

**Book Review:**

**IELTS washback in context:  
Preparation for academic writing in higher education**

(Studies in Language Testing 25)

by Anthony B. Green (2007)

Cambridge University Press & Cambridge ESOL

[pp xiv+386] ISBN 978-0-521-6929-2



The IELTS™ test is used by most universities in the British Commonwealth in the same way as the TOEFL® is by many North American universities, that is to say as a measure of linguistic competence for prospective students. The test consists of four modules: Listening, Speaking, Reading, and Writing. The Listening and Speaking modules are the same for all IELTS examinees, but the Reading and Writing modules exist in two different versions: General Training and Academic. Candidates aiming to enter higher education in an English-speaking country usually take the Academic modules. In 2006, 70% of the 700,000 IELTS examinees worldwide opted to take the these modules.

This book deals with the Academic Writing module (AWM) of the IELTS test as it existed in the period between 2002 and 2004, studying its washback effects. The concluding recommendations of this book helped guide the 2005 revision of the IELTS Academic Writing module.

Since 2006 a computer-based version of the test (CB-IELTS) has been introduced. Currently both the paper-and-pencil and computerized versions of this test co-exist, and they employ the same question types. However, the focus of the book is not so much the test itself as the effectiveness of different types of preparation courses. In other words, are specific IELTS preparation courses the best way to raise IELTS scores, or is a general academic writing course more effective?

The intended audience for this book is people involved in the testing of English for academic purposes (especially of writing) and also graduate students learning about washback. It is complementary to, although more specialized than, Wall and Horak's (2008) washback study of the TOEFL test or Cheng et al's volume on washback research methods (2006).

The volume consists of seven chapters and eight appendices. It opens with a thorough description of washback, contrasting various definitions. An annotated literature review then summarizes washback research over the last three decades. Academic writing is the subject of the second chapter, which discusses the characteristics of academic writing, and how these are operationalized in the IELTS Academic Writing module. The author concludes that academic writing is imperfectly represented in the test.

In the third and fourth chapters we enter the research project itself. The data collection instruments and sampling procedure for this study are described in depth. Four types of data were

employed: student test scores, questionnaire responses (from students and teachers), student focus group results, and classroom observations.

The resulting data was submitted to two types of analysis. In addition to standard data analysis methods such as analysis of covariance and correlation analysis as well as prediction modeling, a less familiar idea known as the “neural network” method, which has been used to predict language course outcomes from multiple sources of data (Hughes-Wilhelm, 1999) was implemented. The non-statistician may find this section difficult, although its goal – to discern which factors correlate with improved scores on the IELTS – is clear.

The main question this text seeks to answer is whether IELTS preparation courses or general academic writing courses do a better job of helping students get higher IELTS scores. The author's initial conclusion is that IELTS courses do produce slightly better scores, but “the additional benefit is limited” (p. 282). Moreover, detailed analysis of the data reveals that course type is not a significant predictor of outcomes; IELTS courses are not significantly better when other conditions are taken into account.

How can we account for the different student scores on the test? This is the substance of chapter six. It shows the relation between the course and learner variables on one hand, such as the type of course taken (IELTS preparation, combination IELTS and general academic writing, and pre-sessional English for academic purposes) and differences in scores on the other. For example, the strongest predictor of the final score was the initial writing score. Simply stated, was taking an IELTS preparation course effective? The chapter is a model of clarity in presentation of statistical approaches to a vast body of data. The variables are classified as presage, process and product, and the statistical procedures were either traditional (linear regression) or artificial intelligence (neural network).

To round the book off, the author then systematically presents his conclusions and their implications. The following three research questions are of particular relevance:

**Question:** *Is the EAP construct better served by IELTS courses or non-IELTS courses?*

**Answer:** General EAP courses train students to handle a wider range of tasks than IELTS courses. For example, on EAP courses students both learn about the question of plagiarism and also learn how to compile bibliographies. Results of students' self-assessed improvement in writing ability shows that more EAP students thought they had made great improvement in their writing than IELTS course students.

**Question:** *Do students on IELTS courses score differently from those on non-IELTS courses?*

**Answer:** Learners on all course types improved their writing scores significantly, but no one course type produced a better score. However, some differences between long and short courses were noted. Long courses produced more improvement among students with low initial writing scores. Strangely, for students with a high score at the start, short courses more often led to a drop in their writing scores, and long courses generally produced only small gains. Green states, “The test-score evidence seems to contradict some of the stronger claims for the value of intensive IELTS preparation made by participants” (p. 306).

**Question:** *Do learners' individual differences interact with instructional differences in predicting outcomes?*

**Answer:** Yes, instructional differences such as course length and individual differences such as region of origin are relevant. Although course content was influenced by the test, teaching and learning methods were based on previous beliefs. As for the actual test results, a meaning-based teaching approach seems to be most successful. This contrasts with an approach stressing the memorization of linguistic formulae.

After these specific recommendations are set out for all the stakeholders concerned – teachers, students, test developers and receiving institutions. For teachers the author offers three recommendations. Firstly, allow more time for EAP instruction. Traditionally the IELTS administration has suggested 200 hours of study for an improvement of one band. However, this study showed that figure was unrealistic. As the author states, “It is very clear that blanket recommendations across proficiency levels are misguided” (p. 309). The second recommendation is to introduce IELTS in the context of EAP. This means informing students of the limited degree of overlap between IELTS writing tasks and the complete range of tasks needed for university study.

The final recommendation for teachers is to inform students of relevant research findings about the IELTS to improve their background understanding of the test. For example, the limitations of an improved IELTS score should be made plain. Green cautions, “An IELTS band score at a given level does not imply that they (students) have nothing further to learn about academic writing in English” (p. 310).

### **The Bottom Line**

Those approaching this text from a pedagogical background are likely to find the introductory chapter on washback particularly enlightening. The later chapters are of specific interest to testing professionals or graduate students, since they explain a large-scale research project with exemplary clarity. Readers who are normally bewildered by statistical explanation may be pleasantly surprised at the way complex bodies of data are analyzed and explained.

Two reservations about this volume are unavoidable. The first is that the research is on a version of the IELTS test that was in use between 2001 and 2004. Since then the IELTS has been revised, although the writing section formats are unchanged. The second reservation concerns the sample size and representativeness. Compared with the 490,000 Academic Module IELTS candidates (2006 figures) the 476 participants used in this study seems like quite a small sample. Moreover, about two thirds of the participants (319) came from two universities. The population in this study seems to differ from general IELTS test takers; 72% of the participants in this study were from East Asian countries, compared with 58% of the actual IELTS candidates.

How does this volume compare with other washback texts, both introductory and specialized? For an introduction to the main ideas in washback, general readers could usefully start with Green and Hawkey (2005). Teachers might prefer to look at Spratt (2005) before examining the slightly more academic works of Bailey (1999), Alderson and Banerjee (2001), or Wall (1997).

On the other hand, for the washback specialist the most significant general book in recent years is by Chen, Watanabe and Curtis (2004). The present volume overlaps with their work in that it summarizes recent scholarship on language testing washback. However, there are clear differences between these two texts. Chen, Watanabe and Curtis give more weight to the relation between washback and curricular innovation. Their text contains relatively brief descriptions of many washback research projects from various parts of the world. However, the Green volume offers a more detailed account of research into of one specific washback case. While Cheng et al aim for breadth, Green gives us depth.

In summary, while this volume can only be seriously nominated as an essential purchase for the providers of and researchers into EAP tests, its introductory material is well worth the time of a wider audience, including general testing students, and anyone connected with large scale academic writing examinations.

- Reviewed by Daniel Dunkley,  
Aichi Gakuin University

### Works Cited

- Alderson, J. C. & Banerjee, J. (2001). Impact and washback research in language testing. In C. Elder, et al. (Eds.) *Experimenting with uncertainty: Essays in honour of Alan Davies*. Cambridge: Cambridge University Press. 150-161.
- Bailey, K. (1999). *Washback in Language Testing*. Retrieved on July 30, 2010 from <http://www.ets.org/Media/Research/pdf/RM-99-04.pdf>
- Cheng, L. Watanabe, J. & Curtis, A. (Eds.) (2004). *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum.
- Green, A.B. and Hawkey, R. (2005). Test washback and impact: What do they mean and why do they matter? *Modern English Teacher*, 13 (4), 66-71.
- Hughes-Wilhelm, K. (1999). Building an adult knowledge base: An exploratory study using an expert system. *Applied Linguistics* 20(4),425-459. DOI:10.1093/applin/20.4.425
- Spada, N. & Frohlich, M. (1995). *COLT Observation Scheme: Coding Conventions and Applications*. Sydney: National Centre for Language Teaching and Research of Macquarrie University.
- Spratt, M. (2005) Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9 (1) 5-29. DOI: 10.1191/1362168805lr152oa.
- Wall, D. (1997). Impact and washback in language testing. In C. Clapham, C & D. Corson (Eds.) *Language testing and assessment*. Amsterdam: Kluwer Academic. 291-302.
- Wall, D. & Horak, T. (2008). The impact of changes in the TOEFL examination on and learning and teaching in central and eastern Europe, Phase 2: Coping with change TOEFL. *iBT report iBT-05*. Princeton NJ: Educational Testing Service.

