

individually by two trained raters; inter-rater reliability (using the Cronbach- α test) was found to be 0.89 for the overall scores and between 0.77 and 0.87 for the individual bands. The paper discusses the value of using an instrument such as WATCH to assess students' choices of websites, as well as potential difficulties faced by instructors when evaluating these sources.

<http://www.elsevier.com>

07-448 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/journals>

Language testing

doi:10.1017/S0261444807004405

07-449 BARBER, RICHARD (Dubai Women's College, UAE), **A practical model for creating efficient in-house placement tests.** *The Language Teacher* (Japan Association for Language Teaching) 31.2 (2007), 3-7.

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 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>

07-450 CHANG, YUH-FANG (National Chung Hsing U, Taiwan), **On the use of the immediate recall task as a measure of second language reading comprehension.** *Language Testing* (Hodder Arnold) 23.4 (2006), 520-543. doi:0.1191/0265532206lt340oa

The immediate written recall task, a widely used measure of both first language (L1) and second language (L2) reading comprehension, has been advocated over traditional test methods such as multiple choice, cloze tests and open-ended questions because it is a direct and integrative assessment task. It has been, however, criticized as requiring memory. Whether and how the requirement of memory biases our understanding of readers' comprehension remains unexplored. This study compares readers' performance on the immediate recall and a translation task in order to explore the effect of memory on readers' recall. Ninety-seven college students participated in this study. All participants were native speakers of Mandarin Chinese whose ages ranged from 20 to 22. The results showed that the translation task yielded significantly more evidence of comprehension than did the immediate recall task, which indicates that the requirement of memory in the recall task hinders test-takers' ability to demonstrate fully their comprehension of the reading passage. The results also showed that the significant difference found in learners' performance between the immediate recall and the translation task spanned the effect of topics and proficiency levels.

<http://arnoldpublishers.com/journals>

07-451 HYUN-JU, KIM (U Seoul, Korea), **World Englishes in language testing: A call for research.** *English Today* (Cambridge University Press) 22.4 (2006), 32-39. doi: 10.1017/S0266078406004068

Native speakers of English are a minority; there are far more non-native speakers in the world (cf. Kachru 1997, Pennycook 2001). In addition, native speakers' standard or 'correct' English, in terms of its grammar and phonology, is not always useful or even appropriate in international contexts (cf. Gisborne 2000, Newbrook 1998, Shim 1999). However, despite global changes in the use of the language, the norms for ENL (English as a Native Language) remain dominant, most notably for the assessment of oral proficiency. Yet it is a major deficiency in the use of international oral tests that the proficiency of non-native speakers is measured against unrealistic and irrelevant standards (cf. Jenkins 1996). The present paper focuses on the need to revisit the testing of English oral proficiency for non-native

speakers, bearing in mind that English is used for world-wide communication and that being able to understand one another (cf. McKay 2002) is the most important goal.

http://journals.cambridge.org/jid_ENG

07-452 MAHON, ELIZABETH A. (Durham Public Schools, North Carolina, USA), **High-stakes testing and English language learners: Questions of validity.** *Bilingual Research Journal* (National Association for Bilingual Education) 30.2 (2006), 479-497.

The purpose of this study was to understand relationships between English proficiency and academic performance for a group of English language learners (ELLs) from 4 elementary schools. Descriptive and inferential statistics were used to examine scores from the Language Assessment Scales, the Woodcock-Muñoz Language Survey, and the Colorado Student Assessment Program. Findings showed that English proficiency was significantly related to English academic achievement, even for ELL students who had been in U.S. schools for 3 years or longer. Furthermore, the 5th grade ELL cohort had greater increases in reading and writing scores compared to all Colorado 5th graders. This led to a slight closing of the achievement gap. Lastly, Spanish achievement, especially when combined with English proficiency, predicted English achievement.

<http://brj.asu.edu>

07-453 MCCOY, DAMIEN (Australian Centre for Education and Training, Vietnam), **Utilizing students' preferred language learning strategies for IELTS test preparation.** *EA Journal* (English Australia) 23.1 (2006), 3-13.

This paper looks at how students' preferred language learning strategies can be used to maximize students' performance in the speaking component of the IELTS test. The research is based on work by O'Malley & Chamot (1990) and was undertaken in the context of an intensive IELTS test preparation course in Vietnam. The study found that using students' preferred learning strategies can be an effective way to help them prepare for high stakes international exams. Students who used the metacognitive strategies of self-evaluation and effective organization performed comparatively well in the test. In contrast, those students who used the cognitive strategy of delayed production generally did not achieve their desired results in the speaking test. A series of recommended activities which had the broadest appeal and most positive outcomes for test candidates are suggested.

<http://www.englishaustralia.com.au>

07-454 MENKEN, KATE (City U New York, USA), **Teaching to the test: How no child left behind impacts language policy, curriculum, and**

instruction for English language learners.

Bilingual Research Journal (National Association for Bilingual Education) 30.2 (2006), 521-547.

In the wake of the federal No Child Left Behind legislation, standardized tests have become increasingly high-stakes. Yet English language learners (ELLs) typically score far below native English speakers, creating pressure to 'teach to the test'. This article shares findings from an intensive year long study in 10 New York City high schools, detailing how high-stakes tests become de facto language policy in schools. Most schools and individual educators have increased the amount of English instruction ELLs receive; however, some have instead increased native language instruction as a test preparation strategy. Curriculum and instruction focuses on test content and strategies, and English as a second language classes have become more like English language arts classes for native English speakers. In bilingual classes, tests are found to promote monolingual instruction with test translations guiding decisions about language allocation.

<http://brj.asu.edu>

07-455 PAE, TAE-IL (Yeungnam U, China) & **GI-PYO PARK, Examining the relationship between differential item functioning and differential test functioning.** *Language Testing* (Hodder Arnold) 23.4 (2006), 475-496.
doi:10.1191/0265532206lt338oa

The present study utilized both the IRT-LR (item response theory likelihood ratio) and a series of CFA (confirmatory factor analysis) multi-sample analyses to systematically examine the relationships between DIF (differential item functioning) and DTF (differential test functioning) with a random sample of 15 000 Korean examinees. Specifically, DIF was detected using the IRT-LR method and the cumulative effect of DIF on DTF was gauged by the multisample analysis technique offered by the LISREL 8.5 program. The results of the current study indicate that item level DIF, once detected, may be carried to the test level bias regardless of the DIF directions, thereby showing mixed evidence to the previous findings reported in the literature. This suggests that the relation of DIF to DTF seems to be much more complex than that reported in the literature and, accordingly, more empirical studies are needed to bridge the gap in the literature about DIF-DTF relationships.

<http://arnoldpublishers.com/journals>

07-456 RIMMER, WAYNE (U Reading, UK), **Measuring grammatical complexity: The Gordian knot.** *Language Testing* (Hodder Arnold) 23.4 (2006), 497-519.
doi:10.1191/0265532206lt339oa

Grammar is central to language description and a posteriori construct validation of language tests

consistently identifies grammar as a significant factor in differentiating between score levels and characterizing overall proficiency. However, there is currently no model of grammatical competence robust enough to be operationalized in tests. Critical to describing grammatical competence is complexity of form and structure, yet grammatical complexity is poorly defined in linguistics and its sub-disciplines, including language assessment. This article argues that the inherent subjectiveness of complexity can be much reduced by a research methodology that combines input from corpora with expert intuition of the linguistic status of the grammatical items retrieved. The significance of the discussion is that, in addition to contributing to the on-going process of construct validity, a corpus-informed study of complexity has the potential to reconcile the tension in theoretical linguistics between grammar as being sentence-bound and grammar as discourse.

<http://arnoldpublishers.com/journals>

07-457 RUPP, ANDRÉ A. (Humboldt U, Berlin, Germany), **TRACY FERNE & HYERAN CHOI, How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective.** *Language Testing* (Hodder Arnold) 23.4 (2006), 441-474. doi:10.1191/0265532206lt337oa

This article provides renewed converging empirical evidence for the hypothesis that asking test-takers to respond to text passages with multiple-choice questions induces response processes that are strikingly different from those that respondents would draw on when reading in non-testing contexts. Moreover, the article shows that the construct of reading comprehension is assessment specific and is fundamentally determined through item design and text selection. The data come from qualitative analyses of 10 cognitive interviews conducted with non-native adult English readers who were given three passages with several multiple-choice questions from the CanTEST, a large-scale language test used for admission and placement purposes in Canada, in a partially counter-balanced design. The analyses show that: there exist multiple different representations of the construct of 'reading comprehension' that are revealed through the characteristics of the items; learners view responding to multiple-choice questions as a problem-solving task rather than a comprehension task; learners select a variety of unconditional and conditional response strategies to deliberately select choices; and learners combine a variety of mental resources interactively when determining an appropriate choice. These findings support the development of response process models that are specific to different item types, the design of further experimental studies of test method effects on response processes, and the development

of questionnaires that profile response processes and strategies specific to different item types.

<http://arnoldpublishers.com/journals>

07-458 VANDERVEEN, TERRY (Kangawa U, Japan), **The effect of EFL students' self-monitoring on class achievement test scores.** *JALT Journal* (Japan Association for Language Teaching) 28.2 (2006), 197-206.

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>

07-459 VAN MOERE, ALISTAIR (Lancaster U, UK), **Validity evidence in a university group oral test.** *Language Testing* (Hodder Arnold) 23.4 (2006), 411-440. doi:10.1191/0265532206lt336oa

This article investigates a group oral test as administered at a university in Japan to find if it is appropriate to use scores for higher stakes decision making. It is one component of an in-house English proficiency test used for placing students, evaluating their progress, and making informed decisions for the development of the English language curriculum. The implementation of a cut-score for students to advance through the university system has recently been proposed, bringing the group oral test component under increased scrutiny. On two successive occasions 113 participants sat the oral test in groups composed of different interlocutors each time. Rasch analysis shows rater fit within acceptable levels considering the length and nature of the test; however, at correlations of .74 inter-rater agreements are lower than has been reported in research on commercially available interview tests. Candidates' scores on the two different test occasions correlate at .61. A generalizability study shows that the greatest systematic variation in test scores is contributed by the person-by-occasion interaction. Topic, or prompt, was not a significant factor. Candidates' performances, or how raters perceive an individual candidates' ability, could be affected to a large degree by the characteristics

of interlocutors and interaction dynamics within the group.

<http://arnoldpublishers.com/journals>

Teacher education

doi:10.1017/S0261444807004417

07-460 DOGANÇAY-AKTUNA, SERAN (Southern Illinois U, USA), **Expanding the socio-cultural knowledge base of TESOL teacher education.**

Language, Culture and Curriculum (Multilingual Matters) 19.3 (2006), 278–295.
doi:10.2167/lcc320.0

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>

07-461 GARRISON, LESLIE (San Diego State U, USA), **Teacher attrition and retention along the Mexican border.** *Bilingual Research Journal* (National Association for Bilingual Education) 30.2 (2006), 365–384.

This study investigated teacher attrition and retention among 527 teachers who earned credentials between 1991 and 1998. They all attended a regional campus of a large university located in a rural area of southern California. Although the area had high poverty and many English language learners, the teacher attrition rate was markedly lower than nationally reported figures. Individual interviews of a randomly selected sample were conducted to determine differences between the conditions these teachers encountered during their first years of teaching and the conditions that contributed to dissatisfaction and attrition among teachers reported in Ingersoll and Smith (2003). No major differences

in teaching conditions were found. Alternative explanations for the high retention rate are explored.

<http://brj.asu.edu>

07-462 HARRISON, JENNIFER (U Leicester, UK), **An examination of the language and interpretations of 'Standard one' for initial teacher training in England: 'Professional values and practice' – outcomes or opportunities?**

European Journal of Teacher Education (Routledge/Taylor & Francis) 29.4 (2006), 431–454.
doi:10.1080/02619760600944753

This is an analysis of school tutors' assessments of aspects of the document *Standard one, professional values and practice* (DfES/TTA 2002) for secondary trainee teachers, together with interview data from a sample of trainee teachers and their school tutors/assessors. The paper examines teachers' interpretations of four sub-categories within this Standard and considers these in terms of three key questions: 1. How is the terminology of the standard being defined and described? 2. How is the evidence for assessment selected and evaluated? 3. What are the interpretative frameworks? While this standard is almost universally welcomed as an important component of teacher education in England, its assessment remains elusive for many teachers and tutors. As a result the standard is highly subject to personal interpretations of professionalism and what this might mean in terms of schools' provision for teacher development in this area. The paper supports the use of a framework of teachers' virtues and personal attributes, professional skills and learning and a set of professional behaviours, which might allow for institutional assessment procedures to take better account of both personal and professional development as well as professional competence. The place of critical reflection on practice is considered as a way of clarifying values and, the paper argues, it should become a central component of this standard.

<http://www.tandf.co.uk/journals>

07-463 LOVE, KRISTINA (U Melbourne, Australia) **& MERLE ISLES, 'Welcome to the online discussion group': A diagnostic framework for teachers.** *Australian Journal of Language and Literacy* (Australian Literacy Educators' Association) 29.3 (2006), 210–224.

State and national curriculum authorities are increasingly encouraging students and teachers to use asynchronous online discussions to communicate both with known and unknown participants. In this paper, we offer a contribution towards the development of a framework that may help teachers design and manage those online discussions concerned with expressing ideas and discussing alternative views about texts in subject English.

<http://www.alea.edu.au>