

***Examining Writing: Research and Practice in Assessing Second Language Writing (Studies in Language Testing 26)***

<b>Author:</b>	Stuart Shaw & Cyril Weir (2007)		
<b>Publisher:</b>	Cambridge: Cambridge University Press		
<b>Pages</b>	<b>ISBN</b>	<b>Price</b>	
Pp. xiv + 344	978-0-521-69293-9 (paper)	\$49.00 U.S.	

A well-developed writing resource is highly valued by employers, educators and the public at large. Consequently, instruments which measure writing are coming under increasing scrutiny from the various stakeholders in assessment, not least the test-takers themselves. *Examining Writing* aims to inform readers of Cambridge ESOL's ongoing work to validate its portfolio of writing tests. The focus is almost exclusively on the Cambridge Main Suite of exams with only occasional reference to other assessment systems, but the research directions and practical outcomes are relevant to anyone interested in testing writing and construct validity in general. Although the main readership will be applied linguists and assessment professionals, teachers of exam courses will benefit from this comprehensive insight into test construction and evaluation.

Every new initiative in assessment has had to react to the Common European Framework of Reference (CEFR) project (Council of Europe, 2001), which describes language performance in a number of criterion-referenced tables. The authors acknowledge the contribution of CEFR to assessment but they believe that the scales are not well-defined enough to underpin test construction in concrete situations (p. 1). They present an alternative socio-cognitive framework containing five validity components—cognitive, context, scoring, consequential, and criterion-related—plus test characteristics. The book is accordingly neatly organized into chapters which explore the rationale of each component and its operationalization in Cambridge exams.

For most readers, cognitive validity will be the most unfamiliar concept. The

cognitive validity of a writing task is defined as "a measure of how closely it represents the cognitive processing involved in writing contexts beyond the text itself, i.e., in performing the task in real life" (p. 34). The cognitive processes discussed include macro-planning (generating ideas and identifying an audience); organization (selecting and sequencing ideas); translation (converting the internalized thinking process to language code on paper). The authors show how the Cambridge exams increase the cognitive demand with level. Take macro-planning: the writing task in the lowest-level exam, KET, requires "knowledge-telling" (p. 47), basically a straight response to very explicit input; the highest-level exam, CPE, demands very careful consideration of the appropriate genre and the expectations of the target audience.

There is no doubt that cognitive validity is a crucial and under-looked consideration in test design but it is surprising that cognition is identified primarily with strategy use. The language processing dimension, how test-takers manipulate their language resources in real-time, is hardly addressed. Even the translation process is largely conceived in behavioural terms, for example circumlocution to compensate for competence gaps (p. 40). Language form and function deserve fuller treatment in a cognitive model. A much-discussed example from applied linguistics is the role of formulaic phrases in language. There is an assumption that the more predictable language is, the larger the degree of prefabrication. Lewis (1997, p. 41) writing on the grammar/vocabulary divide comments that formulaic language has a functional purpose whereas grammar is more expressive. In other words, everyday usage is lexically driven while creative language is distinguished more by grammatical range and/or complexity. The implication for testing is that formulaic language is employed differently at different levels of cognitive attainment, which should impact task type and marking criteria. It is early days in the elaboration of cognitive validity, so it is hoped there will be research into the processing of specific language items during writing tests.

Possibly the most rewarding feature of the book is the case-studies, which report internal research projects and their impact on test design. The first case study (pp. 23-27) provides a fascinating insight into how social and ethical considerations can clash with the demands for reliable tests. Cambridge offers special arrangements for candidates with disabilities. In the case of dyslexia, scripts were exempt from spelling criteria. The appropriateness of this policy was reviewed in the light of statistics showing that 95% of such requests came from a single European country (the largest Cambridge exam, FCE, is taken in about 100 countries). The eventual decision to change the policy for dyslexic candidates was informed by a combination of medical advice, statistical analysis comparing the scripts of dyslexic and non-dyslexic candidates, and a special needs consultant's sociocultural perspective on the country supplying most of these candidates. At present, dyslexic test-takers have an additional time allowance but no dispensation from the standard spelling criteria: a clearly equitable outcome based on very thorough investigation of a sensitive area.

Shaw and Weir do not presume to supply all the answers to key questions in test validation. A case in point is expectations of vocabulary content at different testing levels. Corpora can establish vocabulary frequency and, intuitively, use of less frequent lexis should increase with proficiency. However, a case study

analysing lexical range across the Cambridge exams (pp. 98-104) does not bear this out. There was no clear relationship found between frequency and test level. Amazingly, KET candidates used a higher percentage of low-frequency words (defined as outside the top 2000 words of English) than CPE students, although of course they wrote much shorter texts. The results indicate that "the quantitative measures [of vocabulary] currently available continue to struggle to adequately describe test-taker output" (p. 104). What seems to matter more than what words are used is *how* words are used, that is, style and pragmatics. This conclusion fits into a growing concern in corpus linguistics (for example, Jarvis et al., 2003) that numerical counts fail to capture the real effect of language in context.

Anyone with a remote interest in testing and teaching writing should read this book. The research agenda which Cambridge ESOL has identified and massively contributed to will have resounding benefits for a society ever more dependent on adequate tests. The only shortcoming is that the assessing in the title is a little misleading because the subject matter does not go beyond testing. There are reliable alternatives to formal testing, which in the case of young learners especially are arguably more valid (see McKay, 2006), so it is a shame that space does not allow assessment to be addressed more fully. However, this is merely a limitation of scope, and it does not detract from the book's success in providing a scholarly yet highly readable account of the exciting developments in the field of testing second-language writing.

## References

Council of Europe, Modern Languages Division, Strasbourg. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Jarvis, S. Grant, L., Bikowski, D. & Ferris, D. (2003). Exploring multiple profiles of highly rated learner compositions. *Journal of Second Language Writing, 12*, 377-403.

Lewis, M. (1997). *Implementing the lexical approach*. Hove: Language Teaching Publications.

McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.

Wayne Rimmer  
University of Reading, UK  
<w.m.rimmer@reading.ac.uk>

© Copyright rests with authors. Please cite TESL-EJ appropriately.