doi:10.1016/j.jslw.2007.03.001

Much of the writing research on generation 1.5 students has focused on college students in educational contexts, especially either freshman composition or college ESL writing classes. Relatively little is known about them in K–12 settings, especially high-school students, despite their growing presence in these settings. In addition, there is a lack of understanding of these students’ literacy experiences beyond the classroom. This paper discusses a case study that examined one Korean high school student’s composing practices outside of school. The findings revealed several significant characteristics in her voluntary composing practices: the diversity and richness of her involvement with multiple literacies; her swift and constant movement among multiple genres, contexts, and languages; the public nature of her composing (sharing her writing with peers); her primary focus on L1 writing activities; and her construction of a writerly identity. In presenting her story, the paper attempts to build understanding of out-of-school writing as experienced by students with immigrant backgrounds.

http://www.elsevier.com

07–603 ZHU, YUNXIA (U Queensland, New Zealand; zyunxia@unitec.ac.nz), Understanding sociocognitive space of written discourse: Implications for teaching business writing to Chinese students. *International Review of Applied Linguistics in Language Teaching* (Walter de Gruyter) 44.3 (2006), 265–285. 
doi:10.1515/IRAL.2006.012

Confronted with various issues in teaching business writing to Chinese students in New Zealand, this paper sees the need for bridging the gap between genre-based research and teaching in an intercultural context. Specifically, it develops an intercultural reflective model in the light of Bhatia’s sociocognitive genre study as well as cross-cultural persuasion. As an important part of the model, New Zealand and Chinese experts’ intracultural and intercultural reflections on business writing are solicited and compared and the theoretical implications for teaching and learning business writing are discussed. It has been found, through a case study of analysing English and Chinese business faxes, this model can offer an in-depth understanding about discursive competence across cultures, and provide a link between genre-based theory, teaching practice and professional expertise.

http://www.degruyter.de

doi:10.1177/0265532207071510

In this article, I describe a practical application of the Roussos and Stout (1996) multidimensional analysis framework for interpreting group performance differences on an ESL reading proficiency test. Although a variety of statistical methods have been developed for flagging test items that function differentially for equal ability examinees from different ethnic, linguistic, or gender groups, the standard differential item functioning (DIF) detection and review procedures have not been very useful in explaining why DIF occurs in the flagged items (Standards for Educational and Psychological Testing 1999). To address this problem, Douglas, Roussos and Stout (1996) developed a confirmatory approach to DIF, which is used to test DIF hypotheses that are generated from theory and substantive item analyses. In the study described in this paper, DIF and differential bundle functioning (DBF) analyses were conducted to determine whether groups of reading test items, classified according to a bottom-up, top-down reading strategy framework, functioned differentially for equal ability Arabic and Mandarin ESL learners. SIBTEST (Stout & Roussos 1999) analyses revealed significant systematic group differences in two of the bottom-up and two of the top-down reading strategy categories. These results demonstrate the utility of employing a theoretical framework for interpreting group differences on a reading test.

http://ltj.sagepub.com


Conversation schools usually separate their classes according to ability levels so that a learner’s potential for learning matches the instructional demand and complexity of content of the class. Accurate and efficient placement testing helps to facilitate this process. However, off-the-shelf placement tests rarely match the syllabus of a particular conversation school. This article provides a practical model for constructing a valid and
Language testing

reliable criterion-referenced placement test that matches
the syllabus of a particular conversation school and
which can be administered by both English-speaking
and non-English speaking staff or instructors. The
method of constructing the placement test follows the
process the author actually underwent and is intended to
provide a practical model for other conversation schools.

http://jalt-publications.org/tlt

doi:10.1177/0265532207076363

Results from the Ontario Secondary School Literacy Test (OSSLT) indicate that English as a Second
Language (ESL) and English Literacy Development
(ELD) students have comparatively low success and
high deferral rates. This study examined the 2002 and
2003 OSSLT test performances of ESL/ELD and non-ESL/ELD students in order to identify and
understand the factors that may help explain why
ESL/ELD students failed the test at relatively high
rates. The analyses also attempted to determine if
there were significant and systematic differences in
ESL/ELD students’ test performance. The performance
of ESL/ELD students was consistently and similarly
lower across item formats, reading text types, skills and
strategies, and the four writing tasks. Using discriminant
analyses, it was found that narrative text type, indirect
understanding skill, vocabulary strategy of reading, and
the news report writing task were significant
predictors of ESL/ELD membership. The results of
this study provide direction for further research and
instruction regarding English literacy achievement for
these second language students within the context of
having to complete large-scale English literacy tests
designed and constructed for first English language
students.

http://ltj.sagepub.com

07–608  **DÁVID, GERGELY** (Eötvös Loránd U,
doi:10.1177/0265532207071512

Some educational contexts almost mandate the application of multiple-choice (MC) testing techniques,
even if they are deplored by many practitioners in the
field. In such contexts especially, research into how well
type of MC item perform and how their performance
may be characterised is both appropriate and desirable.
The focus of this paper is on a modified type of MC item
dubbed ‘multitrak’, which was used to test grammar
in the Hungarian national admissions test for English
between 1993 and 2004. Multitrak items were evaluated
in relation to three better-known MC item types
and were shown to provide more information about
candidates in the ability range represented by Levels
B2 and C1 on the Common European Framework of
Reference (CEFR), as well as allowing a focus on more
difficult content than the other item type in the study.

http://ltj.sagepub.com

07–609  **ELDER, CATHERINE** (U Melbourne,
Australia; caelder@unimelb.edu.au), **GARY BARKHUZEN, UTE KNOCH & JANET VON RANDOW**, *Evaluating rater responses to an online training
doi:10.1177/0265532207071511

The use of online rater self-training is growing in
popularity and has obvious practical benefits, facilitating
access to training materials and rating samples and
allowing raters to reorient themselves to the rating
scale and self-monitor their behaviour at their own
convenience. However there has thus far been little
research into rater attitudes to training via this modality
and its effectiveness in enhancing levels of inter-
and intra-rater agreement. The current study explores these
issues in relation to an analytically-scored academic writing task designed to diagnose undergraduates’ English learning needs. Eight ESL raters scored a number of pre-rated benchmark writing samples online and received immediate feedback in the form of a discrepancy score indicating the gap between their own rating of the various categories of the rating scale and the official ratings assigned to the benchmark writing samples. A batch of writing samples was rated twice (before and after participating in the online training) by each rater and Multifaceted Rasch analyses were used to compare levels of rater agreement and rater bias (on each analytic rating category). Raters’ views regarding the effectiveness of the training were also canvassed. While findings revealed limited overall gains in reliability, there was considerable individual variation in receptiveness to the training input. The paper concludes with suggestions for refining the online training program and for further research into factors influencing rater responsiveness.

http://ltj.sagepub.com

In order to motivate university students to improve their English proficiency, the Hong Kong government decided to adopt a common exit English language test for all graduating students. In the process of selecting a suitable measure for this purpose, an empirical study with a sample of over 240 students was conducted to compare two English proficiency tests, the English Test of the Graduating Students’ Language Proficiency Assessment (GSLPA) and the Academic Version of the International English Language Testing System (IELTS). The comparisons covered both speaking and writing components of the two tests, based on information provided by the two tests as well as performances of the participating candidates on the components under study. Results indicate: (1) that GSLPA writing and speaking scores distinguish candidates’ abilities more clearly than the corresponding scores on the IELTS but IELTS overall scores, generated from writing, speaking, reading and listening sub-scores, have a discriminating power similar to that of GSLPA; (2) that the GSLPA and IELTS writing subtests basically measure different skills; and (3) that the constructs of the GSLPA and IELTS speaking subtests, while having their own distinctive features, overlap by about 48%. This paper recommends options for improving the current assessment framework for graduating university students and discusses the possible impact of such a test on English language teaching and learning.

http://rel.sagepub.com

http://ltj.sagepub.com

In the componential approach to modelling reading ability, a number of contributory factors have been empirically validated. However, research on their relative contribution to explaining performance on second language reading tests is limited. Furthermore, the contribution of knowledge of syntax has been largely ignored in comparison with the attention focused on vocabulary. This study examines the relative contribution of knowledge of syntax and knowledge of vocabulary to L2 reading in two pilot studies in different contexts—a heterogeneous population studying at the tertiary level in the UK and a homogenous undergraduate group in Japan—followed by a larger main study, again involving a homogeneous Japanese undergraduate population. In contrast with previous findings in the literature, all three studies offer support for the relative superiority of syntactic knowledge over vocabulary knowledge in predicting performance on a text reading comprehension test. A case is made for the robustness of structural equation modelling compared to conventional regression in accounting for
the differential reliabilities of scores on the measures employed.
http://ltj.sagepub.com


The effect of self-monitoring on an achievement test in an EFL classroom setting was investigated. The results of an experimental study that measured the difference in class achievement test scores between a self-monitoring and external-monitoring group are reported. Participants were 114 first-year students in a Japanese university and college. Target language performance and behaviours were monitored and responses recorded on a protocol by the students in the self-monitoring group and monitored by the instructor in the external-monitoring group. The time required for implementing the protocol or treatment was also measured and was considered to be an important pedagogical factor. At an alpha level of .05, the difference in achievement scores between the groups was not statistically significant.

http://jalt-publications.org/jj

07–614 **Xi, Xiaoming** (Educational Testing Service, USA; xxii@ets.org), *Evaluating analytic scoring for the TOEFL® Academic Speaking Test (TAST) for operational use*. **Language Testing** (Sage) 24.2 (2007), 251–286. doi:10.1177/0265532207076365

This study explores the utility of analytic scoring for TAST in providing useful and reliable diagnostic information for operational use in three aspects of candidates’ performance: delivery, language use and topic development. One hundred and forty examinees’ responses to six TAST tasks were scored analytically on these three aspects of speech. G studies were used to investigate the dependability of the analytic scores, the distinctness of the analytic dimensions, and the variability of analytic score profiles. Raters’ perceptions of dimension separability were obtained using a questionnaire. It was found that the dependability of analytic scores averaged across six tasks and double ratings was acceptable for both operational and practice settings. However, scores averaged across two tasks and double ratings were not reliable enough for operational use. Correlations among the analytic scores by task were high but those between delivery and topic development were lower. These results were corroborated by raters’ perceptions. When averaged across tasks or task types, correlations among the analytic scores were very high, and the profiles of scores were flat. The utility of analytic scoring is discussed, considering both score dependability and whether analytic scores provide diagnostic information beyond that provided by holistic scores.

http://ltj.sagepub.com


This paper argues for the expansion of the knowledge base of TESOL teacher education to integrate greater awareness of the sociocultural and political context of teaching English to speakers of other languages. It is argued that the changing roles of teachers, insights gained from classroom research and recent developments in critical applied linguistics, coupled with the inherent complexity of language teaching, necessitate this curricular modification. The paper outlines three areas of inquiry to be integrated into existing TESOL teacher education curricula: (1) discussion of crosscultural variation in language teaching and learning and tools for investigating this variation; (2) overview of management of pedagogical innovation; and (3) examination of the sociopolitical factors surrounding the teaching of English as an international language. Appropriate junctions in current TESOL teacher education curricula that will allow for the integration of these three areas are suggested throughout the paper, along with resources that can aid teacher educators and curriculum developers in educating more socioculturally and politically aware teachers.

http://www.multilingual-matters.net

07–616 **Gonçalves, Maria de Lurdes** (U Aveiro, Portugal) & **Ana Isabel Andrade**, *Connecting languages: The use of the portfolio as a means of exploring and overcoming frontiers within the curriculum*. **European Journal of Teacher Education** (Routledge/Taylor & Francis) 30.2 (2007), 195–213. doi:10.1080/02619760701274001

The Linguistic European Policy stresses the importance of linguistic diversity and encourages students to learn several languages, believing that the more languages we know, the better we can understand each other and preserve our languages. To preserve diversity and to value all the languages (either included in the curriculum or not) is a first step to develop a plurilingual and intercultural competence in the classroom as a