ‘Evaluations that focus on outcomes’ says Cronbach, ‘can and should be used formatively’ (1982). Parlett and Hamilton go further with their concept of ‘illuminative evaluation’ (1972). This has a primary concern ‘with

description and interpretation rather than measurement and prediction', with how innovation operates, 'how it is influenced by the various school situations in which it is applied; what those directly concerned regard as its advantages and disadvantages; and how students' intellectual tasks and academic experiences are most affected'. Illuminative evaluation 'aims to discover and document what it is like to be participating in the scheme' (see Murphy and Torrance (eds.) 1987:60–61). Levine's concept of evolving curriculum is similarly dynamic, 'where evaluation is an inherent aspect of the curriculum planning process (*evaluation in planning*)' with 'the evaluation process itself a perpetual and self-developmental inquiry process (*evaluation as planning*). The curriculum evaluation process that emerges is flexible, yet methodical, open, yet directive, *and* respectful of the diverse, complex curricular visions, needs and constraints encountered in schools and classrooms' (2002:26).

There would seem to be much to learn from these definitions. The impact studies discussed in this book attempt to combine 'description and interpretation' with 'measurement and prediction'. They seek to investigate the influences of 'the various school situations' in which the IELTS test is prepared for and the principles of the *PL2000* are put into practice. They certainly seek to discover 'what those directly concerned' regard as the 'advantages and disadvantages' of the test and the curriculum reform project, and also how students' 'intellectual tasks and academic experiences are most affected'. We shall also see, throughout this book, that the study of the impact of language tests or programmes, like the evaluation process, tends to be 'perpetual and self-developmental' rather than single and monolithic.

Monitoring

Then there is the term *monitoring*, clearly related to both impact and evaluation, and actually suggested by Italian colleagues participating in the study of the impact of the *PL2000*, as a synonym for impact study. Weiss defines monitoring as '[a]n ongoing assessment of program operations conducted during implementation, usually by sponsors or managers, to assess whether activities are being delivered as planned, are reaching the target populations, and are using resources appropriately' (1998:333). Judging by this definition, there is considerable overlap between monitoring and evaluation, but the fact that monitoring takes place only *during* the implementation of a programme may distinguish it.

A further distinction, suggested by Lynda Taylor (2005, personal communication) sees monitoring as primarily descriptive in function, followed by evaluation, which is, naturally, mainly evaluative in function. As will emerge from this chapter (see Figure 1.5 on page 20) the Cambridge

ESOL model of test development includes, following the establishment of the need for a new or revised test, the stages of:

- design and initial specification
- development through trialling, analysis, evaluation and review
- *monitoring*, mainly through routine descriptive data for analysis, until a decision is made, based on particular monitoring information
- reviewing and *evaluating* the test for possible further revision.

Weiss links monitoring with 'process evaluation', and adds a participant dimension:

... process evaluation is not very different from what is often called monitoring. One key difference is that monitoring is done primarily on behalf of the funders and other high-level officials to hold the program to account (1998:181).

One teacher/administrator participant in the study of IELTS impact seemed to sense that impact studies may be less top-down and judgemental when she described them as more 'user-friendly' than evaluations or monitoring.

In the case of the *PL2000* Impact Study there was, of course, no question of the impact study 'funders', Cambridge ESOL, holding policy-makers, designers, managers or officials of the *PL2000* to account. Rather, the examinations board was concerned with the two-way impacts (see below) of Cambridge exams on participants in the *Progetto*, and of the *Progetto* on these exams. The *PL2000* Impact Study was carried out by Cambridge ESOL as an interested party, selected alongside other international test providers (see below), to provide external certification for students who had been participating in foreign language courses under the *PL2000*. Cambridge ESOL was not the initiator or leader of the foreign language reform project itself; its role, through the *PL2000* Impact Study which it ran with Ministry approval, was to describe impacts rather than to evaluate the Project.

Insider and outsider roles

The question of the 'evaluator-user and insider-outsider interface' is often at issue in the evaluation literature. O'Dwyer summarises as follows:

Evaluators may remain distant and report findings in their own way with the expectation that these may be used to improve a program; or, may be actively involved in the program, working hand-in-hand with those in a program, or stakeholders to the program, with a view to specifying the evaluation focus according to the needs of the users. The profile which an external evaluator may adopt, therefore, could be of a complete outsider to a program, or, towards the other end of the spectrum, of a close 'insider' in relationship to the clients (2005).

The differences in evaluator roles described here would appear to apply to impact studies as well as evaluations. With the design, trialling and implementational phases of the study of IELTS impact, which will be described in detail in Chapters 3–6 below, the outsider-insider roles included, at various stages, both external consultant teams and individuals commissioned by Cambridge ESOL, and validation and implementation expertise from within the organisation. Cambridge ESOL is, of course, one of the three partners in the IELTS test, along with the British Council and IDP Education Australia : IELTS Australia. Both these latter partners are fully informed of the impact studies and themselves contribute to research in support of IELTS through the IELTS funded-research programme (see this chapter and Chapters 4, 6 and 8). In the *PL2000* Impact Study, it will be seen, relationships between the impact study team and participants such as the case study school teachers and heads were close, though not quite ‘insider’.

Impact and washback in foreign language teaching and testing

In this section of the chapter, the concepts of evaluation, monitoring and impact are investigated within the fields of language teaching and testing, where similarities with and distinctions from the general educational literature will be discovered.

In the language teaching and testing literature, the concept of impact as effects or changes still stands but the term co-occurs frequently with the term ‘washback’ (or ‘backwash’) and it is the distinction between the two that is often an issue of debate. In the context of studies of the effects of language programmes or tests on those involved, the concepts of impact and washback/backwash are often considered in terms of their:

- logical location
- definition and scope
- positive and negative implications
- intentionality
- complexity
- direction
- intensity, emphasis
- stakes and stakeholders
- relationships with validity and validation
- relationships with the Critical Language Testing view
- role in impact/washback models.

This chapter attempts below to cover all these aspects of impact and washback.

Washback and impact

‘Washback and the impact of tests more generally has become a major area of study within educational research’ Alderson (2004a:ix) and as the washback and impact net widens, so does the need for agreed labels for the kinds of study we carry out to investigate the effects of tests or programmes *in and beyond the classroom context*. Hamp-Lyons summarises the situation and the terminology well. She finds that Alderson and Wall’s ‘limitation of the term ‘washback’ to influences on teaching, teachers, and learning (including curriculum and materials) seems now to be generally accepted, and the discussion of wider influences of tests is codified under the term ‘impact’ (Wall 1997), which is the term used in the wider educational measurement literature’ (2000:586). In similar vein, Bachman and Palmer 1996 refer to issues of test use and social impact as ‘macro’ issues of impact, while washback takes place at the ‘micro’ level of participants, mainly learners and teachers.

So the term ‘impact’ now appears to be used to describe studies which investigate the influences of language programmes and/or tests on stakeholders *beyond* the immediate learning programme context. An impact study might, for example, investigate the effects of a programme or test on school heads, parents, receiving institution administrators, high-stakes test providers (all these stakeholders included in the two impact studies described in Chapters 3–8 below).

Given that the term ‘impact’ is a word in everyday use in its meaning of ‘influence or effect’ (e.g. *Oxford School Dictionary*, 1994), it is unsurprising to find the term also apparently used non-technically. When Alderson (2004a: ix), for example, writes: ‘We now know, for instance, that tests will have more impact on the content of teaching and the materials that are used than they will on the teacher’s methodology’, is he using the term in its lay sense, since technically the content of teaching and the teacher’s methodology are washback rather than impact matters? Or is he acknowledging that, for some, impact, the broader construct, *includes* washback? Green notes that although ‘the terms have been used to refer to the same concept, backwash is distinguished from test impact by Bachman and Palmer (1996:30) who, with McNamara (1996, 2000), Hamp-Lyons (1998) and Shohamy (2001) place washback within the scope of impact’ (2003:6). This would presumably mean that one could use the term ‘impact’ for all cases of influence from a language test or language programme, whether on teaching and learning or on, say, a university’s admissions policy.

Andrews, writing on washback and curriculum innovation, appears to acknowledge the fragility of the washback : impact distinction:

The term washback is interpreted broadly ... the present chapter uses *washback* to refer to the effects of tests on teaching and learning, the educational system, and the various stakeholders in the education process. Where the word 'impact' occurs in this chapter, it is used in a non-technical sense, as a synonym for 'effect' (2004:37).

In this book we shall try to be consistent in the use of terms:

- to use 'washback' to cover influences of language tests or programmes on language learners and teachers, language learning and teaching processes (including materials) and outcomes
- to use 'impact' to cover influences of language tests or programmes on stakeholders beyond language learners, teachers, except when it is the influences of a test or programme on learners and teachers *outside* their learning or teaching roles, for example on their attitudes to matters beyond language learning; in this case the book will tend to refer to impact e.g. Research Question 4: What is the impact of IELTS on the participants who have taken the test?

In terms of these definitions, the two studies which are the focus of this book cover both washback and impact. They are called 'impact studies' because of this breadth.

Washback/backwash

Hamp-Lyons notes that washback 'is one of a set of terms that have been used in general education, language education and language testing to refer to a set of beliefs about the relationship between testing and teaching and learning' (1997:295). Another of the 'set of terms' is 'backwash', but it would appear that the terms 'washback' and 'backwash' are used interchangeably in the field. '... to clarify the distinction between the terms backwash and washback', Alderson says (2004a:xi), 'there is none'. Hughes admits that there is interchangeable use of the two terms in his work but adds, (2003:57) 'Where "washback" came from I do not know. What I do know is that I can find "backwash" in dictionaries, but not "washback"'. Cheng and Curtis choose to use the term 'washback' 'as it is the most commonly used in the field of applied linguistics' (2004:5). This book will follow suit, preferring the term 'washback' as it does now appear to be in more common use in the field.