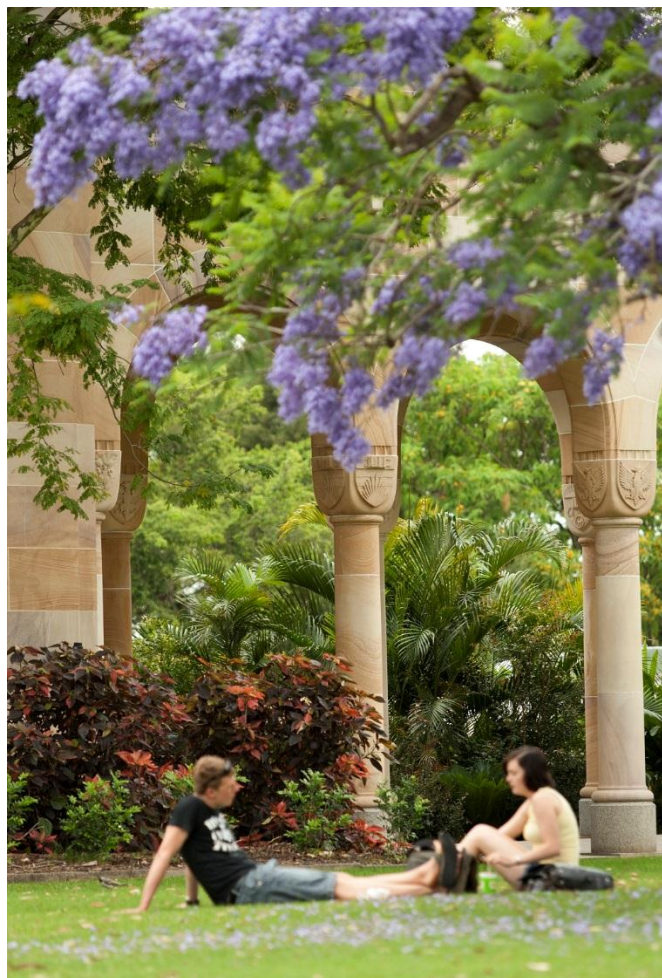




ALTAANZ Conference 2014

**Assessing second languages: Linking theory,
research, policy and practice**



**November 27 - 29, 2014
The University of Queensland
Brisbane, Australia**

TABLE OF CONTENTS

Map.....	3
Welcome.....	4
About ALTAANZ	5
ALTAANZ working parties.....	6
Student Travel Award recipients	7
Sponsors and Supporters	8
Important Information.....	9
Guidelines for presenters.....	9
Useful Information.....	10
Network Group Forum.....	11
Pre-conference workshops: Thursday 27 November.....	12
Pre-conference workshops: Abstracts.....	13
Conference day 1: Friday 28 November.....	15
Conference day 2: Saturday 29 November	18
Work in progress sessions	21
Posters	22
Plenaries.....	23
Symposium	27
Abstracts.....	29
Presenter contact details.....	55

MAP



UQ campus map app available at <http://uq.edu.au/uqnav/>

WELCOME

The ALTAANZ conference is organised every two years by the Association for Language Testing and Assessment of Australia and New Zealand. It is an important means of furthering our mission in training, research, and policy formation/advice. It is also a means of sharing our ideas and excitements, and of course, an occasion for community building among language testing/assessment professionals within and beyond the region. We are delighted that this year's conference has attracted over 100 colleagues not only from Australia, New Zealand, and the Asia Pacific region, but also from as far as Europe and North America.

The conference theme this year is "Assessing second languages: Linking theory, research, policy and practice". We are aware of the need to build bridges between language teachers and language assessment researchers and hope that the conference will go some way towards doing this by focusing on the role of practice in informing theory and as well as the importance of research in illuminating classroom practice and policy formation.

The program promises a range of high quality presentations on a diverse range of topics encompassing both standardized tests and classroom assessments. Our keynote speakers, Professor Micheline Chalhoub-Deville (University of North Carolina at Greensboro), Professor Chris Davison (University of NSW), and Associate Professor Angela Scarino (University of South Australia), bring a wealth of knowledge and experience of language testing and assessment in educational and other professional contexts.

We would like to express our appreciation to Cambridge ESOL, ETS, IDP IELTS, Pearson, English Australia and the Institute of Modern Languages (UQ) whose generous sponsorship for this conference is particularly important for a young organisation such as ours. We also acknowledge the in-kind support provided by the Institute of Continuing and TESOL Education, UQ and the Catholic Education Archdiocese of Brisbane.

Hearty thanks are due to members of the Conference Organizing Committee for the many months of planning and hard work that has gone into organising this event. We are also grateful to the volunteers and session chairs for their role in ensuring the smooth running of this conference.

We wish you a very fruitful professional experience and a pleasant stay in Brisbane! We also look forward to seeing you all in Adelaide next year, where ALTAANZ will be co-organizing a language assessment strand within the ALAA/ALANZ conference.

Catherine Elder
Peter Yongqi Gu
ALTAANZ Co-Presidents

ABOUT ALTAANZ

The purpose of the Association for Language Testing and Assessment of Australia and New Zealand (ALTAANZ) is to promote best practice in language assessment in educational and professional settings in these two countries and to foster collaboration between academia, schools and other agencies responsible for language testing or assessment. Its goals are listed under three broad headings below:

Training

Stimulate professional growth and best practice in language testing and assessment through workshops and conferences.

Research

Promote research in language testing and assessment through seminars, conferences and/or publications (ALTAANZ publishes a web-based journal and a newsletter).

Policy formation/advice

Provide advice on assessment to public and other relevant agencies on assessment-related issues, and advocate on behalf of test-takers, students and other stakeholders whose life chances may be affected by assessment-related decisions.

For further information about the organisation, please visit the website at: <http://www.altaanz.org>.

To become a member of ALTAANZ, please download a membership form from the website and email it to altaanz@gmail.com.

ALTAANZ WORKING PARTIES

Conference Organising Committee

Bernadette Barker, Brisbane Catholic Education
Obaid Hamid, The University of Queensland
Michael Harrington, The University of Queensland
Catherine Hudson, Australian National University, Department of Education, Training & Employment
Noriko Iwashita, The University of Queensland
Iain Mathieson, Institute of Continuing and TESOL Education, The University of Queensland
Lyn May, Queensland University of Technology
Catherine Moore, ETS Global
Paul Moore, The University of Queensland
Janet von Randow, DELNA, The University of Auckland

Best Student Presentation Award Committee

Martin East (Chair), The University of Auckland
Micheline Chalhoub-Deville, University of North Carolina, Greensboro
Susy Macqueen, The University of Melbourne

Abstract Reviewers

Karin Ashton, Massey University
Ana Maria Ducasse, RMIT University
Martin East, The University of Auckland
Cathie Elder, The University of Melbourne
Rosemary Erlam, The University of Auckland
Kellie Frost, The University of Melbourne
Peter Gu, Victoria University of Wellington
Michael Harrington, The University of Queensland
Kathryn Hill, The University of Melbourne
Peter Keegan, The University of Auckland
Ute Knoch, The University of Melbourne
Susy Macqueen, The University of Melbourne
Tim McNamara, The University of Melbourne
Sally O'Hagan, The University of Melbourne
Aek Phakiti, The University of Sydney
John Pill, The University of Melbourne
Katherine Quigley, Victoria University of Wellington
John Read, The University of Auckland
Carsten Roever, The University of Melbourne

Student Volunteers

Felipe Juan Pablo Araos Munoz, The University of Queensland
Kate D'Ozario, Queensland University of Technology
Ngoc T. H. Hoang, The University of Queensland
Silvana Flores, The University of Queensland
Fei Geng, Queensland University of Technology

Susan Shanshan Lu, The University of Queensland
Tanya Kelly, The University of Queensland
Hang Mai, The University of Queensland
Khoi Ngoc Mai, The University of Queensland
Kevin Insik Oh, The University of Queensland
Constanza Pena, The University of Queensland
Zeliha Arshad Sethi, The University of Queensland
Duyen T. Tran, Queensland University of Technology
Claudia Vasquez, The University of Queensland
Alimin Waloyo, The University of Queensland
Xiaodong (Denny) Yao, The University of Queensland

STUDENT TRAVEL AWARD RECIPIENTS

Chao Han, Macquarie University
Naoki Ikeda, The University of Melbourne
Nguyen Thi Huynh Loc, Victoria University of Wellington, NZ

SPONSORS AND SUPPORTERS



Cambridge English

PEARSON



School of Languages and
Comparative Cultural Studies

School of Education

Institute of Continuing and
TESOL Education

Institute of Modern Languages



Australian
National
University



IMPORTANT INFORMATION

- **Registration**

- Thursday 27 November 12pm-1pm, 3pm-4pm & 5:30pm-6:30pm in the foyer in front of Room 302 and 303 in Social Sciences Building (No. 24)
- Friday 28 November and Saturday 29 November: 8am-9am in Level 1 foyer in Sir Llew Edwards Building (No. 14)

- **Workshops** – 1pm-3:30pm and 4pm-6:30pm at Room 303/302, Social Sciences Building (No. 24) on Thursday 27 November

- **Conference** – Friday and Saturday 28-29 November on Level 1 and 2 in Sir Llew Edwards Building (No. 14)

- **Opening Reception** – 6:45pm on Thursday 27 November at UQ Art Museum (James and Mary Emelia Mayne Centre, Building 11)

- **Conference Dinner** – 6:30pm at St Leo's College, College Road St Lucia ph (07) 3878 0600 (J13 on the campus map)

- **Newcomers' session** – At the reception and morning tea on Friday 28 November

GUIDELINES FOR PRESENTERS

1) **Paper presentation** – The paper presentations cover a range of topics and focus on both research and assessment from researchers and practitioners from diversified sectors. Each session is 20 minutes followed by five minutes discussion. Please save your slides on USB and upload to the computer well before your talk.

2) **Work in Progress (WIP)** – This session gives researchers and teachers the opportunity to share and discuss aspects of research projects they are currently involved in or classroom assessment tasks that they have developed. In doing so, they will receive feedback from conference attendees. The one hour session will be structured in three 20 minute sections, so that each presenter has the opportunity to discuss their project/task with 3 groups of interested conference attendees. Each room will have several presenters, and each presenter will be stationed at a table in a different area in the room.

3) **Poster presentation** – All poster presentations will be held at lunch time (12:05-1pm) on Saturday, 29 November. You will be expected to prepare for a poster before the conference and at the conference you will be given a space to place your poster. During lunch time, presenters are expected to stand next to the poster to explain their study and answer any questions.

USEFUL INFORMATION

Facilities on campus (All information is available on the conference website)

- **Banks** – Commonwealth bank and ANZ Bank and ATMs (9-4pm weekdays) (Building No. 21A and B)
- **Post Office** – 9-5pm weekdays (Building No. 61)
- **UQ Sport** (Open to public) (Building No. 25)
- **Car park** (Open to public)
- **Health Service** (Open to public) (Building No. 32)
- **Campus Pharmacy** ph 07 3870 1509, 3870 3199 (Internal: 51732) (Building No. 21B)
- **Internet access for visitors** – You will have access to 1GB of download quota for use on the wired network to free of charge (see: <http://uqconnect.net/visitor>). If you wish to use more than 1GB, you can [purchase a one-off block of additional download](#). More information will be available at the registration desk.
- **Cafes on campus** – Merlo and Saint Lucy are the closest to the conference venue; All may be open on Saturdays
 - *Merlo* in the Duhig Building (No. 2)
 - *Wordsmiths* – The Writer's Café (next to the Co-op Bookshop)
 - *Saint Lucy Café Et Cucina* (in UQ Tennis Centre)
 - *Darwins* (in Building No. 94)

Restaurants near the campus

There are a number of restaurants with a variety of cuisines near the campus (approx. 20 mins walk from the conference venue) as well as on campus. More information is available below:

St Lucia <http://www.urbanspoon.com/n/337/48125/Brisbane/St-Lucia-restaurants>

Campus <http://www.uq.edu.au/about/places-to-eat>

Please note that lunch and morning/afternoon tea will be served at the conference venue during the conference.

Supermarket

IGA in St Lucia (6am-midnight everyday) (20 mins walk from the conference venue)
242 Hawken Dr, St Lucia QLD 4067 ph 07 3371 6466

Medical Emergency - Brisbane public hospitals

- Royal Brisbane Hospital, Butterfield Street, Herston, ph 07 3646 8111
- Princess Alexandra Hospital, Ipswich Road, Woolloongabba, ph 07 3176 2111
- Mater Hospital, Raymond Terrace South Brisbane ph 07 3163 8111

Do not go to the emergency centre of a Queensland Public Hospital if you have a **minor illness**. This should only be used for an emergency. You can call 1300HEALTH, a government sponsored medical helpline available 24 hours per day.

For medical emergencies **off campus** dial 000 for an ambulance or visit the emergency room of a nearby public hospital (which provide free emergency care to Australian residents). Expect long delays.

NETWORK GROUP FORUM

Meeting of Minds: Educators reveal issues in their diverse classroom contexts for language testing and assessment researchers

In this forum, chaired by Associate Professor Cathie Elder from the University of Melbourne, teachers will explain the language assessment and teaching issues that they encounter in their particular classroom contexts. These will be collected prior to the conference from the (classroom) teachers who are registering to attend, enabling them to voice their interest in language assessment. This forum will give researchers at the conference significant insights into language assessment across diverse contexts in schools.

The aim of the forum will be to implement the ALTAANZ vision of fostering research collaboration and best assessment practice in schools. The role of the researchers in the forum will be to listen to and understand the issues presented by the teachers. Where relevant, researchers with specific expertise may be asked by the chair to respond briefly to issues raised by the teachers. Researchers will be able to augment their understandings of schooling contexts as a first step in building useful collaborations to foster research and development activities which serve the interests of the teaching profession. The forum will address diverse language assessment issues relating to different age stages, programs and language situations, drawing links to policy, school curriculum, and assessment and testing requirements at state and federal levels.

After the forum, participants will be sent a summary of the issues discussed. Possibilities for research collaborations will be further explored via ongoing interactions on the ALTAANZ Facebook page.

Cathie Elder has worked as a language teacher both in Australia and overseas and also as teacher educator. Her PhD thesis completed in 1997 examined the performance of learners from different backgrounds on school language examinations. Since then she has published widely on language testing and has conducted a range of language assessment projects relating to language learning in schools and universities in Australia, New Zealand and elsewhere.

PRE-CONFERENCE WORKSHOPS: THURSDAY 27 NOVEMBER

Registration	12:00-1:00 3:00-4:00	Foyer in front of Room 302 and 303, Social Sciences building (no. 24)	
		Room 303 (Social Sciences building (no. 24))	Room 302 (Social Sciences building (no. 24))
Workshops	1:00-3:30	Workshop 1 Large-scale, standardized and classroom assessments <i>Professor Micheline Chalhoub-Deville</i>	
Afternoon tea	3:30-4:00	Foyer in front of Room 302 and 303, Social Sciences building (no. 24)	
Workshops	4:00-6:30		Workshop 2 The role of interpretation in assessment for learning in languages education <i>Associate Professor Angelo Scarino</i>
Workshops	4:00-6:30	Workshop 3 Improving EAL learning outcomes through enhancing teacher assessment literacy <i>Professor Chris Davison</i>	
Registration	5:30-6:30	Foyer in front of Room 302 and 303 in Social Sciences building (no. 24)	
Opening reception	6:45-	UQ Art Museum (James and Mary Emelia Mayne Centre, Building 11)	

PRE-CONFERENCE WORKSHOPS: ABSTRACTS

Workshop 1

Professor Micheline Chalhoub-Deville

University of North Carolina, Greensboro

Large-scale, standardized and classroom assessments

The workshop is intended for anyone who is interested in expanding their conceptualization of how to document the quality of their assessments and related results/scores. The workshop will focus on validation that attends to standardized testing programs as well as classroom-based assessments. Workshop participants will broaden their understanding of validity conceptualization and develop an understanding of approaches that guide validation efforts. They will also become familiar with a methods typically used to explore and document different aspects of validity. Participants will have the opportunity to explore research (data collection and analyses) in this area, if interested. Participants are encouraged to bring samples of validation plans, instruments, and/or studies conducted to share. The interactive workshop is structured to include presentations, discussions, and opportunities to work in small groups. Participants in the workshop will be able to:

- enrich their understanding of the role of validation as part of quality assurance of assessments;
- develop an appreciation of the state of knowledge with regard to validity and validation;
- become familiar with a variety of methods, which can be used to document different aspects of the validity process;
- receive input on how to improve their assessment documentation;
- identify reference sources to consult in this area;
- outline validation/research plans they can pursue on their own; and
- share materials and insights.

Workshop 2

Associate Professor Angela Scarino

University of South Australia

The role of interpretation in assessment for learning in languages education

In this interactive workshop, designed primarily for teachers of language K-12, we examine the role of interpretation in conceptualising, eliciting and judging student performance and progress in learning an additional language. Drawing on examples from a range of research studies conducted with teachers of languages, working in diverse languages and contexts, participants will be invited to critically examine a range of tools (including models, rubrics, texts student products) developed by and for teachers. They will discuss the extent to which these tools provide the possibility of understanding not only the students' products but also their interpretations of meanings in learning within episodes and across learning trajectories. The workshop will conclude with a discussion of the implications of working with students' interpretations in assessing language learning.

Workshop 3
Professor Chris Davison
The University of New South Wales

Improving EAL learning outcomes through enhancing teacher assessment literacy

The School of Education at the University of New South Wales (UNSW), in partnership with Educational Assessment Australia, has been commissioned by the Victorian Department of Education and Early Childhood Development (DEECD), the Catholic Education Commission, Victoria (CECV) and Independent Schools Victoria (ISV), to research and develop an innovative e-based toolkit of assessment tools and advice to assist teachers in developing pedagogically sound and useful approaches to assessing the English language development of EAL students across all years in Victorian schools. Research has demonstrated that to improve learning and teaching, teachers of such students need to have a deep understanding of assessment for learning principles as well as access to a bank of appropriate assessment tools and tasks, exemplars and work samples, and networks of supportive and engaged peers so that they can make trustworthy assessment decisions. This interactive workshop will explore some of the issues and challenges in building such online assessment for learning communities in the EAL field.

The workshop will first highlight the theoretical background to the project and the various components of the project designed to develop and support EAL teachers' assessment literacy, at the same time providing selected findings on various aspects of the research and development process. The workshop will then focus on the tools being developed to assess students' language skills. Samples of EAL learners' work in response to some of the assessment tasks will be presented, and participants will be involved in rating the performance of learners, using the draft criteria and documentation being developed in the project. Participants will compare their ratings of the students' performances, and reflect on the usefulness of the task and accompanying documentation. The implications for classroom teaching and assessment will be explored. The workshop will also explore how these assessment principles and the tools developed in the project can be adapted for use by all ESL/EAL teachers regardless of the sector in which they work.

CONFERENCE DAY 1: FRIDAY 28 NOVEMBER

Registration	8:00-9:00	Level 1 Foyer, Sir Llew Edwards Building (no. 14)			
Welcome	9:15-9:30	Auditorium (Room 212, Sir Llew Edwards Building [no. 14])			
Plenary address	9:30-10:30	The new frontiers of validity and L2 constructs <i>Professor Micheline Chalhoub-Deville, University of North Carolina, Greensboro</i>			
		Room 115	Room 116	Room 217	Room 216
Parallel Session A	10:35-11:05	Language assessment policy and practice: ethics and accountability Validating test-takers' voices <i>Liying Cheng</i>	Research methodology for exploring language assessment Assessment of students with complex language contact backgrounds in Australia and the United States: the need for a contextualised research methodology <i>Denise Angelo, Catherine Hudson & Rosalie Grant</i>	Standardised tests: development, implementation and/or use by institutional stakeholders Involving test-takers in validating uses of high-stakes language assessments: Opportunities and challenges <i>Ngoc T. H. Hoang</i>	Comparing the predictive power of regression, CART, and Artificial Neural Networks: An English reading comprehension study <i>Vahid Aryadoust</i>
Morning Tea	11:05-11:30	Level 1 Foyer, Sir Llew Edwards Building (no. 14)			
Parallel Session B	11:30-12:00	Pleading guilty without understanding? The challenges of assessing English comprehension in a forensic context <i>Catherine Elder</i>	The interface between test validation and research methodology: The pledge of mixed-methods research (MMR) <i>Mehdi Riazi</i>	High-stakes test preparation across different contexts: Australia, China and Iran <i>Shahrzad Saif, Jia Ma, Lyn May & Liying Cheng</i>	The reliability of automated rating machines in predicting the quality of reflective blog posts based on limited linguistic features <i>Brad Blackstone, Wu Siew Mei & Vahid Aryadoust</i>

		Room 115	Room 116	Room 217	Room 216
Parallel Session C	12:05-12:35	Linking theory to practice in the assessment of second language spoken proficiency: Teachers' and students' perspectives on different assessment types <i>Martin East</i>	Measuring rater variability in interpreter performance testing: Using classical test theory, generalizability theory and Rasch measurement <i>Chao Han</i>	NAPLAN and the problem LBOTE category: Patterns in LBOTE student test answers and their English language proficiency level <i>Sue Creagh</i>	Assessing language for academic purposes What makes a good performance? – analysis of test-taker discourse of the IELTS Academic Speaking Task <i>Noriko Iwashita & Claudia Vasquez</i>
Parallel Session D	12:40-1:05	The place of English in the University Entrance Examination in China: An update <i>Peter Yongqi Gu</i>	Unleashing the power of the Halliday-an third eye; 'tone' in letter writing tasks made visible <i>Johanna Motteram</i>	Second language proficiency models and implications for developing assessment tasks Tripartite conceptualizations of English language proficiency in Australian higher education and the role of post-entry language assessment <i>Paul Moore & Michael Harrington</i>	"An ability-in writer-in context": An interactive model of L2 academic writing performance <i>Mehdi Riazi</i>
Lunch and ALTAANZ AGM	1:05-2:15	Lunch - Level 1 Foyer, Sir Llew Edwards Building (no. 14) AGM – Room 116, Sir Llew Edwards Building (no. 14)			
Works in Progress	2:15-3:15	1 and 2	3, 4 and 5	6, 7 and 8	9 and 10

		Room 115	Room 116	Room 217	Room 216
Parallel session E	3:15-3:45	Living with IELTS: A narrative inquiry into the lived experiences of IELTS test candidates. <i>Megan Yucel</i>		Standardised tests: development, implementation and/or use by institutional stakeholders Efficacy: Student perceptions of English language tests as indicators of academic preparedness <i>John Gardiner & Stephen Howlett</i>	Assessing language for academic purposes Coherence and cohesion in academic writing: An observational study of discussions between raters <i>Janet von Randow & Helen Basturkmen</i>
Afternoon tea	3:45-4:10	Level 1 Foyer, Sir Llew Edwards Building (no. 14)			
Parallel Session F	4:10-4:40	Listening to test-taker voices <i>M. Obaidul Hamid</i>	Effects of corrective feedback on L2 pragmatic development <i>Nguyen Thi Thuy Minh, Do Thanh Ha, Nguyen Tuan Anh & Pham Thi Thanh Thuy</i>		Integrating pragmatic and interactional features of language use into academic language assessment: What discursive features are measurable? <i>Naoki Ikeda</i>
Parallel Session G	4:45-5:15	Washback effect on Chilean teachers of English: The case of SIMCE test <i>Silvana Flores</i>	How do pre-service teachers do language testing and assessment in their practicum? <i>Nguyen Thi Huynh Loc</i>		Predicting integrated writing performance from lexico-grammatical knowledge and reading proficiency using Structural Equation Modeling <i>Vahid Aryadoust & Susan Tan</i>
Group photos	5:20	TBA			
Conference Dinner	6:30-	St Leo's College, College Road, St Lucia (J13 on the campus map)			

CONFERENCE DAY 2: SATURDAY 29 NOVEMBER

Registration	8:00-9:00	Level 1 Foyer Sir Llew Edwards Building (no. 14)			
Plenary address	9:00-10:00	Auditorium (Room 212, Sir Llew Edwards Building [no. 14]) Expanding language assessment practices: The possibilities and challenges of assessing the interpretation and exchange of meaning in language learning in the context of linguistic and cultural diversity <i>Angela Scarino, University of South Australia</i>			
		Room 115	Room 132	Room 217	Room 216
Parallel Session H	10:05-10:35	Classroom-based assessment What does classroom interaction look like in high-stakes exam classes? <i>Dina Tsagari</i>	Assessing language for professional purposes In-service teachers' responses to a language assessment and training program in Vietnam <i>Khoi Ngoc Mai</i>		Symposium Teacher Responses to Assessment at ICTE-UQ Introduction to ICTE-UQ context Iain Mathieson 'Stand and deliver': reflections on the content and assessment of a bespoke English language academic programme <i>Maree Sandeford</i>
Morning Tea	10:35-11:00	Level 1 Foyer, Sir Llew Edwards Building (no. 14)			
Parallel Session I	11:00-11:30	Assessment for learning in the EFL classroom: A case study of practices in a Vietnamese university <i>Duyen T. Tran</i>	Building a validation argument for the Occupation English Test (OET): Investigating test taker processes on the listening subtest <i>Kellie Frost & Ute Knoch</i>		Symposium (cont.) Auditory Thinking: a scaffold for helping learners answer knowledge-based test questions on connected speech <i>Arizio Sweeting & Martin Dutton</i> It's a musical language: The effect of intervention on pronunciation assessment outcomes <i>Vicki Bos & Megan Yucel</i>

		Room 115	Room 132	Room 217	Room 216
Parallel Session J	11:35-12:05	Assessing the impact of teacher feedback on accuracy in the writing of EFL learners: A longitudinal study <i>Ali Rastgou</i>	Transitioning from university to the workplace: Stakeholder perceptions of academic and professional writing demands <i>Ute Knoch, Lyn May, Susy Macqueen, John Pill & Neomy Storch</i>	Validation of a web-based test of ESL sociopragmatics <i>Carsten Roever, Catriona Fraser & Catherine Elder</i>	Symposium (cont.) Using a web-based interactive tool for formative assessment via mobile devices <i>Paul Forster</i> Q & A
Lunch & Posters**	12:05-1:00	Level 1 Foyer, Sir Llew Edwards Building (no. 14)			
Parallel Session K	1:05-1:35	Assessment literacy of foreign language teachers: Findings of a European study <i>Dina Tsagari & Karin Vogt</i>	Stakeholder perceptions of test impact: The case of IELTS and overseas trained teachers <i>Judie Cross, Jill Murray & Ken Cruickshank</i>	Network Group Forum Meeting of Minds: Educators reveal issues in their diverse classroom contexts for language testing and assessment researchers <i>Cathie Elder, Denise Angelo, Cath Hudson (coordinators)</i>	
Parallel Session L	1:40-2:10		Language testing in medical registration procedures: Perceptions of the role of language and communication skills <i>John Pill & Sally O'Hagan</i>		
Parallel Session M	2:15-2.45	Identifying and assessing the full cohort of year two English language learners in a regional town of northern Queensland <i>Courtney Farley & Denise Angelo</i>	Writing demands of healthcare professionals <i>Susy Macqueen, Sharon Yahalom, Hyejeong Kim & Ute Knoch</i>	Assessing language for academic purposes What can you see in those clouds? : An image based PELA <i>Craig Baird & Patricia Dooley</i>	

Afternoon tea	2:45-3:15	Level 1 Foyer, Sir Llew Edwards Building (no. 14)
Closing plenary	3:15-4:15	Auditorium (Room 212, Sir Llew Edwards Building [no. 14]) Teacher-based assessment of English as a second or additional language: Challenges in theorizing our practice <i>Chris Davison, The University of New South Wales</i>
Conference closing	4:15-4:35	Includes announcement of Best Student Presentation award
Group photos	4:40	TBA

WORK IN PROGRESS SESSIONS

Presentation rooms:

Tables 1-2: Room 115

Tables 3-5: Room 116

Tables 6-8: Room 217

Tables 9-10: Room 216

Table	Title	Presenters
1.	Indigenous assessment: criteria discussed by raters marking undergraduate classroom based speaking tasks in diverse academic contexts	<i>Annie Brown & Ana Maria Ducasse</i>
2.	'What is it you want to mean?' Assessing strategic competence in English as an international language	<i>Douglas Meyer</i>
3.	On the psychometric quality and criterion validity of self- and peer-assessments of the oral presentations of first-year science students	<i>Vahid Aryadoust</i>
4.	A validation framework for an online English language Exit Test: A case study using Moodle as an assessment management system	<i>Zakiya Al Nadabi</i>
5.	Response validity of the English Language Examination for university entry in Shanghai.	<i>Matthew Book</i>
6.	The impact of teachers' beliefs on their formative assessment practices	<i>Rashid AlMaamari, Gillian Wigglesworth & Ute Knoch</i>
7.	Improvements of assessment practices at Defence Force School of Languages	<i>Ana Saldanha & Beth Foster</i>
8.	Replacing new for old: How to implement a change to a writing-from-sources test in an EAP course	<i>Sascha Mitchell</i>
9.	Development and implementation of a method for validation of EAP Reading and Listening Assessment	<i>Kerensa Townsend</i>
10.	Helping Vietnamese EFL learners to develop learner autonomy through self-assessment	<i>Thao Pan</i>

POSTERS

Title	Presenters
The school language data profile	<i>Denise Angelo, Henry Fraser & Cassy Nancarrow</i>
Assessment practices of secondary schools in Chile: A report on three different contexts	<i>Felipe Araos & Constanza Pena</i>
Development of mainstream classroom teacher tool to self-assess ability and confidence when implementing differentiated units for EAL/D learners	<i>Courtney Farley, Rose Callaghan, Bernie Bliss & Evelyn Griffiths</i>
Defining effective communication in the psychiatric interview	<i>Kathryn Hill, Jonathon Crichton, Joel King & Andrew Gleason</i>
Identifying the language demands of assessment tasks for classroom teaching purposes	<i>Cath Hudson, Ellie Brownlow & Erika Charola</i>
Introducing the Cambridge English Scale	<i>Carmen Le</i>
Alternative approaches to Chinese language assessment in a middle-school classroom	<i>Jennifer Russell</i>

PLENARIES

Plenary 1: *The new frontiers of validity and L2 constructs*

Professor Micheline Chalhoub-Deville, University of North Carolina, Greensboro

Day: Friday 28 November

Time: 9:30 – 10:30

Room: 202 in Sir Llew Edwards Building (no. 14)

Contact: mbchalho@uncg.edu



Abstract

Theoretical and operational developments in measurement and language testing demand a reassessment of our validation and L2 construct representation. In terms of validation Messick's (1989) emphasis on construct judgments is giving way to Kane's (2006) "interpretation/use argument" (IUA) justification. In language testing, this shift is exemplified in the works of Bachman and Palmer (2010), Chapelle, Enright, and Jamieson (2008) Chapelle (2012). The talk will discuss tenets presented in construct-based and justification-anchored validity formulations and reflect on the implications for L2 testing. A discussion of validity, especially given the nature of the language testing profession, necessitates that we also pay close attention to L2 construct representation. The talk will address the theoretical underpinnings of the L2 construct and invite the profession to consider an expanded view of our theoretical orientation. The talk will take up socially grounded frameworks, which advocate for localized theories of the L2 construct (Chalhoub-Deville, 2003), integrate consequences into technical quality documentation, and call for more realistic delineations of stakeholders and their roles.

Bio

Micheline Chalhoub-Deville (Ph.D., The Ohio State University) is Professor of Educational Research Methodology and Co-Director of the Coalition for Diverse Language Communities at the University of North Carolina at Greensboro (UNCG). Professor Chalhoub-Deville has been awarded the 2000 Outstanding Young Scholar Award by the Educational Testing Service—TOEFL Program and the UNCG 2013 School of Education Outstanding Senior Scholar Award, both of which recognize her contributions to the field of language testing and assessment. Professor Chalhoub-Deville has also received the 1998 Best Article Award from the International Language Testing Association.

Professor Chalhoub-Deville is the editor/co-editor of two books: *Issues in Computer Adaptive Testing of Reading Proficiency* (Cambridge University Press) and *Inference and Generalizability in Applied Linguistics: Multiple Research Perspectives* (John Benjamins Publishing Company). Her scholarship also appears in a number of peer-reviewed journals, including *Language Testing*, *Annual Review of Applied Linguistics*, *Foreign Language Annals*, *Language Learning and Technology*. In 2006, Professor Chalhoub-Deville co-authored a chapter for *Educational Measurement*, a compendium of critical articles from across the measurement field that is published every fifteen years. In 2011, she was the co-editor of the *Language Testing*, 28(3) Special Issue—Standards-based Assessment in the United States. Professor Chalhoub-Deville has been the lead developer of foreign language proficiency-based and ELL accountability assessment batteries.

Professor Chalhoub-Deville has been a keynote/plenary speaker at a number of conferences and meetings, including meetings in Australia, Korea, Lebanon, UAE, UK, and the U.S.A. In 2013, she was invited by the American University of Cairo, Egypt as the Distinguished Visiting Professor to give a series of talks, workshops and consultations with various departments on campus. In April, 2015, she is scheduled to give a keynote address at the annual meeting of the Language Training and Testing Center (LTTTC) International Conference, held in Taipei. The theme of the conference will be

“Critical Reflections on Foreign Language Education: Globalization and Local Interventions.” Throughout her professional career, Professor Chalhoub-Deville has received approximately \$3M to support her scholarship and R&D projects.

Professor Chalhoub-Deville has served in various leadership roles in the profession. She is a member of the editorial boards of both *Language Testing* and *Language Assessment Quarterly*. She is the founder of the Mid-West Association of Language Testers (MwALT). She is a former member of both the TOEFL Policy Board, as well as the TOEFL Committee of Examiners, where she also served as chair for three years. She currently serves on the ETS K-12 English Learner Technical Advisory Committee and on the British Council Assessment Advisory Board.

Plenary 2: *Expanding language assessment practices: The possibilities and challenges of assessing the interpretation and exchange of meaning in language learning in the context of linguistic and cultural diversity*

Angela Scarino, University of South Australia

Day: Saturday 29 November

Time: 9:00 – 10:00

Room: 202 in Sir Llew Edwards Building (no. 14)

Contact: Angela.Scarino@unisa.edu.au



Abstract

Globalisation is changing the very nature of multilingualism and of language, culture and language learning. These changes call for (1) an expansion of the goals of language learning to include both communication and an understanding of the act of meaning making and exchange in communication and (2) a more sophisticated understanding of the processes of language learning as interlinguistic and intercultural. These changes have important implications for assessment. In this paper I discuss the nature of the theoretical expansion and the possibilities and challenges that it poses for language assessment. Drawing on case studies from research with teachers of diverse languages K-12, I will discuss the need for assessment practices that recognise and support an understanding of communication and language learning as interpretive.

Bio

Angela Scarino is Associate Professor in Applied Linguistics and Director of the Research Centre for Languages and Cultures, University of South Australia. Her research expertise is in languages education in linguistically and culturally diverse societies, second language curriculum design and assessment, intercultural language learning and second language teacher education. She has been a Chief Investigator on a number of research grants, for example, Assessing the intercultural and language learning (ARC Linkage 2006-2009) and Student Achievement in Asian Languages Education (DEEWR, 2009-2011). Her most recent books include: *Intercultural Language Teaching and Learning* (with AJ. Liddicoat, Wiley-Blackwell, 2013), *Languages in Australian Education: Problems, Prospects and Future Directions*, co-edited with AJ Liddicoat, (Cambridge Scholars) and *Dynamic Ecologies. A Relational Perspective on Languages Education in the Asia Pacific Region* co-edited with N. Murray (Springer 2014). She is currently the Chair of the Multicultural Education Committee, an advisory committee to the Minister for Education in South Australia.

Plenary 3: *Teacher-based assessment of English as a second or additional language: Challenges in theorizing our practice*

Chris Davison, The University of New South Wales

Day: Saturday 29 November

Time: 3:15 – 4:15

Room: 202 in Sir Llew Edwards Building (no. 14)

Contact: c.davison@unsw.edu.au



Abstract

The development and use of teacher professional judgment has been the centrepiece of much reform in the assessment of English as a second or additional language in recent years, and has long been the mainstay of school-based assessment systems. However, teacher judgment has always been seen as one of the weakest links in the assessment process with concerns raised about the trustworthiness and reliability of teacher assessment decision-making processes and the teacher's ability to be both "accurate" and "fair". Some of these concerns are misdirected, the result of the widespread tendency to evaluate teacher-based assessment systems with traditional psychometric criteria developed for formal testing programs, but some of the concerns highlight the need for us to better theorize teacher-based assessment practice. This presentation will first explore some of the concerns and challenges in using teacher professional judgment in large scale school based assessment – both theoretical and practical issues - and then demonstrate how the development of teacher judgment can be supported and enhanced by online professional interaction, including the sharing of student work samples, and benchmarking of assessments. The implications for individual schools as assessment communities will also be discussed.

Bio

Professor Chris Davison is Head, School of Education, UNSW, an Honorary Professor, University of Hong Kong and a past president of ACTA. She has worked in teacher education, including at Melbourne and La Trobe University, for over 30 years, and before that, as an ESL teacher and consultant in AMES, TAFE, secondary schools and English language centres. Chris has published extensively on ESL development, language and content curriculum, and English language assessment. She is a series editor for Springer's new series on English language education and also editor and contributor to a two volume handbook on teaching English internationally (with Jim Cummins). She has undertaken large-scale curriculum and assessment projects in Hong Kong, Singapore and in Brunei, and is currently leading the research and development of a teacher-based assessment framework for EAL learners in Victorian schools, funded by the Department of Education and Early Childhood (DEECD), the Catholic Education office and Independent Schools Victoria.

SYMPOSIUM

Teacher Responses to Assessment at ICTE-UQ

In this Symposium, ESOL teachers from the Institute of Continuing and TESOL Education at the University of Queensland (ICTE-UQ) will share classroom based experiences and action research projects that explore teacher responses to the assessment requirements on a variety of different English language programs at ICTE-UQ. ICTE-UQ offers a wide range of English programs including General English, Academic English, University Pathway Programs, TESOL programs and customised programs. This means that a wide variety of assessments are used. The washback of these different assessment tasks requires teachers to respond in a variety of different ways, which in turn impacts on teachers' classroom practice and learning intervention strategies. This symposium will consist of four presentations. In each the presenter(s) will explore how they responded to particular features of the course assessment, and reflect on the success of the approaches they used.

Maree Sandeford

Stand and deliver: reflections on the content and assessment of a bespoke English language academic programme

This paper outlines informal action research and teacher reflection regarding one English language teaching programme, and has two parts. The first describes the 'journey' made by a team of three teachers during the development, delivery and assessment of an English for Academic and Research Communication (EARC) programme, specifically designed for a group of Science Without Borders (SWB) Brazilian undergraduate students, and delivered at the University of Queensland's Institute of Continuing and TESOL Education (ICTE-UQ), over a 5-week period in January and February, 2014. In particular, the assessment of speaking and writing will be explained. The second part discusses one teacher's reflections regarding the teaching and assessment challenges for this programme's speaking and writing components. Issues around student motivation, needs and test burnout, and assessment applicability to future tertiary academic study.

Arizio Sweeting & Martin Dutton

Auditory Thinking: a scaffold for helping learners answer knowledge-based test questions on connected speech

The dynamics of speech can be a challenge for English language learners and teachers alike. When sounds change shape in natural speech, they break the rules of what Cauldwell (2001: 18) calls the "careful speech model" or the "emulation model", rendering the written sound representation an unreliable referential. This phenomenon can be a particular issue for learners taking exams such as the Knowledge About English (KAL) test, a Cambridge English language knowledge-based test, in which learners are assessed on their ability to answer questions about connected processes without any aural or oral support. As an attempt to specifically improve learners' ability to answer the questions on connected speech (e.g. elision, intrusion, assimilation, weak forms, or consonant-vowel linking), we conducted a five-week action research project in which we developed and tried a range of activities involving different interaction patterns, macro-skills work and learner styles.

At the end of our project, we concluded that speaking and listening practice were a useful way of developing the learners' ability to hear the sounds between connected words in their minds and simultaneously strengthen their knowledge of connected speech for the purposes of a KAL test. We conceptualized this strategy as "auditory thinking" (Dutton & Sweeting, 2014: p. 8). In this session, we will draw on the pronunciation literature and our own classroom experience to explore the benefits of "auditory thinking" as a scaffold for helping learners to develop greater awareness of sound representations, or its misrepresentations thereof, through the use of receptive and productive activities during a KAL test preparation course.

References:

- Cauldwell, R. (2001). *Streaming speech: Listening and pronunciation for advanced learners of English* [CD-ROM/web-based materials]. Birmingham, England: Speech In Action.
- Dutton, M. & Sweeting, A. (2014). *Preparing students for answering connected speech questions in the Cambridge English Teaching Knowledge Test*. ISSN 1756-509X, 7.

Vicki Bos & Megan Yucel

It's a musical language: The effect of intervention on pronunciation outcomes

In this presentation, we will report on an action research project conducted in an English language centre attached to an Australian university. Our project focused on students in an academic pathway program at the Institute of Continuing and TESOL Education, the University of Queensland (ICTE-UQ). Through diagnostic testing, we identified thirty students who displayed poor performance in speaking as a result of their pronunciation. Our aim was to assist these students through a direct intervention program, consisting of remedial pronunciation workshops and choral singing.

We will begin by giving a brief overview of our research project. Key elements of the project will be highlighted, such as the use of formative assessment in determining learner needs, the purpose and design of an intervention program for pronunciation, and the unconventional approach of using choral singing as an effective tool for improving pronunciation.

We will go on to outline how we designed and conducted this intervention program, including the syllabus, materials and resources used in the remedial workshops. Sample recordings of student work will be analysed, and the theoretical underpinnings of the recording task design will be briefly looked at. Through the use of contrastive spoken and sung recordings, we will also introduce the concept of using choral singing to aid pronunciation and fluency in spoken discourse.

In the final part of this workshop, we will detail the results of our research, and reflect on the effects of the project on the language teaching and learning process, from the point of view of both students and educators. We will show how the aforementioned teacher intervention seemed to contribute to improved results for the participants in their end-of-course speaking assessment.

Paul Forster

Using a web-based interactive tool for formative assessment via mobile devices

While traditional paper-based forms of assessment appear to remain as the mainstay of classroom teaching, the development of digital educational technology tools has begun to afford teachers more opportunities to conduct several types of assessment in dynamic ways. Further to this, many modern language learners are digitally literate, mobile-driven and keen to engage in more paperless tasks. This presentation will focus on a student response system used by teachers to gather information from language learners via technology; thus maximising the use of mobile devices (ie. tablets, and mobile phones) in the language classroom. Programmable responses, multiple question types and automatically generated results spreadsheets are amongst some of the affordances that will be explored during this presentation. Drawing on classroom experience, this session will explore how the tool allows teachers to build customised quizzes and assists teachers in providing feedback and grading of student's work even beyond the classroom.

ABSTRACTS

Rashid AlMaamari, Gillian Wigglesworth & Ute Knoch

The University of Melbourne

The impact of teachers' beliefs on their formative assessment practices

The impact of formative assessment on students' learning and teachers' instruction has been widely acknowledged in the literature on language assessment. Teachers' beliefs have been reported to play an essential role in screening, influencing and sometimes determining the choice of classroom practices. This study utilized a mixed-method approach consisting of a questionnaire, observations, interviews and stimulated-recalls to investigate the impact of teachers' beliefs on their formative assessment practices. This study aims at promoting quality classroom assessments to improve teachers' instruction as well as students' learning through the effective use and implementation of formative assessment practices. In order to achieve this, the study sought a sound understanding of teachers' beliefs about the importance and implementation of formative assessment. It also attempted to investigate how these beliefs came to existence, how these beliefs impact on teachers' formative assessment practices and the congruence between teachers' beliefs and their classroom assessment practices. Finally, the study determined the congruence between teachers' and students' beliefs and whether they share the same beliefs about formative assessment. The study sought answers for the following research questions:

1. What beliefs do English language teachers at the Language Centre (Sultan Qaboos University) hold about formative assessment?
2. What factors are responsible for shaping these teachers' beliefs and formative assessment practices?
3. What is the influence of teachers' beliefs about formative assessment on their classroom assessment practices?
4. Do teachers and students share the same beliefs about formative assessment?

Zakiya Al Nadabi

University of Queensland/Language Centre, Sultan Qaboos University, Oman

A validation framework for an online English language Exit Test: A case study using Moodle as an assessment management system

A work-in-progress PhD research plan will be shared. This is a case study of using Moodle as an assessment management system for authoring, administration, scoring, recording of test performance, and generation of test results and item analysis statistics of an online English language Exit Test in an English-medium university. The pre-piloting phase will witness the design and development of the new test, leading to pilot-testing and post-piloting phases which will be the focus of the main study using a mixed-method research design involving quantitative and qualitative data collection and analysis procedures. This will be a validation study employing a new validation framework formulated from ideas of two evidence-based interpretive argument models, namely: 1) the Assessment Use Argument (Bachman, 2005; Bachman & Palmer, 2010); and 2) the socio-cognitive model (Weir, 2005) in order to provide multiple sources of validity evidence supporting the validity argument that the new test is a valid and reliable measurement tool for its intended purpose. Scoring validity manifested in internal reliability estimates based on classical and modern Rasch item analyses will be one component of the validity argument. Scoring validity will also be established by comparing human marking reliability (inter- and intra-rater) with the Moodle scoring engine reliability of marking constructed-response test items. Criterion-related validity will be examined by correlating test takers' scores on the new test with their scores on other language tests. Another evidence concerns the practicality and impact of using Moodle to assess language in high-stakes large-scale testing contexts, as reflected by retrospective verbal protocols of testees and invigilators in questionnaires and semi-structured interviews. This study will contribute a new unified validation research framework to the validation theory literature, particularly in technology-enhanced web-based language assessment, and will highlight the potentials of using Moodle for large-scale high-stakes testing situations along with recommendations for future applications.

Denise Angelo, Henry Fraser & Cassy Nancarrow
Australian National University

The school language data profile

Australian schools are operating in a data-driven, evidence-based improvement agenda environment. Language data however usually plays little part in the standard canon of school data collection and accountability processes. The school language data profile is a tool which has been developed as a process of identifying data sources and assembling what is and is not known about the "landscape" of a school and its students and the surrounding community. As part of the Bridging the Language Gap project (2011-2013), the school language data profile was used to drive training of school nominated personnel for implementing practices supportive of the full cohort of English as an Additional Language/Dialect learners (EAL/D).

Working with the school language data profile at school sites confirms the findings of Dixon and Angelo (2014) that school language data remains typically confused and inconsistent. As a result, students' EAL/D learning characteristics are obscured by a plethora of mismatching data. The profile is a tool for encouraging school staff and a trainer/researcher to engage in this challenging space, thereby teasing out and clarifying the issues with this language data. As a result, the reasons for invisibility of EAL/D language learning needs are made evident and can –with time and expertise– gradually be rectified. The school language data profile has been found to give teachers working in this confused data-driven space a counter-narrative which provides evidence of their linguistically complex classrooms. It empowers teachers by acknowledging learning characteristics of their students which might otherwise be invisible to onlookers relying on standard and highly privileged data sources.

Denise Angelo, Catherine Hudson & Rosalie Grant

Australian National University, Canberra, Australia; Wisconsin Center for Education Research at the School of Education, University of Wisconsin-Madison

Assessment of students with complex language contact backgrounds in Australia and the United States: the need for a contextualised research methodology

In both Australia and the US there is very little English language proficiency research on school students with complex language contact backgrounds. Consequently Indigenous Australian and American Indian students with such backgrounds receive little recognition and servicing of their language needs. When operationalised in school contexts, current theoretical 'givens' and definitional assumptions can 'misdirect' towards false-negative and false-positive interpretations of student (language) data. These 'misdirections' are generated by descriptions of 'typical' L2 acquisition and development in L2 assessment frameworks and in language and literacy education documents. Redressing this situation necessitates increasing the validity of L2 assessment frameworks for these cohorts through a rigorous redesign of 'customary' research methods into a contextualised research methodology which responds to complex, dynamic language ecologies. This paper describes a variety of case studies illustrating the need for collaborative, non-linear, iterative research methodologies: a local school project (Queensland), a World-Class Instructional Design and Assessment (WIDA) feasibility study (US) and the national Australian Early Development Index (AEDI). The paper proposes a 'consultation plus' approach that goes beyond usual ethical protocols. Engaging stakeholders, especially 'insiders', is not only vital for professional and ethical standards but also for research validity. Informed consent and valid research require a knowledge building process from the outset and throughout research phases. The paper discusses tools developed to contextualise the research method, such as a language awareness continuum, a three way linguistic conceptualisation and revisions to typical L2 proficiency descriptions involving teachers. In a context of policy driven assessment in schools, the paper calls for augmented professional and ethical guidance and research methodologies in language assessment. This paper has implications for future revisions of the ILTA Code of Ethics and Guidelines for Practice.

Felipe Araos, Constanza Pena & Sebastian Perez
The University of Queensland

Assessment practices of secondary schools in Chile: A report on three different contexts

The main objective of this presentation is to report on the assessment practices in the English subject at secondary school level in Chile. Particularly, it depicts the three local realities of secondary education in the country: public, subsidized private, and private schools. To begin, it is presented a description of the national curriculum standards established by the Ministry of Education for teaching English, which sets the minimum language contents and objectives that schools in the country must achieve. Moreover, it is explained how schools can freely design and apply individual syllabus and assessment practices if they fulfil the minimum standards. Thus, the three realities present differences in regards to language testing and assessment and how the objectives set by the national curriculum are achieved. Firstly, public schools tend to emphasize assessment on receptive skills - especially reading, closely following the national guidelines. Similarly, subsidized public schools show a predisposition to fulfil assessment objectives, emphasizing receptive and/or productive skills, partly using the national curriculum guidelines. Thirdly, private schools create their own assessment standards through personalized curriculum (including all four skills), producing higher results than the other schools. These differences on assessment practices, which are tightly linked to curriculum design, have a negative impact on students' English proficiency, since there is a disparity among public, subsidized private, and private schools' learners, being public school's students the ones with the lowest knowledge of the language. Finally, these observations can be substantiated by the results of the English national test – SIMCE – which show that students from private schools obtain significantly better scores than students from the other two realities. This issue may set the basis for future research in the field in order to identify the causes which may be involved in this phenomenon and provide possible suggestions to narrow the existing gap.

Vahid Aryadoust
National University of Singapore

Comparing the predictive power of regression, CART, and Artificial Neural Networks: An English reading comprehension study

Studies show that language learners' reading proficiency is predictable by their linguistic resources such as grammar and vocabulary. Learners with rich and flexible resources achieve high reading comprehension scores but learners with lower grammar and vocabulary knowledge typically perform poorly on reading and listening tests. This study examines the relationship between reading proficiency and language components (i.e., lexico-grammatical knowledge) by comparing three data mining approaches including linear regression, classification and regression tree (CART), and a multilayer perceptron artificial neural network (ANN). The independent (input) variables are the test performance of 300 English language learners on multiple grammar, vocabulary, and written expression subtests. The dependent variable includes students' reading comprehension measures.

The data-mining models in this study take a reductionist approach. They explore the data to identify the patterns that can explain the variance. The algorithms used in each of the models have three components: representation, evaluation, and search. Representation is the postulated relationships among data, which are to be tested. The postulations are evaluated in the following stage where fit statistics are generated to determine the validity of the postulations. These two stages are iterative and often yield a number of models, among which the optimal model is selected in the final stage. The study shows that the ANN outperformed the other two models, although CART achieved a fairly high level of accuracy. The study provides guidelines for model selections in different research designs.

Vahid Aryadoust
National University of Singapore

On the psychometric quality and criterion validity of self- and peer-assessments of the oral presentations of first-year science students

Development and assessment of science students' oral presentation skills has been recognized as a primary requirement of many university communication courses. This study examines the

psychometric quality and concurrent validity of self- and peer-assessment of oral presentation skills of first-year science students. The participants from three intact classes ($n = 40$) received training for 12 weeks on delivering effective presentations. They also received instruction on using a tertiary-level English oral presentation scale comprising three subscale (verbal communication, non-verbal communication, and Connect and Organization) measured by 18 items. For their final projects, each student was given 10 to 12 minutes to present on one of the five compulsory science books for the module and was rated by the tutor, peers, and himself/herself. Many-facet Rasch measurement (MFRM), correlation, and analysis of variance were performed to mine the data collected. The results show that the student raters, tutor, items, and rating scales achieved fairly high psychometric quality, though a small number of assessments exhibited bias. While all biased self-assessments were underestimation of presentation skills, peer- and tutor-assessment bias had a mixed pattern. In addition, self-, peer-, and tutor-assessments had low to medium correlations on the three subscales and a statistically significant difference was found between the mean scores of the assessments. Educational and assessment implications are discussed.

Vahid Aryadoust & Susan Tan
National University of Singapore

Predicting integrated writing performance from lexico-grammatical knowledge and reading proficiency using Structural Equation Modeling

Integrated writing skills assessment has a fairly long history at the National University of Singapore, although this type of assessment has only recently been researched in language assessment. The present study seeks to examine whether the integrated writing test performance can be predicted by test takers' lexico-grammatical knowledge and reading comprehension performance. The integrated writing test emulates academic writing tasks: it comprises two reading passages (word count approximately 1200 words) which students are required to read and subsequently write a 500-to-700-word essay in response to an argumentative writing prompt. The essays are marked on an analytical scale including Content, Organization, and Language and validated through many-facet Rasch measurement. The test takers are also required to take a reading and lexico-grammatical test battery comprising 80 multiple choice items, which are also validated through Rasch measurement. Using structural equation modeling, we used students' reading and lexico-grammatical test scores to predict Content, Organization, and Language. Findings show that lexical knowledge significantly predicted writing performance operationalized as a latent variable indicated by Content, Organization, and Language, though grammatical knowledge only predicted the Language scores. Grammatical knowledge, however, predicted the reading proficiency of test takers significantly but reading proficiency itself had a weak effect on integrated writing test performance. In addition, lexical knowledge was a moderate but significant predictor of the Organization scores. These results indicate that integrated writing performance is primarily influenced by test takers' lexical knowledge, although reading proficiency and grammatical knowledge also have some direct or indirect contribution.

Craig Baird & Patricia Dooley
Curtin University, Western Australia

What can you see in those clouds? : An image based PELA

The research and development of the Post Entry Language Assessment (PELA) tool described here was driven by new English language proficiency policies introduced by the university at which the authors lecture. This paper presents a new approach to testing and assessing student language abilities of higher degree students by means of a diagnostic test having everyday images rather than written instructions to stimulate understanding and communication. Here, the authors propose using a library of photographs covering all aspects of life, coupled with the most minimal of written instructions, to encourage students to communicate their interpretation of the situations depicted or to develop stories around those images that allow for individual ways of thinking and expressing ideas. Student written responses to simple images will be analysed in terms of understanding or interpretation of the topic image, writing skills, creative thinking and basic grammar. The proposed PELA will provide a level test platform that can be readily changed for every individual taking the assessment, or tailored for cultural groups via the choice of images. Instructions will remain the same across all tests, but having unique images will avoid potential shared responses. It is proposed here that the PELA will take the form of a three column table in which the first column will have a series of

images; the second column will have a simple instruction; and the third column will be for the student to write their responses. Two additional columns will only be visible to the assessor. These will provide the assessment criteria and grade boxes respectively. This format can be readily applied to paper based or electronic delivery methods at all levels of study, including transnational students. Assessment of the student responses will involve common test tools for language testing.

Brad Blackstone, Wu Siew Mei & Vahid Aryadoust

National University of Singapore

The reliability of automated rating machines in predicting the quality of reflective blog posts based on limited linguistic features

Though significant discussions in writing assessment literature focus on understanding the relationship between the quality of second language (L2) students' texts measured by human judges and linguistic features identified by automated rating engines such as Coh-Metrix, little attention (if any) has been given to assessing reflective essays presented as individual student blog posts in a tertiary level communication course. The present study examines the relationship between the linguistic features of the reflective blog posts of Asian university learners enrolled in a professional communication course as measured by Coh-Metrix and these posts' quality as assessed by human raters in discrete assessments. Rather than using traditional linear regression methods, the data was subjected to classification and regression trees (CART) to address this specific research question:

How might Coh-Metrix indices of linguistic features including lexical diversity, syntactic complexity, word frequency, and grammatical accuracy relate to the assessment of these reflection essays made by the instructor?

This study uses the data from 104 tertiary students enrolled in a writing module. They completed four writing tasks at four time points (i.e., Pre-Course, Mid-1-Course, Mid-2-Course, and End-Course), yielding 416 essays, which were marked holistically by both human raters and Coh-Metrix. 84 linguistic features for each essay (including vocabulary sophistication, lexical diversity, syntactic sophistication, and cohesion statistics) were recorded.

A description of the nature of the reflective blog posts will be presented, along with the rationale for this study, more details on the methodology used for analyzing each post and preliminary findings. It will also be argued that using CART modelling to predict essay quality from linguistic features is novel. Unlike linear regression models, CART relaxes normality assumption, optimizing the predictive power of the analysis.

Matthew Book

Victoria University of Wellington, New Zealand

Response validity of the English Language Examination for university entry in Shanghai.

This presentation presents findings from a study that aimed to discover the response validity of a past English Language Exam paper for university entry in Shanghai. Think aloud method is widely used in research to collect test-takers' interaction with a test. The think-aloud tool is used in this study to capture the "online" thinking behind the processing of each test task. In other words, the think-aloud process reveals if a test item or task indeed engages a test-taker's language ability the way it is designed to engage this ability.

A group of eight students taking a past English Language Exam paper for university entry in Shanghai are asked to verbalise their task performance and thought processes. Each participant takes the Exam separately and sits in front of the researcher. While they are taking the Exam, they are asked what is going through their minds. The process is audio-recorded while they are taking the tests.

Data are now being collected and will be analyzed. Preliminary findings from the analysis will be reported in the presentation. Results of the study would have important implications for test developers and other stakeholders in Shanghai.

Annie Brown & Ana Maria Ducasse

ACER; RMIT University, Melbourne, Australia

Indigenous assessment: criteria discussed by raters marking undergraduate classroom based speaking tasks in diverse academic contexts

The integration of problem-based learning and graduate capabilities such as team work and oral communication into the undergraduate curriculum has heightened the importance of spoken language in academic classrooms and in academic assessments. The criteria used to assess these classroom based academic oral assessment tasks are likely to be broader than those used to assess high-stake performances such as the TOEFL for example (c.f., McNamara's [1996] distinction between strong and weak tests, and Jacoby and McNamara's [1999] concept of indigenous criteria). Given the increasing role of spoken language in assessed academic performance, research that compares how raters assess academic speaking performance for classroom tasks in different academic disciplines is sorely needed in order to increase understanding of accepted levels of oral communication in higher education at undergraduate level. Thus, this study focuses on assessment of speaking in three disciplines and reports on similarities or differences in the characteristics of speaking raters identify and use to judge performance in different academic contexts.

Forty five first year undergraduate students and seven lecturer/raters were recruited from three faculties. Each student was video-recorded performing in-class oral assessment. The activities were assessed as normal by the lecturer using indigenous assessment criteria. Next, the lecturers were recorded while participating in paired moderation sessions with a peer, where they applied and justified their assessments after watching and marking the videos of the students from their discipline. The transcriptions of 45 rater discussions during moderation sessions provide the data for a content analysis of the rater focus while applying indigenous assessment criteria.

The characteristics that raters identify and discuss when marking undergraduates speaking in diverse academic contexts make up the findings which have implications for high stakes test developers, university diagnostic entrance tests and academic English curricula to expand their understanding of speaking in different academic disciplines.

Liyong Cheng

Queen's University, Kingston, Ontario, Canada

Validating Test-takers' Voices

Test-takers' interpretations of validity as related to test constructs and test use have been widely debated in large-scale language assessment. This study contributes further evidence to this debate by examining 59 test-takers' perspectives in writing large-scale English language tests. Participants wrote about their test-taking experiences in 300 to 500 words, focusing on their perceptions of test validity and test use. A standard thematic coding process and logical cross-analysis were used to analyze test-takers' experiences. Codes were deductively generated and related to both experiential (i.e., testing conditions and consequences) and psychometric (i.e., test construction, format, and administration) aspects of testing. The main themes reflect concerns over construct representation (e.g., test administration/conditions, timing, test structure and content, and scoring) and construct-irrelevant variance (e.g., aspects of test use/consequences such as coaching/preparation, emotions/self-efficacy, and misuses of test scores vs. test purpose). Furthermore, several codes appeared to be linked. The theme of preparation and test-taking strategies co-occurred with a relatively high frequency to the themes of test administration and testing conditions (18.2%), test purpose (19.4%), and test structure and content (19.5%). This results suggest that test preparation and test-taking strategies have an impact on experiences related to test structure and content, purpose, and administration. The findings of the study indicate that there is evidence for the value of gathering test-takers' perspectives on the impact and role of test consequences on perceptions of test validity. This study offers test-takers' voices on fundamental aspects of language assessment, which bear implications for test developers, test administrators, and test users. The study also demonstrated the need for obtaining additional evidence from test-takers for validating large-scale language tests.

Sue Creagh

The University of Queensland

NAPLAN and the problem LBOTE category: Patterns in LBOTE student test answers and their English language proficiency level

Test results on the National Assessment Program: Literacy and Numeracy (NAPLAN) are disaggregated and reported according to a range of classifications in annual NAPLAN reports produced by the Australian Curriculum Assessment and Reporting Authority (ACARA). The category Language Background Other Than English (LBOTE) is one such category and captures the group of students who speak a language other than English at home, or, whose parents speak a language other than English at home. Because of the broad scope of this category definition, and its failure to take account of language proficiency level, the category LBOTE is devoid of diagnostic information; in fact, according to ACARA data, the average NAPLAN performance of LBOTE is as strong as, or better than that of non-LBOTE, English speaking students. However, the spread of performance within the LBOTE category is more indicative of the range and is greater than that of non-LBOTE. The policy and pedagogy implications of these data are dire for English as an Additional Language/Dialect (EAL/D) students, who constitute a significant subset of LBOTE students who have English language learning needs, but who are well hidden in the broad LBOTE category. In an effort to unpack this problematic category, this paper will interrogate the actual test answers of a group of Year 9 LBOTE students of migrant and refugee background, who participated in the 2011 test. Their performance will be examined in relation to their English proficiency level, as measured by the ESL bandscales. Patterns will be explored in the students' answers to determine whether there is an association between language proficiency and NAPLAN result. The practical systemic challenges of achieving such an analysis will be discussed, as well as the policy implications suggested by the outcomes of the analysis.

Judie Cross, Jill Murray & Ken Cruickshank

University of Wollongong, Australia; Macquarie University, Australia; University of Sydney, Australia

Stakeholder perceptions of test impact: The case of IELTS and overseas trained teachers

This paper reports on an IDP funded investigation of test use and impact. In New Zealand and most states of Australia the preferred way for overseas trained teachers (OTTs) to demonstrate professional level English is through the academic module of IELTS, which was originally designed to measure the language skills required for academic or training contexts. While it is not unusual for the application of test scores to extend beyond the confines of their intended use, monitoring is necessary to ensure that scores obtained do justify the conclusions that are being drawn from them.

In this qualitative study we first explored the perceptions of principals, drawn from a broad range of teaching contexts. We interviewed these stakeholders individually about their experiences with OTTs who had worked in their schools. Principals also provided their perspectives on the changing communication needs of teachers.

The second part of the study investigated the principals' reactions to current IELTS benchmarks. Focus groups were conducted in the Canberra, Sydney, Melbourne, Hobart and Auckland. Participants evaluated samples of candidates' spoken and written responses at bands 6 – 8 in the light of insider knowledge of language requirements of the school as a professional workplace. While not unanimous, there was a consensus that the benchmark of 8 for speaking was appropriate, but concerns were consistently raised about whether 7 was in fact an acceptable writing score.

As a result of this study we are in a position to make recommendations for better guidelines to be set for employing bodies as to the significance of an IELTS score as well as what it does or does not claim to demonstrate. The outcomes are also potentially beneficial in raising awareness in users of test scores, encouraging them to take responsibility for areas the test does not claim to measure.

Martin East

The University of Auckland, New Zealand

Linking theory to practice in the assessment of second language spoken proficiency: Teachers' and students' perspectives on different assessment types

This paper compares and contrasts teachers' and students' views on two different ways of assessing second language (L2) speaking. The context is recent assessment reform in New Zealand's high schools, and the roll-out of a revised assessment system in a three-year programme from 2011 to 2013. Under the former model, students' L2 spoken proficiency was assessed using a traditional summative assessment called 'converse'. At the end of the year, students would take part in a single teacher-initiated and teacher-controlled interview, with the teacher as examiner. 'Converse' often led to contrived, rehearsed and stilted performances. The new model, called 'interact', aims to capture for assessment purposes a series of genuine and unrehearsed opportunities for student-initiated, learner-centred peer-to-peer spoken L2 interactions as they take place throughout the year.

Theoretically, the new 'interact' assessment arguably represents, in comparison with 'converse', a more authentic model for spoken assessment that measures a broader construct of spoken communicative proficiency. Nevertheless, the new assessment represents a radical departure from established practice that has not been without controversy. In order to gauge stakeholder reactions to the new assessment during its period of roll-out, a survey of teachers ($n = 152$) was carried out towards the end of 2012 (two years into the assessment reform) and designed to gauge teachers' comparative perceptions on 'converse' and 'interact'. Two surveys of students in their final year of schooling were also used to gauge senior students' perceptions on the original 'converse' assessment in its final year of operation (2012, $n = 30$) and on the new 'interact' assessment in its first year of operation (2013, $n = 119$). In this paper, and drawing on different theoretical models for the assessment of speaking, I present findings from the different surveys, comparing and contrasting teachers' and students' views on 'converse' in comparison with 'interact'.

Catherine Elder

The University of Melbourne

Pleading guilty without understanding? The challenges of assessing English comprehension in a forensic context.

Language tests are generally administered to groups of test takers to measure current abilities either as indicators of achievement in specific educational or training programs or as predictors of future performance in particular academic or workplace settings. In forensic contexts, by contrast, language assessments serve as a means of examining linguistic evidence produced in the past by a single person for specific communications. In such situations, the language tests and/or testing procedures devised for large populations or for particular programs may have limited utility (Van Naerssen 2014, forthcoming).

In this paper I consider the role of language assessment in a case where a legal verdict was appealed on the grounds that the defendant had not understood the import of his guilty plea. The claim of misunderstanding was supported by linguistic evidence derived from a one-on-one post-trial assessment of the defendant's English proficiency in relation to the language demands of legal interactions and documentation surrounding the case.

The paper considers the challenges facing the language assessor in assembling credible evidence for the case including: 1) the task of determining how much proficiency was required for understanding situational demands; 2) the problem of locating appropriate assessment tasks; 3) the need to ascertain that the defendant was displaying his true proficiency during task performance; and 4) the need to report assessment outcomes in terms that would be meaningful to a non-testing audience. The assessor's efforts to address these challenges in the interests of justice for all parties are outlined.

The paper concludes by considering key differences between the approach taken here and procedures adopted in more traditional testing situations. It argues that while the particularities of each forensic case call for creative individualized solutions, protocols should be established to ensure

that language testing for such purposes is both theoretically and ethically defensible.

Courtney Farley & Denise Angelo

Language Perspectives, Education Queensland; Australian National University

Identifying and assessing the full cohort of year two English language learners in a regional town of northern Queensland

This paper reports on a project undertaken with classroom teachers of year 2 students (aged approximately 6.5 - 7.5 years old) from 6 schools in a regional centre in northern Queensland. The project involved classroom teachers in the processes of identifying all the children who were learners of English as an Additional Language or Dialect (EAL/D) and assessing their levels of proficiency in Standard Australian English (SAE) in the four macroskills of speaking, listening, writing and reading/viewing. This project was instigated due to growing concerns that school language data inadequately reflected the reality of linguistically complex teaching and learning environments of many classrooms in the region. Poor language data meant that EAL/D interventions could be overlooked in data-driven approaches to improving student achievement.

This paper describes the research processes of selecting focus students and mentoring teachers about second language learning, language data collection and student assessment. It reports that a close examination of learner attributes by these mentored classroom teachers suggested the number of EAL/D learners to be triple what had been represented in school data. The majority of newly identified EAL/D learners were Aboriginal students with contact language backgrounds. In addition, this paper outlines the experiences of classroom teachers as they re-viewed student behaviours as possible indicators of second language proficiency levels. These findings are significant for language assessment and testing in school contexts, as they reveal gaps in school language data, classroom teacher capacity and second language proficiency tools.

Courtney Farley, Rose Callaghan, Bernie Bliss & Evelyn Griffiths

Language Perspectives, Education Queensland

Development of mainstream classroom teacher tool to self-assess ability and confidence when implementing differentiated units for EAL/D learners

Teachers in Queensland state schools are tasked by the English as an Additional Language /Dialect (EAL/D) policy to include all EAL/D learners in learning and achieving in mainstream classroom curriculum by differentiating their pedagogy. As part of an ongoing project to scaffold classroom teachers' abilities to differentiate their pedagogy, the Language Perspectives team has developed Language for Learning (L4L) units, a suite of 5 week text-based units for each primary year of the ACARA English curriculum with target language teaching strategies. Ongoing implementation of L4L units throughout 2012-13 in a wide cross-section of Queensland state schools revealed the need for a teacher reflection tool to accompany these units. The purpose of collecting this information has been to feed back up the line of school administration to inform them of the sticking points in teacher self-assessed ability to teach and assess their EAL/D learners. This poster presents the rationale behind the parameters of the tool which have been included through on the ground consultation with teachers, curriculum personnel, principals and project team members.

Silvana Flores

The University of Queensland

Washback effect on Chilean teachers of English: The case of SIMCE test

The Chilean Ministry of Education uses education quality measurement tests, Sistema de Medición de la Calidad de Educación (SIMCE), to assess the achievement of learning objectives of different subjects across the curriculum. In 2010, English SIMCE was introduced for the first time to evaluate the students' English level; this evaluation has been administered every two years across the country since then. The test assessed the students' receptive English language skills; reading and listening as well as their grammar knowledge. The results obtained nationally by the secondary Chilean students have been considered poor.

The effects a test, especially high-stake tests, may exert on teaching and learning is known as washback phenomenon (Alderson & Wall, 1993), and a considerable amount of work has been done investigating different aspects of washback in various contexts both focusing on students and teaching content. However, the impact or effects this examination may have caused on EFL teachers in Chile remain unknown.

To fill the gap, the current study explored a possible washback effect on secondary Chilean EFL teachers concerning teaching methodologies and content since the introduction of English SIMCE. A large-scale questionnaire survey and follow-up interview with a small number of teachers were conducted to collect information from 30 secondary school teachers about their perceptions of the test, their teaching methodologies and teaching contents. Results of qualitative and quantitative analysis of the data revealed presence of both positive and negative washback. Moreover, teachers' perception of the test is mainly negative. The study will contribute to a better understanding of the influences exerted by the test on ESL teachers in terms of methodology and teaching content as well as providing useful information for in-service and pre-service training.

Kellie Frost & Ute Knoch
The University of Melbourne

Building a validation argument for the Occupation English Test (OET): Investigating test taker processes on the listening subtest

The purpose of the OET is to provide test users with a valid and reliable assessment of candidates' English language ability in a health-related context. The listening subtest consists of two parts, and is designed to reflect the sorts of listening demands health professionals are likely to encounter in the workplace. The two parts are intended to draw on different skills, thereby capturing a more comprehensive and authentic listening construct.

In this study, a verbal reporting methodology was used to investigate if the listening processes that test-takers engage in on the subtest resemble those that the test tasks are designed to elicit, and to verify that the two parts of the test are tapping into different listening skills. The aim was to evaluate if the hypothesised construct, articulated in the test specifications, is in fact being operationalized by the test tasks encountered by candidates. In terms of building an overall validity argument for the OET, the study provides construct-related evidence in support of the explanation inference, one of the six principal inferences underlying the interpretation of test scores in an argument-based approach to test validation. The explanation inference relies on the assumption that the listening processes, skills and strategies elicited by the tasks are consistent with theoretical expectations.

Findings support the taxonomy of abilities in the existing test specifications, and provide evidence that many of these abilities are important for distinguishing between test takers with different levels of listening proficiency. Further, the data provide strong evidence that the two parts of the OET listening test made different and appropriate demands on test takers. Evidence also indicates that the different task types create an appropriate range of difficulty on the test. This study thereby provides valuable evidence in support of the overall validation argument for the OET.

John Gardiner & Stephen Howlett
The University of Sydney Centre for English Teaching, Australia

Efficacy: Student perceptions of English language tests as indicators of academic preparedness

We present a comparative analysis of students' qualitative responses to four current tests acceptable for university entrance requirements. Students in a university direct entry English language course were surveyed on their perceptions of test efficacy before and after test-taking, and throughout their first semester in university. Test-takers' interpretations of the efficacy of English language tests as related to the institutional construct of a 'task in process' were monitored.

Participants were invited to discuss with the researchers their test-taking experiences, focusing on their perceptions of the test as an indicator of preparedness for university study. A thematic coding process and logical cross-analysis was used to analyse perceptions. Codes were deductively generated and related to both the student experience in the language centre, and at university.

Efficacy provides a qualitative review of the students' perceptions of the capacity of English language tests to produce a meaningful indicator of preparedness for university study and graduate attributes. The study contributes further evidence to the debate on English language testing for international students by listening to student voice and examining students' perceptions and attitudes towards the efficacy of these tests in preparing them for learning at university.

The outcome produces an improved perspective of the utility of English language testing using a variety of tests. The findings offer test-takers' voice on fundamental aspects of language assessment as a prerequisite for international students' study at university in Australia.

Efficacy challenges beliefs and attitudes about teaching and learning language. The 'take home' is evidence from the first survey of test-takers' voice on fundamental aspects of language assessment as a prerequisite for international students' study at university in Australia. The implications of student voice in this context are relevant to test developers and administrators, test users, and educators. The results also demonstrate the need for obtaining additional evidence from test-takers for validating large-scale language tests.

Peter Yongqi Gu

Victoria University of Wellington, New Zealand

The place of English in the University Entrance Examination in China: An update

The University Entrance Examination (UEE) in China is the make-or-break point in reform for basic education. After more than two decades of inefficient calls for "quality education" that aims for the whole-person development of every child to his/her best potential, there is a realisation that the high-stakes UEE has become a hurdle which has to be overcome. From the end of 2013, a series of policy signals have been released, pointing to some very bold gestures in assessment reform, including the final removal of the test of English language in China's UEE system. The Ministry of Education announced in January 2014 that a consultation version of "An Overall Reform Framework for the Examination and Enrolment System" will be released in the first half of 2014.

This presentation will briefly introduce the overall assessment system, and a history of the UEE in China since 1977, and the rationale for the current assessment reform. I will next focus on the most important details in the consultation document that should have been released and formalised before November. An analysis of the policy implications in relation to the role of English in the UEE will follow. I will finally engage in a discussion of policy implementation and practice pertaining to the English language in China's basic education system.

M. Obaidul Hamid

The University of Queensland

Language assessment policy and practice: Listening to test-taker voices

Some of the recent theoretical developments including Critical Language Testing have placed emphasis on drawing on test-takers' perspectives in language testing and assessment. Although empirical work on test-takers' views and perceptions being informed by these theoretical frameworks has started to contribute to the field, very few studies have focused on test-takers of globally operative high-stakes tests. This presentation will draw on test-taker perceptions data from an IELTS study that was informed by Critical Language Testing and World Englishes. The main source of the data was a survey which was responded to by 430 IELTS test-takers (female= 45%, male = 55%) of 49 different nationality backgrounds. The presentation will specifically focus on these IELTS test-takers' voices on two policy issues: the test re-take policy and the duration of the validity of test scores. Analyses of test-takers' views indicate that the two policies are not founded on convincing theoretical, empirical or pragmatic grounds. In the absence of this important policy logic, test-takers attributed mercenary motives that are believed to guide these policies which potentially raise questions of validity, integrity and ethical basis of the test. The gaps between the policy and test-taker perspectives in terms of the two key policy issues provide justification for research on test-taker views. It will be emphasized that understanding test-taker voices may lead to reducing the gaps and formulating assessment policy and practice that are more acceptable to stakeholders of global proficiency tests including their producers and consumers.

Chao Han

Macquarie University, Australia

Measuring rater variability in interpreter performance testing: Using Classical Test Theory, Generalizability Theory and Rasch Measurement

Objectives: To examine rater variability in interpreter performance assessment, drawing upon research in second language testing;

To compare three approaches to rater variability, namely, classical test theory (CTT), generalizability theory (GT) and multifaceted Rasch measurement (MFRM).

Method: 32 interpreters participated in an assessment of English-to-Chinese simultaneous interpreting (SI) which consists of four performance tasks. Nine raters evaluated performance by each interpreter in each task, using three criteria (i.e., information completeness/InfoCom, fluency of delivery/FluDel, and target language quality/TLQual). Thus, a fully-crossed measurement design was implemented. Such programs as SPSS, mGENOVA and FACETS were employed for relevant statistical analysis.

Results: CTT: Overall, the rater group displayed high internal consistency across all task-by-criterion conditions, although differences existed across criteria and SI tasks, respectively.

Multivariate GT: For the rater facet, the non-negligible variance estimates (VEs) indicate that raters differed somewhat in terms of overall severity across three criteria. Based on covariance estimates (CVEs), raters who provided severe/lenient ratings using InfoCom did not behave similarly when using FluDel. For rater-by-interpreter interaction, the non-trivial VEs and positive CVEs suggest that raters disagreed somewhat in rank-ordering interpreters, and that this disagreement was to some extent consistent across three criteria. For rater-by-task interaction, the VEs were negligible and CVEs were close to zero. Additionally, both p^2 and Φ were above 0.80, in both univariate G analyses and multivariate G analyses using different weighting schemes.

MFRM: Overall, homogeneity statistic shows that not all raters were equally severe ($\chi^2 = 307.6$, $df = 8$, $p < 0.01$); a high separation reliability index (0.97) means heterogeneous degree of overall rater severity. Additionally, bias analyses including rater-by-interpreter, -criterion and -task interaction pinpoint differences of rater behavior.

Discussion: Results were discussed in the context of interpreter performance testing. Moreover, comparisons were made concerning utility of CTT, GT and MFRM in investigating rater variability.

Kathryn Hill, Jonathan Crichton, Joel King & Andrew Gleason

La Trobe University, Australia; University of South Australia, Australia; University of Melbourne, Australia; Melbourne Health, Australia

Defining effective communication in the psychiatric interview

Psychiatric conditions are very common and encountered in most areas of medicine. As diagnosis of psychiatric disorders is accomplished almost exclusively through the interview (e.g., Couper, 2005) it would seem important for all medical students to acquire basic interviewing skills in this discipline. However, teaching in psychiatry tends to concentrate on clinical knowledge, leaving the achievement of the necessary communication skills to some form of 'osmosis'.

This study will use discourse analysis and expert feedback on learner performance in two simulated psychiatric interviews to identify the skills underpinning effective communication in the psychiatric interview.

The main research questions are:

1. What criteria do specialists use to evaluate communication in the psychiatric interview?
2. How is the interaction achieved and how can this be made accessible to learners?

3. Is the simulated interaction sufficiently authentic for relevant learning to take place?

Two third year MD student volunteers and seven clinical psychiatrists involved in clinical education have been recruited for the study.

Data will include video-recordings and verbatim transcripts of two history-taking role-plays (with the students alternating in the junior doctor and patient roles), student commentary on video footage of the performances, and verbal and written feedback from the clinical psychiatrists on features of the communication.

Thematic content analysis will be used to identify and code emergent themes in the student commentary and clinician feedback. A detailed discourse analysis will be performed on transcripts of the performances.

The results will be used to produce a set of features of communication illustrated by relevant examples from the data. It is hoped the results of this study will inform teaching and assessment of medical students completing their specialty rotation in psychiatry.

Ngoc T. H. Hoang, Obaid Hamid & Richard Baldaulf Jr.

The University of Queensland, Australia

Involving test-takers in validating uses of high-stakes language assessments: Opportunities and challenges

In this presentation, we argue that test-takers can play a significant role in validating institutional uses of high-stakes tests.

First, the values of involving test-takers in validation research will be discussed drawing on Messick's (1989) validity theory. It is noteworthy in this conceptualisation that: 1) validity resides in the score interpretation and use rather than the test itself; 2) consequences form an integral part of validity evaluation; and 3) validation requires collecting and analysing all relevant evidence from all stakeholder groups for and against a particular interpretation and use. Therefore, additional validity evidence from test-takers is essential for a balanced judgement of validity issues because language testing literature currently comprises ample evidence from test developers and test users but scant evidence from test-takers. Moreover, as test-takers are the most profoundly affected by institutional uses of scores, these insiders' views and experiences provide vital insights into consequences of these uses. Additionally, engaging test-takers in evaluating the use of scores for making life-changing decisions about them is a needed response to the call for more ethical language testing.

Second, empirical evidence will be presented to illustrate the numerous ways test-takers can contribute to validity inquiry. For example, they can provide particularly useful evidence concerning consequences (washback and impact) unavailable from other stakeholders. Notably, test-takers' perception of test purposes can be evidence of face validity while comparison of test-takers' reactions can be used as evidence of perceived fairness of the test system. Furthermore, test-takers with relevant experience and knowledge can be consulted as a kind of "experts" regarding construct validity and content validity.

Finally, challenges to involving test-takers in validation research such as difficulty in accessing their views, making sense of and reconciling their contradictory opinions will be outlined. Nevertheless, the considerable benefits of this practice are well worth overcoming these difficulties.

Cath Hudson, Ellie Brownlow & Erika Charola

Australian National University

Identifying the language demands of assessment tasks for classroom teaching purposes

Mainstream classroom teachers in Queensland state schools have been given the role of differentiating the delivery of classroom curriculum for English as an Additional Language/Dialect (EAL/D) learners. For this cohort, differentiation involves not only diagnostic assessment and tailored pedagogies, but also specifically targeted language content. Classroom teachers often operate in

contexts with a multiplicity of EAL/D proficiency levels yet are required to deliver mainstream curriculum and assessment tasks. In order to differentiate their pedagogy effectively, classroom teachers need to identify the language demands of the curriculum and match these to their EAL/D students' current "outputs".

This poster outlines the Language for Learning (L4L) framework which has been developed for identifying significant language demands in assessment tasks. A front-ended "aspirational" (A-level) exemplar of a mainstream assessment task which has been analysed with this framework provides teachers with a multi-layered tool for scaffolding their ability to differentiate their teaching for EAL/D learners. The framework models the selection of significant language structures which go beyond the usual "vocabulary" of academic language approaches and deeper than the groupings of participant, process and circumstance. Penetrating below the "word group level" has been significant for recognising language demands for speakers of English-lexified contact languages who benefit from explicit teaching of morphosyntax in SAE. The framework enables teachers to better compare their students' output to a linguistically analysed A-level exemplar, thus assisting them to see the kind of language their students have and have not produced. Furthermore each language feature selected in the framework is briefly explained and accompanied by an example teaching activity.

Naoki Ikeda

The University of Melbourne

Integrating pragmatic and interactional features of language use into academic language assessment: What discursive features are measurable?

Pragmatic ability is a crucial part of real life communication as has been formally recognised as a part of communicative competence (Bachman & Palmer, 2010). Accordingly, it should be defined in test constructs if the test tries to measure the abilities to communicate. Pragmatic ability, as Ross and Kasper (2013) state, is therefore "a key domain in language assessment" (p.1). With the growing interest in assessment of L2 pragmatics (Hudson et al., 1992, 1995; Roever, 2005 and others), the construct of L2 pragmatics has expanded to discursive pragmatics (Kasper, 2006; Roever, 2011).

The present study aims to investigate (a) what features of the broader construct of L2 pragmatics can be elicited from 12 intermediate (with equivalent to IELTS 5.5 to 6.5) and advanced (IELTS 7.0 or above) L2 speakers of English and (b) to what extent the findings are consistent with empirical studies of L2 pragmatic and interactional competence.

Six dialogue role-plays simulating university activities were developed. Conversation analysis revealed that several common traits discussed in relation to the literature are displayed in students' performances. These included laying pre-expansion (Schegloff, 2007); mitigating imposition of the act; laying closing; understanding the on-going context; command of organising the conversation and choices of linguistic resource (e.g., bi-clausal structure, modal verbs). Many of these features correspond to the features of discursive pragmatics and the general findings are consistent with those in recent empirical studies (Al-Gahtani & Roever's, 2012; Youn, 2013 and others).

Possible construct-irrelevant factors in the task administration procedures were also examined during the interviews with the participants. The findings will be discussed in terms of the discursive pragmatic features that are measurable and can be included in the construct of an academic English test. The discussion will also focus on task design and the administration procedures to elicit the discursive features.

Noriko Iwashita & Claudia Vasquez

The University of Queensland

What makes a good performance? – analysis of test-taker discourse of the IELTS Academic Speaking Task

The paper presents a study comparing features of discourse and lexical competences observed in the IELTS speaking test performance across levels. When considering the construct of L2 speaking in test validation, it is widely acknowledged that different aspects of the language contribute differently to overall language proficiency (e.g., Higgs & Clifford, 1982). Also a strong link between linguistic

competence (grammar and vocabulary) and speaking proficiency has been recognised (e.g., Harley et al., 1990). Furthermore, discourse competence, which involves the ability to interpret and produce appropriate language beyond the sentence level, is considered to be an essential aspect of communicative language ability (Bachman & Palmer, 1996). The essence of discourse is the organisation of meaning, and as such, conceptually, lexis is at the heart of meaning-making (Halliday & Hasan, 1976).

The current study has undertaken detailed analysis of the test-taker oral discourse in order to identify the distinctive features of performance at the three IELTS speaking levels. The transcribed 60 speech samples (20 at each level) of the IELTS Academic Speaking task was analysed in terms of coherence (i.e., content schemata, thematization and staging, temporal continuity, management of old and new information, propositional structures and their organizational sequences), cohesive devices (use of reference, substitution, ellipsis, conjunction, parallel structure) and lexical richness (i.e., word-token, word-type, and diversity).

The in-depth analysis has shown more sophisticated use of conjunctions and organisation and a wider range of vocabulary were observed in the task performance by higher level test-takers than lower level test-takers, but other features (e.g., use of reference, thematization) was not clearly distinctive according to the levels. Especially other factors such as test-takers' L1 might have had an impact on some aspects of coherence regardless the proficiency levels. These findings contribute to further understanding of the nature of oral proficiency and communicative language ability.

Ute Knoch, Lyn May, Susy Macqueen, John Pill & Neomy Storch
The University of Melbourne

Transitioning from university to the workplace: Stakeholder perceptions of academic and professional writing demands

The number of international students and local students whose first language is not English studying in English-medium universities has increased significantly in the past decade. Although most incoming students need to achieve a certain minimum English entrance requirement for admission to tertiary level study, little is known about the writing standards necessary for graduating students once they enter their respective professions, despite recent interest in English language exit testing from universities. The transition point from university to the workforce remains under-researched; nevertheless, some employers have suggested that graduates are entering the workforce with insufficient language skills.

This study aimed to gain a detailed insight into the changing writing demands from the last year of university study to the first years in the workforce for two groups of professionals, accountants and engineers, and relate these to the demands of the writing component of IELTS, which is increasingly used for exit testing, although not expressly designed for this purpose. Data were collected in several phases, using a multi-method design. Interviews were held and questionnaire data collected from final year students as well as their lecturers. These interviews focused on students' and lecturers' perceptions of the writing demands both at university and in the workplace, and the students' perceptions of their preparedness for the language demands of their chosen profession, as well as the similarities and differences between IELTS language, university language and the perceived language demands of the profession. Similarly, interviews were conducted with new graduates in the workforce and employers of new graduates. These focussed on perceptions of the preparedness of young graduates for the writing demands in the workforce and the relevance of the IELTS tasks to the writing conducted in the professions. Wherever possible, participants from all four groups also provided tasks and writing samples from their respective domains and these were subjected to a domain analysis.

The study showed that university courses are only able to provide new graduates with limited preparation for the writing demands of the workforce. Many skills are only learned at work and often only several years into the profession, although this varied largely from context to context. The IELTS tasks, while generally perceived as relevant by employers and lecturers, were judged to not completely represent the writing abilities required by new graduates. The domain analysis supports the judgements by both stakeholder groups. The findings are discussed in terms of validity theory and exit testing for new graduates.

Carmen Le

Cambridge English Language Assessment

Introducing the Cambridge English Scale

Cambridge English exams are aligned to the Common European Framework of Reference for Languages (CEFR) – the international standard in measuring language ability. The Cambridge English Scale is a range of scores used to report results for Cambridge English exams, and will be introduced in January 2015. The scale was developed according to the well-documented and researched links between performance on different tests (using data from millions of candidates) and the processes by which we define and maintain standards.

Many universities and higher education institutions in Australia, New Zealand, and around the world accept Cambridge English: Advanced (CAE) as proof of English language proficiency. The Australian Department of Immigration and Border Protection (DIBP) has accepted CAE for student visas since 2011. From January 2015, DIBP will extend CAE recognition to include Temporary Graduate, Skilled, Former Resident, and Work and Holiday visa programmes.

The Cambridge English Scale is an enhancement to results reporting that will allow us to report numerical scores for multiple Cambridge English exams on one consistent scale. This scale, which is the product of many years of research, will replace the standardised score and candidate profile used for our exams taken before 2015. We welcome delegates to see us at the Cambridge English stand to find out more about the Cambridge English Scale and our exams.

Susy Macqueen, Sharon Yahalom, Hyejeong Kim & Ute Knoch

The University of Melbourne

Writing demands of healthcare professionals

In comparison to general language tests, tests of Language for Specific Purposes (LSP) should, at least theoretically, predict the way in which the candidate will perform in the target domain with greater accuracy (Elder, 2001). For the Occupational English Test (OET), a test which assesses the English proficiency of overseas-qualified health professionals wishing to practise in an English-speaking context, the issue of specificity is particularly relevant because the sub-tests for the productive skills, speaking and writing, are profession-specific. The advantage of occupation specificity is that the claims about candidates' performances are made based on writing and speaking activities which are related as closely as possible to the target professional contexts. Therefore, the validity of OET score use relies on evidence of a strong link between test tasks and the corresponding real-world requirements of health professionals' written communication. However, real-world communication is a dynamic phenomenon; tests which claim to reflect it need to be reviewed periodically to ensure that the test still captures relevant language skills. The research aim for this study was to ascertain the current relevance of the OET writing task (a referral letter). We investigated the frequency of different written tasks carried out by a range of health professionals as well as the qualities of the key genres. Taking a mixed methods approach, we conducted qualitative interviews with a small, representative sample and subsequently, developed an online survey which elicited information from a larger sample. Findings revealed that formal letters, especially referral letters, remain a common form of written communication across relevant professions in the Australian healthcare industry. Referral letter construction emerged as a complex process which is interconnected with other healthcare genres, e.g. patient notes. These findings and the implications for LSP test task design will be explored with reference to samples of real-world written communication from hospital settings.

Douglas Meyer

Victoria University of Wellington, New Zealand

'What is it you want to mean?' Assessing strategic competence in English as an international language

This work in progress will focus on the assessment of English as an international language (EIL) using strategic competence as a central construct. An assessment tool has been developed to elicit

strategic competence through an interactive communicative activity that makes several accommodations for the participants. This tool aims to measure strategies such as conversation repair, rephrasing, clarification, speech accommodation, and working towards a mutually shared understanding – key elements of EIL discourse. The results of this assessment will attempt to answer the question “What constitutes effective international communication?”

The participants are government officials from developing Asian and African countries who are representative of educated multi-lingual EIL users and who operate in a variety of international contexts. In such contexts, it has been suggested that accommodations be made for EIL users, such as penalizing only the errors which result in miscommunication, using EIL users as raters of an interactive communicative activity, and evaluating the performance in terms of successful task fulfillment (Canagarajah, 2006; Elder & Davies, 2006). These tasks and activities were created using a top-down approach based on the literature, and bottom-up approach based on the results of a survey and interview questions. It is argued this method has created an authentic and representative sample of the target language domain. Of course, any assessment tool must address important issues of validity, reliability, and practicality. These issues will be addressed, as well as any inferences that can be made from this test about the effective and strategic use of English as an international language.

Sascha Mitchell

University of Queensland International College, Brisbane, Australia

Replacing new for old: How to implement a change to a writing-from-sources test in an EAP course

The current writing assessment of a local English for Academic Purposes Direct Entry course is comprised of two single-prompt timed essay tests and a research assignment. The timed writing tests are often viewed by teachers and students on the course as far less relevant or authentic than the research assignment, yet are usually given a great deal more classroom teaching time. Considering that much academic writing in a university context requires, to a large extent, writing based on input from source texts (Zhang, 2013), a writing test that more fully represents the constructs of academic writing whilst still maintaining high levels of practicality and reliability has been proposed for the course.

This new test would take the form of a writing-from-sources essay test and would replace at least one of the existing single prompt essay tests. However, such a change to the assessment of the course would need to be based on justifiable evidence. Thus, comparative data must be collected on examinees' test performance and test users' perceptions of test usefulness, in particular, the degree to which the new test enhances construct validity (Bachman & Palmer, 1996).

This session will look at the feasibility of replacing a current single-prompt writing test with the proposed writing-from-sources test; the practicalities that constrain the design and implementation of the new test, as well as the complexities of gathering and analyzing comparative data on both test types.

Paul Moore & Michael Harrington

The University of Queensland

Tripartite conceptualizations of English language proficiency in Australian higher education and the role of post-entry language assessment

Policy-makers and practitioners in Australian higher education increasingly view English language proficiency (ELP) as a tripartite construct consisting of English proficiency for academic study, interpersonal communication, and more recently for the workplace. This paper examines this assumption and identifies the significant implications it has for the increasingly urgent issue of ELP assessment and support in Australian universities. The main focus for this support is international students who do not use English as their first language, but it is also increasingly relevant for domestic students.

The presenters will first review the evidence for positing the three distinct types of ELP. Although the assumption has great intuitive appeal, only limited research has actually examined how the three are

defined and discriminated. Similarly, only limited evidence is available concerning how the types interact with English learning and use by these students during their university study in Australia.

Concerns about construct validity arising from the paucity of research in these areas are then discussed. Although the three-way distinction suggests that ELP is best understood as a family of proficiency constructs, the validity of the respective constructs and how they relate to each other remain open questions with broad implications for research and practice. We specifically focus on the case of post-entry (or post-enrolment) language assessment (PELA) and its relationship to policy and practice which aim to support the development of English language proficiency in so-called 'at-risk' students in Australian higher education.

Johanna Motteram

The University of Adelaide

Unleashing the power of the Halliday-an third eye; 'tone' in letter writing tasks made visible

Linking an explicit theory of language in society with research into an element of the scoring criteria of the IELTS General Module Written Task 1 (a letter writing task), this presentation offers a way of seeing how interpersonal role relationships are expressed in texts written in response to a test prompt.

Within the test prompt the purpose of the letter and the relationship between the author and the audience of the letter are specified. To achieve a maximum score, candidates must compose a response which "fully satisfies the requirements of the task", with a "fully developed response", "with the tone consistent and appropriate" (www.ielts.org/PDF/UOBDS_WritingT1.pdf). Responding to a question first raised by a frustrated IELTS candidate, this study focusses on 'tone' in the context of the IELTS scoring criteria.

The gap between the prompt and the inclusion of appropriate 'tone' in the scoring criteria is bridged by the understanding, shared between well-socialised, educated, mature, proficient speakers of a language, that the prompt prescribed context carries implicit expectations of the text to be submitted. These expectations are inherent in the purpose (the social action to be undertaken, or field), the specified social relationship between the author and addressee (the role relationships, or tenor), and the selection of a letter (the symbolic organisation of the text, or mode).

I will present Systemic Functional Linguistic based analysis of tenor (Poynton 1985, 1990, Martin 1995, Martin and White 2005, Berry, Thompson and Hillier no date, amongst others) to make visible the myriad ways candidates in language tests express the relationship between themselves as authors and the audience as anticipated by the task prompt. Using colour to emphasise the cumulative nature of the deployment of tenor resources within texts, the differing patterns of management of tenor resources between high and mid scoring texts are also illustrated.

Khoi Ngoc Mai

The University of Queensland

In-service teachers' responses to a language assessment and training program in Vietnam

The paper discusses appropriateness and wash-back effect of a language assessment policy in Vietnam. As a part of a national project aiming to improve the quality of language teaching and learning, the English proficiency of Vietnamese primary and secondary school teachers of English were assessed by the Ministry of Education and Training (MOET). The results were disappointing as it was reported by MOET officials that 97%, 93%, and 98% of teachers at primary, lower secondary, and upper secondary respectively were underqualified. Due to such disappointing results, training programs were organized to help teachers reach the prescribed standard of linguistic competence. This study aims to raise awareness of the importance of listening to the voices of these in-service teachers who were the subjects of MOET's testing and assessment policy. While this research is motivated by the practical problems in Vietnam and the lack of research in the literature regarding non-native English teachers' process of language development, it also contributes insights to the refinement of narrative inquiry as a method still very new in language testing and assessment. The study employed a narrative inquiry with the supplement of a self-assessment language proficiency

survey and semi-structured interviews. 350 and 42 in-service teachers participated in the survey and interviews respectively. Among the 42 participants, five were chosen to be further interviewed in multiple times over a ten week period. The results reveal participants' dissatisfaction with the language assessment and training policy, describing it as superficial and impractical. Despite their enthusiasm towards joining a professional development courses, the participants expressed disappointments regarding how the wash-back effect manipulates the course content and practicality. A discussion of a matrix of interrelated difficulty challenging teachers' development of language proficiency led to the call for a holistic approach with better collaboration among different forces at different levels to resolve language proficiency related issues in order to draft meaningful and long term supporting plans in this context.

Nguyen Thi Huynh Loc

Victoria University of Wellington, New Zealand

How do pre-service teachers do language testing and assessment in their practicum?

Doing testing and assessment is among the routine tasks of a language teacher. As pre-service language teachers will start their teaching career after graduation, assessment knowledge is necessary to provide a strong foundation for their future profession. Research has proven the close link between teachers' assessment literacy, their knowledge of testing and assessment, and notable benefits in students' learning and teaching improvement. Teachers' assessment literacy is also believed to help teachers avoid misdirected and inappropriate educational decisions. Despite its importance, teachers' training in assessment literacy has not received enough attention in teacher education.

This presentation addresses a gap in the literature regarding how pre-service language teachers learn to do testing and assessment during their practicum. Multiple research tools including observation, in-depth interviews and stimulated recalls were employed to look at what kinds of assessment, and how a group of four Vietnamese pre-service language teachers improved their testing and assessment during their practicum. Triangulated data analysis showed the influence of the pre-service teachers' beliefs and previous experiences with testing and assessment on how they did classroom assessment. Even though they showed a clear improvement in their language testing and assessment, they felt a lack of language testing and assessment knowledge and abilities. They also expressed their need for more language testing and assessment training in teacher education program.

Understanding how the pre-service teachers carry out testing and assessment during their practicum is important for language teacher trainers because it reflects the pre-service teachers' readiness for their future job as language teachers. The findings are useful for raising the awareness of the interested stakeholders such as policy-makers, teacher trainers and universities in Vietnam about the importance of language testing and assessment training in their teacher education program.

Nguyen Thi Thuy Minh, Do Thanh Ha, Nguyen Tuan Anh & Pham Thi Thanh Thuy

Nanyang Technological University, Singapore; Vietnam National University, Vietnam

Effects of corrective feedback on L2 pragmatic development

Although there has been a growing interest in research on the effects of particular types of corrective feedback (CF) on L2 development, much of this body of research has focused on the role of CF in grammar development while giving far less attention to the role of CF in pragmatic development (Lyster et al, 2013). Several studies have investigated the effects of instruction on pragmatic development, including CF as part of the instruction but few have explored pragmatic development in relation to CF alone. This project therefore sets out to investigate the relative effectiveness of four types of CF, i.e. metalinguistic cues, recasts, explicit correction and clarification requests, on raising L2 pragmatic awareness and improving L2 pragmatic performance. It focuses on one particular written speech act, i.e. writing request emails to professors.

Four intact classes of EFL intermediate level students were recruited. The classes were randomly assigned to different CF conditions although they received similar instruction, were exposed to similar instructional materials and participated in similar learning activities. Another intact class was randomly

selected as a control group who received neither equivalent instruction nor any types of CF on their pragmatic performance. Treatment was conducted in six consecutive 45 minute sessions in which the students received instruction on three types of request emails which varied in levels of imposition (i.e. requests for an appointment, requests for feedback on draft and requests for an extension of a deadline). The students also produced several drafts of each type of request email and received the CF type assigned for their groups. Students' pragmatic performance was measured by means of a pre-test, an immediate and delayed post-tests, which consisted of a production and a judgment tasks. Findings will be discussed with pedagogical implications for teaching and assessing pragmatics in the classroom context.

Thao Pan

Queensland University of Technology

Helping Vietnamese EFL learners to develop learner autonomy through self-assessment

This paper reports on research in progress which explores the role of self-assessment in the development of learner autonomy in the Vietnamese higher education context. It focuses on how assessment of English language learning is currently conducted and how this relates to learner autonomy. Despite the official requirement from the Vietnamese government for learner autonomy, the academic context of Vietnamese higher education currently provides little space for autonomous learning. The idea of self-assessment is relatively unfamiliar, and has therefore been little investigated in this context. However, based on data from one study in an educational research about the understanding of learner autonomy in Vietnamese higher education, this paper suggests that self-assessment may be a useful operational approach for fostering learner autonomy in Vietnam and the East Asian region. There is a corresponding need to raise awareness of principles of self-assessment and to recognise its validity in the local context.

John Pill & Sally O'Hagan

The University of Melbourne

Language testing in medical registration procedures: Perceptions of the role of language and communication skills

Migrating doctors' training and work experience can differ from what is expected in the destination country. The language of their previous study and practice may not match the new context either. Regulatory authorities in the new jurisdiction seek to ensure that quality of care and patient safety are maintained through the implementation of registration pathways which include assessment of language and communication skills as well as of professional knowledge and clinical skills.

Different practices have developed in different countries. In the United States assessment of spoken English proficiency is integrated into a test of clinical skills taken by all medical graduates regardless of place of training, while Australian law requires language skills to be assessed separately from professional competence. Revised policies are under consideration in the United Kingdom to manage tension between European Union freedom of movement legislation and professional accountability, following high-profile cases of malpractice involving doctors trained in other jurisdictions and media coverage about poor language skills among such doctors.

As members of a research team on a large project in Australia, we investigated what doctors value in doctor-patient consultations and subsequently elicited their opinions to determine passing standards for the speaking component of a specific-purpose English test for health professionals. In workshops held over the course of the study, doctors expressed their views on the relationships between language, communication and professional skills. Our paper seeks to relate the doctors' perspectives to the different testing regimes used in countries where English is the dominant language. We discuss the various views of language and communication skills implied in the assessment procedures. We reflect on issues fundamental to language testing regarding, for example, the definition of the construct being measured and test fairness for test takers. The complex nature of language testing in practice is highlighted.

Ali Rastgou

The University of Melbourne

Assessing the impact of teacher feedback on accuracy in the writing of EFL learners: A longitudinal study

Since Truscott's (1996) argument against the application of written corrective feedback (WCF) in English-as-a-foreign/second-language (EFL/ESL) writing, it has been argued that research should continue before a generalizable decision is reached for or against it. And indeed, a large number of studies have been spurred by Truscott's claims albeit producing mixed results. One reason for such mixed results may be attributed to research design issues. There has been growing concern over methodologies adopted to conduct such studies, such as the volume of feedback offered, number of treatment occasions, ethical problems with the control group, and student dedication to teacher feedback. This study aimed to address these concerns.

Adopting a longitudinal pre-test/post-test/delayed post-test design, the study investigated the efficacy of different types of written feedback on written accuracy, complexity, fluency, content and organization. One hundred EFL students of an intermediate proficiency participated in this study in an Iranian English Institute. Participants were divided into four groups: three treatment groups and a control group. The treatment groups received different types of feedback on weekly writing tasks over a 10-week period. Group 1 received WCF; Group 2 received feedback on content/organization; Group 3 received both WCF and feedback on content/organization. Students in the treatment groups had to revise their writing after receiving feedback. The control group revised their writing without being given feedback.

Using two measures of accuracy, results showed that grammatical accuracy improved significantly only in Groups 1 and 3, which received WCF. Furthermore, improved accuracy occurred without impairing syntactic complexity and fluency. The study and its findings have implications for EFL/ESL writing teachers and for researchers. The study illustrates a more ecologically valid research design which addresses principal methodological concerns with research to date on WCF. The findings are discussed in terms of its implications of teacher WCF in the language classroom.

Mehdi Riazi

Macquarie University, Australia

"An ability-in writer-in context": An interactive model of L2 academic writing performance

Using results of qualitative interview data analysis from 17 postgraduate coursework students from four universities in New South Wales, Australia, this study proposes a social interactional model of L2 academic writing. This study, which is part of a larger scale research project funded by the Educational Testing Service (ETS), aimed at documenting the processes postgraduate students went through and the strategies they used to perform on real-life academic writing tasks. The data reported here were collected through stimulated recall interviews which followed a semi-structured format in which participants' academic assignments were used to trigger recollections of their experiences of composing each text in terms of the processes and strategies. Results of the qualitative data analysis enabled us to describe and understand the dynamic processes and provide evidence for the interactive model of L2 writing performance.

Drawing on Chalhoub-Deville's (2003) approach which views L2 construct from a social interactional perspective as "an ability-in language user- in context", this study presents a parallel interactive model of L2 academic writing performance as "an ability-in writer-in context". Such a model accounts for the synergic interaction among the writer's abilities, the writing task features and the context in which the task is defined and carried out. The capacity of the model lies in the fact that it reveals how the participants co-construct language and knowledge as they try to solve a rhetorical problem, the academic writing task. Implications of the model for L2 writing assessment in large-scale proficiency tests will be discussed.

Reference

Chalhoub-Deville, M. (2003). Second language interaction: Current perspectives and future trends. *Language Testing*, 20(4), 369-383.

Mehdi Riazi

Macquarie University, Australia

The interface between test validation and research methodology: The pledge of mixed-methods research (MMR)

Drawing on McNamara (2001), this paper discusses how the research methodology used in test validation studies has expanded over time embracing more complex procedures in recent studies. The discussion will be framed within the tripartite framework of ontology, epistemology, and methodology, in research paradigms and through examples of test validation studies. It will be argued that the research methodologies used to collect and provide evidence in test validation studies reflect researchers' conceptualisation of the object of study, i.e., validity, and validation framework. The purpose is to show the link between ontology (conceptualisation of the object of study), and methodology (data collection and analysis procedures).

The second point to be discussed in this paper is how mixed-methods research in which quantitative and qualitative methodologies are incorporated in a single research design so as to provide more comprehensive evidence, offers test validation researchers greater and more grounded opportunities to investigate issues related to test validation. I will argue how the twin research methodologies, one informed by a (post)positivist paradigm with an explanatory mandate, and the other constructivist with a mission of seeking understanding, which have often been regarded as incompatible, can be drawn on complementarily to address different layers and often latent dimensions in test validation studies. To substantiate this argument, I will draw on argument-based approaches to test validation and will show how mixed-methods research can provide evidence for different inferences incorporated in the validation framework.

Finally, I outline opportunities and challenges for researchers in test validation field in the light of the two arguments presented.

Reference

McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing*, 18(4), 333-349.

Carsten Roever, Catriona Fraser & Catherine Elder

The University of Melbourne

Validation of a web-based test of ESL sociopragmatics

Pragmatics broadly describes language use in social settings (Bardovi-Harlig, 2013), and is widely acknowledged as a central component of overall communicative competence (Bachman & Palmer, 2010; Canale & Swain, 1980; Purpura, 2004). However, not many tests exist that assess second language learners' pragmatic ability, and testing of sociopragmatics (knowledge of social rules and norms about politeness, appropriateness, and conventionality) has been especially limited and suffers from problems of practicality (Roever, 2011).

In this study, we developed and validated a web-based test of sociopragmatics for second language learners of Australian English. Following an argument-based validation approach (Chapelle, 2008; Kane, 2006, 2013), we defined our target domain as everyday language use in social, workplace and service settings. We designed a web-based test delivery and rating system, which tested sociopragmatic knowledge employing productive and receptive item types.

We ran the test with 447 EFL and ESL learners in Chile and Australia as well as 50 native speakers of Australian English. We attained a reliability of Cronbach's alpha = .82, and found significant, medium strength correlations between sections as well as a four-factor structure. As expected, the native speaker group outscored both learner groups, and the ESL group outscored the EFL group. A factorial ANOVA showed that a model incorporating length of residence, proficiency, and amount of daily interaction in English best explained the data, which is in line with findings from previous research in sociopragmatics (Bardovi-Harlig & Bastos, 2011; Matsumura, 2003; Taguchi, 2011; Wang, Brophy & Roever, 2013). Our empirical findings support the generalization, explanation and extrapolation inferences of the argument-based validation approach. In terms of utilization, we

concluded that our test is most appropriate for pedagogical and diagnostic decisions with limited utility for higher stakes decisions.

Jennifer Russell

Anglican Church Grammar School

Alternative approaches to Chinese language assessment in a middle-school classroom

Current middle-school curriculum guidelines allow for greater flexibility in assessment tasks. As such, middle-school tasks in the Chinese language department at Churchie have started shifting in delivery and content design.

Macroskills are being combined in assignment-based assessment pieces. Thus far the year seven classes have combined reading with speaking, and are about to embark upon a writing and speaking task involving a children's picture book.

Combining macroskills is hardly innovative when considering recent changes within language classrooms, what is interesting though is the level of control the students had over their assessment. Students are now excited about their assessment tasks rather than nervous; attitudes are changing. Students are seeing their assessment as an opportunity for them to explore language and to share what they're capable of. This sharing though is not with their teacher, but rather their parents.

Parents were read an excerpt of The Three Little Pigs completely in Mandarin followed by a 10 minute discussion in English in which parents, using a set of provided questions as a guide, acted as facilitator as their child explained, discussed and shared their opinions about the story and their knowledge of the language thus far. After the dialogue, parents and students then completed a form, assessing their child (the parent) or reflecting on their knowledge and ability to explain (the student). The entire process was recorded and then sent to the teacher to assess pronunciation.

This task enabled transparency across the wider community, gave the boys a sense of pride and allowed me as a teacher to step back and see what the boys could do when they were in the roll of language expert.

I would like to share not only the task but some of the student and parent feedback, as the level of success was felt far beyond the classroom.

Shahrzad Saif, Jia Ma, Lyn May & Liying Cheng

Universite Laval, Quebec, Canada; Queen's University, Ontario, Canada; Queensland University of Technology, Australia

High-stakes test preparation across different contexts: Australia, China and Iran

Test preparation for high-stakes English language tests has increasingly been a research focus in language testing (e.g., Green, 2007; Tsagari, 2009), with an emphasis on the effects of test-takers' test preparation practices on their test performance. However, previous studies revealed mixed findings regarding test preparation effects (e.g., Elder & O'Loughlin, 2003; Liu, 2011). These mixed findings suggest that more empirical studies are needed to examine the nature and the context of test-takers' test preparation practices and make legitimate interpretations of test preparation effects. Therefore, this study, with a multiple-case design, investigates the nature of test preparation for high-stakes English language tests across diverse contexts (i.e., Iranian and Chinese EFL contexts and the Australian ESL context) to reach a more nuanced understanding of the interplay of contextual factors and interactions among different stakeholders within test preparation centres.

Differing from previous empirical studies using test preparation classes as the unit of analysis, this study focused on test preparation centres and collected multiple data (i.e., semi-structured interviews, classroom observations, focus group discussions) from diverse groups of stakeholders (i.e., students, teachers, administrators) in each test preparation centre for the IELTS in Australia and Iran, and for the TOEFL iBT in China. The results of data analysis within each case revealed that high-stakes test preparation practices were predominantly oriented to the test demands. The results across cases revealed important contextual evidence highlighting the differences of the nature of test preparation

practices across the three contexts, including stakeholder expectations, linguistic emphasis, and sociocultural and cognitive engagement. The findings highlight the importance of contextual factors in determining the nature of test preparation specific to and across contexts and provide empirical evidences for future studies examining test preparation under similar contexts (e.g., EFL/ESL) or involving similar groups of test-takers (e.g., Chinese/Iranian).

Ana Saldanha & Beth Foster

Defence Force School of Languages, Australia

Improvements of assessment practices at Defence Force School of Languages

Language curriculum at the Australian Defence Force School of Languages has recently been reconfigured from general proficiency to competency based Language for Specific Purposes (LSP). We teach 20 languages per year and have two assessment regimes running concurrently: proficiency assessment on legacy and short Refresher courses and CBTA on new competency-based courses. Our session will cover the process of revision of our assessment instruments and procedures as well as a number of introduced improvements we are currently trialling on most of our courses. We'll talk about equivalency and moderation across languages, relative weighting of assessment components, reading test format and conditions of assessment, distinguishing errors, challenges in gaining common understanding of proficiency descriptors and improved format of our CBT assessment tasks.

Kerensa Townsend

Queensland University of Technology

Development and implementation of a method for validation of EAP Reading and Listening Assessment

A particular university language centre offers a number of English for Academic Purposes (EAP) courses, including Direct Entry for entry into Faculty programs, EAP Plus as a preparation program for Direct Entry, and EAP Advanced for progression into Nursing or allied Health courses. Assessment tasks within the English for Academic Purposes courses offered by the institution are typically written in-house by interested teachers. Since the successful completion of both EAP Direct Entry and EAP Advanced allows students to progress directly into faculty courses, it is imperative that assessment tasks be both valid and reliable. In addition, due to considerable external pressure, student results need to accurately reflect student ability, be a reasonable predictor of student success within university contexts and effectively gate-keep.

Due to these factors, there was a need to examine the overall performance of EAP assessments currently being used and to determine whether individual items on the tests were effectively discriminating between students. Whilst data had been collected in the past, much of this was paper-based or even anecdotal. Hence, a method that allowed teachers to easily input student responses into Excel spreadsheets and that would enable basic statistical analysis of individual items was suggested and developed. Ideally, if successful the method would then be implemented across the whole of the college. The hope was that this project would also improve assessment literacy among teachers, and arouse further interest in test writing and development.

To date, Excel spreadsheets have been created for all tasks, teachers have attended Professional Development sessions and have been inputting item responses into the spreadsheets. For each test completed, reports have been prepared, along with specific recommendations for improvement of items and the test overall. Statistics generated have been used to remove some tests from circulation and have prompted the creation of others.

Duyen T. Tran

Queensland University of Technology

Assessment for learning in the EFL classroom: A case study of practices in a Vietnamese university

Assessment has the potential to support the learning of English as a Foreign Language (EFL) in Higher Education. In Vietnam, limited research has been conducted to explore this potential, given the sociocultural context in which it takes place. Adopting a sociocultural theoretical perspective, this

qualitative case study investigates the assessment practices carried out by three EFL teachers in a public university. The aim of the study was to explore how assessment within this context might support the learning of EFL and to identify impeding or facilitating factors.

The data collection methods included video-taped classroom observations with field notes and audio recordings of students working in pairs and groups, semi-structured interviews with three teachers and one executive officer, and focus group interviews with students. Related documents and students' writing with feedback from peers and teachers were also analysed, enabling a multifaceted exploration of students' involvement in assessment, the helpfulness of feedback and the use of assessment results in the teaching and learning of EFL.

Findings indicate the potential of assessment to enhance students' learning through the facilitation of the students' understanding of learning objectives, their learning engagement and reflection. The findings also reflect strong contextual influences: particularly dominant were the cultural imperative of face-saving, respect for harmony and the educational tradition, and the institutional mandated reliance on text books and the institutional use of standardized tests to certify students' learning achievements and evaluate education quality. Within this context, several assessment-related pedagogical practices appeared to have facilitative impacts on students' EFL learning. These practices include organizing and monitoring students' work in pairs and groups, giving explicit feedback and creating a supportive learning environment by providing students with continuous support in class tasks.

Dina Tsagari

The University of Cyprus

What does classroom interaction look like in high-stakes exam classes?

The language testing field places a great importance on the impact of examinations and encourages researchers to pay particular attention to the 'washback' of high-stakes language tests on teaching and learning (Alderson & Wall, 1993). However, the scope of research of test washback on teaching and learning accumulated over the past three decades has been somewhat restricted in that researchers have considered either specific aspects of exam preparation programmes on particular skills, or studied the influence of tests on teacher perceptions, attitudes or general teaching practices rather than examine details of teachers' instructional behavior or teacher-student interaction.

Motivated by constructivist and socio-cultural theories of learning (Brown, 2007; Lightbown & Spada, 2006) and the influence of Vygotsky's work (1978) in language teaching and learning, the present study explores types of feedback used in high-stakes language exam classes (FCE - Cambridge English Language Assessment) in Cyprus to support learning by transferring, correcting and expanding learners' knowledge during their preparation for the exam.

Using a qualitative research design, four language teachers preparing their students for the FCE exam were observed and audio-recorded. A total of twenty six audio-recorded lessons were transcribed and analyzed using a special coding scheme based on relevant literature in the field of test washback, feedback and classroom interaction. Analysis was supported by the use of qualitative analysis software (Atlas.ti) that was employed to analyse teacher orientations to teaching and type of feedback (evaluative, descriptive and corrective) to scaffold learning.

Based on the results, the paper discusses the complexities involved in classroom interaction in high-stakes exam contexts and makes research and pedagogical recommendations with the aim of contributing to washback research and enhancing teachers' awareness and practices of exam preparation as part of their broader assessment literacy.

Dina Tsagari & Karin Vogt

The University of Cyprus; University of Heidelberg, Germany

Assessment Literacy of Foreign Language Teachers: Findings of a European Study

Training of pre- and in-service teachers constitutes one of the most important aspects in the quality assurance of Language Testing and Assessment (LTA). For instance, foreign language (FL) teachers have to deal with standardised tests as well as their own classroom-based assessment procedures.

This means they need the necessary expertise that can be provided by training measures. In order to gauge the current level of FL teachers in LTA literacy and identify their training needs in this area, data from seven European countries were collected in a mixed methods study that used questionnaires (n=853) and qualitative data from teacher interviews (n=63) in selected countries across Europe.

Despite the small differences across countries, the results show that only certain elements of teachers' LTA expertise are developed such as testing microlinguistic aspects (grammar/vocabulary) and language skills (e.g. reading and writing). LTA aspects such as compiling and designing non-traditional assessment methods (e.g. self-/peer-assessment or portfolio assessment), grading and placing students onto courses as well as establishing quality criteria of assessments (e.g. validity, reliability or using statistics) are not. To compensate for insufficient training, teachers seem to learn about LTA on the job or use teaching materials for their assessment purposes. Teachers overall express a need to receive training across the range of LTA features identified in the study with varying priorities depending on their local educational contexts. The results of the study offer a sound basis on which teachers and trainers (academic/professional) can build a solid foundation for the enhancement of teachers' LTA literacy.

Janet von Randow & Helen Basturkmen

The University of Auckland, New Zealand

Coherence and cohesion in academic writing: An observational study of discussions between raters

Coherence and cohesion are two criteria that generally appear on rating rubrics for assessing academic writing (Knoch, 2007; Lumley, 2002) as is the case in the rubric used in the diagnostic English language needs assessment (DELNA) at the University of Auckland. Assessors typically find using the criteria of coherence and cohesion more difficult than other criteria (Cotton & Wilson, 2011) and this has been the case for the DELNA raters. This paper reports a study that aimed to develop an understanding of why the raters in this context find the criteria of coherence and cohesion difficult to use and to identify the potential limitations of the descriptors of coherence and cohesion in our current rubrics. In the study regularly scheduled training sessions for experienced raters were recorded. In these sessions the raters, in line with normal procedures, discussed scripts they had just rated with reference to the rating rubric. Transcriptions of the recordings were imported into nVivo, software that has been developed for the qualitative analysis of text. The raters' comments about cohesion and coherence (which reveal their understanding of these constructs) and about using the descriptors were analysed. In the analysis we identified the specific aspects of coherence and cohesion the raters discussed and the aspects they described as difficult to assess in general or with reference to the current descriptors. Finally, the paper describes how we are using the findings from this study to inform future rater training and in refining the descriptions of coherence and cohesion given in the current rubric.

Megan Yucel

The University of Queensland

Living with IELTS: A narrative inquiry into the lived experiences of IELTS test candidates

The IELTS test was taken by approximately 1.7 million candidates around the world in 2011 for the purposes of academic entry into tertiary institutions and immigration. This research project investigates the impact of the test through a narrative inquiry into the stories of test candidates. It explores the lived experiences of those candidates and examines the personal significance and meaning that the test holds for them.

In order to better understand the perceptions of IELTS test takers, this project has adopted the Narrative Inquiry approach, a qualitative research method which studies the stories of research participants, with an emphasis on the collaborative aspect of research between researcher and participant over time and in a sociocultural context. Whilst narrative inquiry has been used extensively in educational research, it has not been applied to English language proficiency testing and in particular, to the context of the IELTS test and its effect on English language teaching and learning. In terms of test washback and impact, this particular form of qualitative research provides a unique and illuminating view of an arguably somewhat neglected group of test users, the candidates.

Five research participants were selected from a pool of prospective candidates for the IELTS test. Data was collected from the candidates over the study period, from which the researcher produced learner narratives. The data was interpreted using a whole-text analytical approach and placed in a wider theoretical context, in order to build up a complete picture of each individual's story, and compare and contrast the experiences of the group in order to uncover themes and patterns. The relationship between learner narratives and test scores was examined in order to focus on candidates' perceptions of the test and how this relates to test performance.

PRESENTER CONTACT DETAILS

Corresponding presenter	Email address
Rashid AlMaamari	rashidalmaamari@gmail.com
Zakiya Al Nadabi	zakiya.alnaddabi@gmail.com
Denise Angelo	denise.angelo@anu.edu.au
Vahid Aryadoust	vahid.aryadoust@nus.edu.sg
Craig Baird	c.baird@curtin.edu.au
Brad Blackstone	elcbb@nus.edu.sg
Matthew Book	matthew.book@vuw.ac.nz
Liyong Cheng	liying.cheng@queensu.ca
Sue Creagh	screagh@uq.edu.au
Judie Cross	jcross@uow.edu.au
Ana Maria Ducasse	anamaria.ducasse@rmit.edu.au
Martin East	m.east@auckland.ac.nz
Catherine Elder	caelder@unimelb.edu.au
Courtney Farley	courtney.farley@dete.qld.gov.au
Silvana Flores	floressilvana1@gmail.com
Kellie Frost	kmfrost@unimelb.edu.au
John Gardiner	john.gardiner@sydney.edu.au
Peter Yongqi Gu	peter.gu@vuw.ac.nz
M. Obaidul Hamid	m.hamid@uq.edu.au
Chao Han	chao.han2@students.mq.edu.au
Kathryn Hill	k.hill@latrobe.edu.au
Ngoc T. H. Hoang	ngoc.hoang1@uqconnect.edu.au
Catherine Hudson	cathehudson@hotmail.com
Naoki Ikeda	niked@student.unimelb.edu.au
Noriko Iwashita	n.iwashita@uq.edu.au
Ute Knoch	uknoch@unimelb.edu.au
Carmen Le	Le.c@cambridgeenglish.org.au
Susy Macqueen	susym@unimelb.edu.au
Lyn May	lynette.may@qut.edu.au
Douglas Meyer	douglas.meyer@vuw.ac.nz
Sascha Mitchell	sascha.mitchell@qut.edu.au
Paul Moore	p.moore5@uq.edu.au
Johanna Motteram	johanna.motteram@adelaide.edu.au
Khoi Ngoc Mai	khoi.maingoc@uqconnect.edu.au
Nguyen Thi Huynh Loc	huynhlocnguyen@gmail.com
Nguyen Thi Thuy Minh	thithuyminh.nguyen@nie.edu.sg
Sally O'Hagan	sohagan@unimelb.edu.au

Thao Pan	phanthao@thanhdo.edu.vn
Constanza Pena	constanza.penabarahona@uq.net.au
Ali Rastgou	arastgou@student.unimelb.edu.au
Mehdi Riazi	mehdi.riazi@mq.edu.au
Carsten Roever	carsten@unimelb.edu.au
Jennifer Russell	Jennifer.Russell@churchie.com.au
Ana Saldanha	Ana.Saldanha@defence.gov.au
Kerensa Townsend	kj.townsend@qut.edu.au
Duyen T. Tran	tranduyen1178@gmail.com
Dina Tsagari	dinatsa@ucy.ac.cy
Janet von Randow	j.randow@auckland.ac.nz
Megan Yucel	m.yucel@icte.uq.edu.au

Papers in Language Testing and Assessment:

An international journal of the Association for Language Testing and Assessment of Australia and New Zealand.

CALL FOR PAPERS

Papers in Language Testing and Assessment (PLTA) is a peer reviewed international journal and one of only five journals dedicated exclusively to language testing and assessment. It is published in open access format at: <http://www.altanz.org/> and <http://ltrc.unimelb.edu.au/>.

PLTA publishes original research papers, essays/discussion papers, research digests and reviews. Submissions from both new and experienced researchers are welcome at any time.

Information about the journal, editorial board and submission guidelines can be found via the above links. Submissions and enquiries can be sent to the Editors, Sally O'Hagan and Ute Knoch c/- plta.editor@gmail.com

Forthcoming special issues:

Language assessment in higher education settings

Look out for two special issues of PLTA in December 2014 and February 2015. Guest-edited by Kathryn Hill, and considering English and languages other than English in a diverse range of contexts, the special issues feature papers on:

- pre-entry testing
 - post-entry diagnostic & placement testing
 - in-course language assessment
 - exit-level language assessment
-

Volume 3, Issue 1 now available:

Articles

- *Test of English as a Foreign Language (TOEFL): Interpretation of multiple score reports for ESL placement.* Kateryna Kokhan & Chih-Kai (Cary) Lin, University of Illinois at Urbana-Champaign
- *The effect of assessment of peer feedback on the quantity and quality of feedback given.* Rachael Ruegg, Akita International University
- *Concepts underpinning innovations to second language proficiency scales inclusive of Aboriginal and Torres Strait Islander learners: a dynamic process in progress.* Catherine Hudson & Denise Angelo, Australian National University

Book review

The Cambridge Guide to Second Language Assessment. C. Coombe, P. Davidson, B. O'Sullivan & S. Stoyloff. Reviewed by Naoki Ikeda, University of Melbourne

80+ LANGUAGE COURSES



The University of Queensland Institute of Modern Languages

Non-degree language courses
in over 30 languages.
Assessment is optional.
Up to 6 levels plus conversation.
Courses are taught by qualified tutors
with native proficiency.

Comprehensive translation and
interpreting service with NAATI
accreditation, providing official
documents for your legal,
medical and business needs.

OFFICIAL TRANSLATION



NAATI INTERPRETING



Language courses

Asia
Europe
Latin America
Middle East
Africa
and more..

Early Bird Discounts
Enrol and pay online

<https://enrolments.iml.uq.edu.au>

CUSTOMISED TUITION



Translation and interpreting

uqtranslation@uq.edu.au
Phone (07) 3346 8207

Email iml@uq.edu.au
Phone (07) 3346 8200

Level 1, Joyce Ackroyd Building no.37 Blair
Drive, St Lucia

Web www.iml.uq.edu.au

Facebook www.facebook.com/Institute.of.Modern.Languages.UQ

QUALIFIED AND ACCREDITED STAFF



MORE THAN 30 LANGUAGES



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

Institute of
Modern Languages