

# ALTAANZ Online Research Forum 2021 – Full Programme

Current Issues in Language Assessment & Testing  
2–3 December 2021

Day 1 (Thu 2 Dec) Session A		Session B
Session time AEDT (UTC+11)	Event type • Title • Presenters	Event type • Title • Presenters
8:15am	Informal gathering/ Coffee with newcomers to ALTAANZ	
8:45am	Official opening/ Welcome session	
9am	Research Study • Understanding rater perceptions of on/off-topic performance • <a href="#">Olena Rossi &amp; Judith Fairbairn</a>	
9:30am	Research Study • Comparison of performance on paper-based testing (PBT) and computer-based testing (CBT) by English-majored students in China • <a href="#">Wenjing Yu</a>	
10am	Coffee break	
10:30am	Research Study • Test Takers' Perspectives on the Current IELTS Practice and Their Test Preparation • <a href="#">Okim Kang, Vito Miao, Kevin Hirschi, Hyunkee Ahn &amp; Yongkook Won</a>	
11am	Research Study • “We prioritise her study over our work”: Understanding Parents' Role in Mediating Test Impact • <a href="#">Yan (Niles) Zhao</a>	
11:30am	Research Study • Setting standards: Using Subject Matter Experts to Decide How Much is Enough • <a href="#">Maria Treadaway</a>	
12pm–1pm	Lunch break	
1pm	Research Study • Text complexity in short Chinese essays by L1 and L2 college learners and its validity in predicting language proficiency levels • <a href="#">Qin Xu</a>	Assessment Practice (WiP) • Classroom-based assessment to inform ESOL funding: English Language Learners in primary and secondary schools • <a href="#">Jannie van Hees</a>
1:30pm	Research Study • Reading and incidental vocabulary acquisition • <a href="#">D. R. Rahul</a>	Conceptual Paper (WiP) • Intervening in the Early Childhood Education policy process: Guiding principles to support valid language assessment • <a href="#">Denise Angelo &amp; Catherine Hudson</a>
2pm–3pm	Student gathering	Teacher networking session
3pm	Coffee break	
3:30pm	Research Study (WiP) • Assessment in English pathway programs: Practices, influences and implications • <a href="#">Kyle Smith</a>	Research Study • Formative Assessment in English Language Teaching: Exploring the Enactment Practices of Teachers within Three Primary Schools in Singapore • <a href="#">Kiren Kaur</a>
4pm	Assessment Practice • Learning-Oriented Assessment: Practical Applications for the Language Classroom • <a href="#">Peter Davidson &amp; Christine Coombe</a>	Research Study (WiP) • What's Developed in Teacher Continuous Professional Development in Classroom Assessment Literacy? • <a href="#">Jiayi Li &amp; Peter Gu</a>
4:30pm–5:30pm	<b>Keynote Speech • The private and public lives of test constructs</b> • <a href="#">Susy Macqueen</a>	
5:30pm–6pm	ALTAANZ Social event	

# ALTAANZ Online Research Forum 2021 – Full Programme

Current Issues in Language Assessment & Testing  
2–3 December 2021

Day 2 (Fri 3 Dec)	Session A	Session B
Session time AEDT (UTC+11)	Event type • Title • Presenters - Zoom link	Event type • Title • Presenters - Zoom link
8:15am	Informal gathering/ Coffee with newcomers to ALTAANZ	
9am	Research Study • Language teachers' views on the future of classroom assessment • <a href="#">Ferit Kiliçkaya &amp; Agnieszka Borowiak</a>	
9:30am	Research Study • A matter of design: A case study on responding to classroom teacher views whilst maintaining a hard-won system-wide inclusive EAL/D scale • <a href="#">Denise Angelo &amp; Catherine Hudson</a>	
10am	Research Study • Oral reflection tasks: Spanish L2 learner insights on Emergency Remote Teaching assessment practices in a higher education context • <a href="#">Ana Maria Ducasse</a>	
10:30am	Coffee break	
11am	Assessment Practice • Fairer for all: innovative approaches to assessing English Language Learners in the mainstream classroom • <a href="#">Rosemary Erlam</a>	
11:30am	Assessment Practice (WiP) • Developing, validating and using a descriptor scale for academic English language proficiency in science and technology • <a href="#">Jianqiu Tian</a>	
12pm–1:15pm	ALTAANZ AGM	
1:15pm–2:15pm	Lunch break	
2:15pm–3pm	Session with PLTA editors John Read & Jason Fan	
3pm	Research Study • Constructive alignment of assessment practices in English language classrooms • <a href="#">Lynrose Jane D. Genon &amp; Chezka Bianca P. Torres</a>	Research Study • Student Engagement with Feedback from an Automated Writing Evaluation System for Formative Assessment of EFL Writing • <a href="#">Limei Zhang</a>
3:30pm	Research Study (WiP) • The CSE-based rubrics and assessment for learning in EFL classroom • <a href="#">Jiuliang Li &amp; Noriko Iwashita</a>	Research Study • Developing a Computer-Assisted English Language Test for Undergraduates for Diagnostic and Placement Purposes: Challenges and the Way Forward • <a href="#">Radhika De Silva &amp; Subhashi Wijesinghe</a>
4pm	Coffee break	
4:30pm	Research Study (WiP) • Examining the validity of the CSE-based scale for a writing assessment using the argument-based approach to validation • <a href="#">Noriko Iwashita &amp; Jiuliang Li</a>	
5pm	Research Study • Program Evaluation of University Language Placement Tests: Examining Impact of Placement for French and Chinese • <a href="#">Ivy Chen &amp; Ute Knoch</a>	
5:30pm–6pm	Awards Ceremony & Closing remarks	

Day 1 (Thu 2 Dec) Session A		Understanding rater perceptions of on/off-topic performance
Session time AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<b>Day 1</b> <b>9AM</b> <b>Session A</b>	Research Study <ul style="list-style-type: none"> <li>• Understanding rater perceptions of on/off-topic performance</li> <li>• <a href="#">Olena Rossi, Lancaster University, UK</a> &amp; <a href="#">Judith Fairbairn, British Council, UK</a></li> </ul>	<p>In tests of speaking/writing, test-takers' performance is often rated on the task achievement (fulfilment, completion) criterion. The criterion might ask the rater to decide whether the response is on-topic or off-topic. It is, however, unclear how raters make such decisions and how reliable the decisions are because little research has been done into the issue. Therefore, this presentation aims to explore the rater decision-making process when making on/off-topic decisions: what makes the decisions valid? Does the decision-making style have an effect on the reliability of the judgements?</p> <p>The presentation draws on two studies of rater perceptions of the on/off-topic construct in speaking and writing tests. In the first study, thirty raters marked 45 test-taker speaking responses and provided explanations for 135 on/off-topic decisions (three decisions per response), as well as for the CEFR level for each of the 45 responses. In the second study, three raters quality-reviewed five speaking and five writing test versions and provided expert judgements about the quality of each test task on a range of evaluation criteria, including whether the task might generate unintended – but legitimate - off-topic responses. The judgements were submitted via an online questionnaire.</p> <p>Findings from the first study demonstrated that the raters' decision-making process has an effect on the on/off-topic decisions the raters make. Findings from the second study revealed that even highly-experienced senior raters did not gain the same understanding of what test-takers are expected to address in a test task just from considering the task and in the absence of additional guidance from test designers. The studies' findings might be helpful in understanding the decision-making process behind rating on/off-topic responses. The presentation might also be of help to those involved in operational rating as suggestions on improving the validity and reliability of on/off-topic decisions will be made.</p>

Day 1 (Thu 2 Dec) Session time AEDT (UTC+11)	Session A Event type • Title • Presenters	Comparison of performance on paper-based testing (PBT) and computer-based testing (CBT) by English-majored students in China Abstract
<b>Day 1</b> <b>9.30AM</b> <b>Session A</b>	Research Study <ul style="list-style-type: none"> <li>• Comparison of performance on paper-based testing (PBT) and computer-based testing (CBT) by English-majored students in China</li> <li>• <a href="#">Wenjing Yu, The University of Queensland, Brisbane, Australia</a></li> </ul>	<p>Computer-based testing (CBT) has been widely used in large English proficiency tests worldwide. In China, considerable attention has been invested in the implementation of CBT in high-stakes language testing. However, limited research is available concerning whether CBT can be used for Test for English Majors-Band 4 (TEM 4), a test specifically targeted at English-majored undergraduate students in tertiary education in China.</p> <p>This study aimed to investigate whether testing mode could impact TEM 4 score and to identify factors that might explain performance on CBT of TEM 4. In total 92 Chinese undergraduate students were randomly assigned to one of the groups, i.e. CBT or paper-based testing (PBT) and took the test. A mixed method was employed to analyse the test performance data in the two modes and CBT group participants' computer familiarity and attitudes towards the mode via questionnaire survey (quantitative analysis) and semi-structured interviews to the CBT group participants (qualitative analyses). The results revealed that (1) test scores in the two modes were comparable; (2) two items in the computer familiarity questionnaire, i.e., comfort level of reading articles on the computer and forgetting time when using computers, positively correlated with CBT scores; and (3) participants' attitudes towards CBT did not impact test performance.</p> <p>The comparable test scores and students' positive attitude toward CBT suggest that CBT could be a viable solution to respond to the trend of employing computers in language testing and to better address both students' requirements and institutional needs. The positive correlation between participants' computer familiarity and test scores was not found in many previous studies. This implies that computer literacy is an evolving notion that should be an enduring concern of assessment designers, given the rapid advancement of diverse digital platforms and technologies (laptops, desktops, tablets, etc.).</p>

Day 1 (Thu 2 Dec)

Session A

Test Takers' Perspectives on the Current IELTS Practice and Their Test Preparation

Session time  
AEDT (UTC+11)

Event type • Title • Presenters

Abstract

**Day 1**  
**10.30AM**  
**Session A**

Research Study

- Test Takers' Perspectives on the Current IELTS Practice and Their Test Preparation

- Okim Kang, Northern Arizona University, USA, Vito Miao, Northern Arizona University, USA, Kevin Hirschi, Northern Arizona University, USA, Hyunkee Ahn, Seoul National University, South Korea & Yongkook Won, Seoul National University, South Korea

Understanding test-takers' perceptions helps access their lived experiences of validity which may lead to more socially responsive enactment of language testing and assessment (Hamid, Hardy, & Reyes, 2019; O'Sullivan & Green, 2011; Weir, 2005). The study explored test takers' perspectives on the current practices in IELTS and their needs especially in relation to the fast-growing technology as well as the global pandemic situation. It also examined to what extent these contexts affected test-takers' IELTS preparation in different geographic locations. The study involved 241 potential IELTS test takers from 3 different geographic locations (i.e., Korea=81, China=80, and the U.S. = 80). Participants who had recently taken or planned to take the IELTS tests) completed a survey which took about 60-90 minutes. The survey included background questionnaires, online and offline resource use for test preparation, and their pandemic-related perspectives. Brief online interviews were included at the end of the survey. Descriptive statistics, ANOVAs, and Linear Mixed-Effects Models were performed. The results showed that 45-60% of the participants reported their needs for changing test locations and schedules, due to the pandemic situation. Those in the U.S (M=60) looked for online resources or changed the testing centre significantly more than those in Korea (M=23) and China (M=30) ( $F=13.52, p=.000$ ) and they also preferred the computer-based test at home (M=78) significantly more than those in Korea or in China ( $F=27.33, p=.000$ ). The geographic locations yielded various test-attitude differences and explained about 17% of the variance in these perception scores (Marginal  $R^2=.168$ ). The findings offer important implications to IELTS test development and administration, and language assessment and learning. The findings of learners' perception of test preparation styles and methods can be of use to curriculum development and instructional planning in general.

Session time AEDT (UTC+11) Event type • Title • Presenters Abstract

**Day 1  
11AM  
Session A**

Research Study

- “We prioritise her study over our work”: Understanding Parents’ Role in Mediating Test Impact
- Yan (Niles) Zhao, The University of Melbourne, Australia

In language testing, the impact of high-stakes tests has drawn increasing attention and much research has explored the impact on teachers, students and other stakeholders. Despite parents being identified as one stakeholder group (e.g. Booth, 2017; Garner & Huang, 2017), few studies have focused on them. Specifically, little is known regarding how parents mediate the test impact on their children and how they experience the impact themselves. This study illuminates the role of parents in test impact by investigating the parents of candidates of the National Matriculation English Test (NMET), a language test used to inform university admissions in China.

Research questions include 1) What are parents’ perceptions of the NMET? 2) How the NMET exerts impact on parents? 3) How parents mediate the impact of the NMET on their children? A qualitative approach was adopted to address the questions. In-depth interviews were conducted with five parents of Year 12 students who had taken the NMET and were triangulated by interviews with test takers. Interviews were transcribed verbatim and analysed thematically.

This study found that parents’ perceptions of the NMET were surprisingly unanimous, supporting the fairness of the test and its use for university admissions, although their backgrounds and familiarity with the NMET varied. Additionally, the impact on parents was more salient than they perceived it to be. While parents claimed that the NMET had little impact on them, in reality, their lives largely revolved around their children as they prepared for the test. Finally, parents shaped the test impact on students differently depending on their extent of involvement in students’ English learning.

This study has important implications. Theoretically, the findings necessitate the inclusion of parents in future test impact models. At a more practical level, educators should actively engage parents in helping students cope with the test impact.

Day 1 (Thu 2 Dec) Session A		Setting standards: Using Subject Matter Experts to Decide How Much is Enough
Day • Time • Session AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<b>Day 1</b> <b>11.30am</b> <b>Session A</b>	Research Study <ul style="list-style-type: none"> <li>• Setting standards: Using Subject Matter Experts to Decide How Much is Enough</li> <li>• <a href="#">Maria Treadaway, University of Auckland, New Zealand</a></li> </ul>	<p>This presentation highlights key validity considerations when determining the standard of performance required in an LSP test. In this case, the test was developed to measure the linguistic readiness of NNES student pilots to begin their practical flight training. A standard-setting exercise was undertaken using the Ebel method to determine cut scores on the listening and reading tests. The method elicits judgements by subject matter experts (SMEs) on the difficulty and importance of test content within the Target Language Use domain as well as estimations of test taker performance at different levels, as defined in performance level descriptors (PLDs). These PLDs had been generated empirically in previous stages and were subsequently refined during the Ebel procedures. The presentation will report on the preparatory training activities and procedures undertaken by the SMEs.</p> <p>Twelve SMEs were divided into two standard setting panels so that the replicability and reliability of the procedures and cut scores could be compared. The standard deviations of the final four cut scores set by the two panels were not significantly different. The interrater reliability estimate of performance across all 12 panellists as measured by an intraclass correlation was excellent (.986), suggesting that these judgements can be generalised to the wider population of SMEs within the TLU domain. In addition, feedback surveys indicated that training materials were effective in preparing panellists, that they were confident in their cut score recommendations, and that the process was perceived as robust. These results suggest that stakeholders can have confidence in the cut scores arrived at as a result of these standard setting activities.</p>

Day 1 (Thu 2 Dec) Session time AEDT (UTC+11)	Session A Event type • Title • Presenters	Text complexity in short Chinese essays by L1 and L2 college learners and its validity in predicting language proficiency levels Abstract
<b>Day 1</b> <b>1PM</b> <b>Session A</b>	Research Study <ul style="list-style-type: none"> <li>• Text complexity in short Chinese essays by L1 and L2 college learners and its validity in predicting language proficiency levels</li> <li>• <a href="#">Qin Xu, Osaka University, Japan</a></li> </ul>	<p>I. Objectives</p> <ol style="list-style-type: none"> <li>1. To find out whether or not text complexity in short essays will vary significantly across topics and across Chinese proficiency levels;</li> <li>2. To decide which measure(s) of text complexity in short essays can serve as valid predictors of Chinese proficiency levels.</li> </ol> <p>II. Methods</p> <p>Short essays on two topics were collected through mobile apps from volunteers of 80 Chinese L2 college students in Japan and 43 Chinese L1 college students in China. Seven measures of text complexity, namely, lexical density, lexical diversity, lexical sophistication, mean length of sentence, mean length of T-unit, T-unit ratio, and error-free T-unit ratio, were tallied in the short essays. Two-way MANOVA then determined whether or not Chinese proficiency level and essay topic (including their interactions) could explain the variations across the seven measures of text complexity. Finally, multinomial logistic regression modelling identified valid measures of text complexity, in terms of predicting writers' level of Chinese language proficiency.</p> <p>III. Findings</p> <p>Topic has no effects on text complexity in short essay writings in this study. Proficiency level has strong effects on text complexity. Out of the seven measures of text complexity, lexical diversity, lexical sophistication, mean length of sentence, and error-free T-unit ratio are valid predictors of Chinese proficiency level, with an overall accuracy of 72.8%.</p> <p>IV. Conclusions</p> <p>Particular measures of text complexity can effectively reflect language proficiency level of writers, and this is true for Chinese short essay writings.</p> <p>V. Implications</p> <p>The latest format of Chinese Proficiency Test (administered by The Society for Testing Chinese Proficiency, Japan) includes short essay writing. Current study pioneers in validating short essay as desirable instrument of Chinese proficiency level. Lexical diversity, lexical sophistication, mean length of sentence, and error-free T-unit ratio are variables deserve consideration in its scoring.</p>



Day 1 (Thu 2 Dec)	Session B	Classroom-based assessment to inform ESOL funding: English Language Learners in primary and secondary schools
Session time AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<p><b>Day 1</b> <b>1PM</b> <b>Session B</b></p>	<p>Assessment Practice (Work in progress)</p> <ul style="list-style-type: none"> <li>• Classroom-based assessment to inform ESOL funding: English Language Learners in primary and secondary schools</li> <li>• <a href="#">Jannie van Hees, Linguaging Minds Ltd, Auckland, New Zealand, University of Auckland, New Zealand</a></li> </ul>	<p>English Language learners (ELLs) are eligible for Ministry of Education (MoE) ESOL funding for up to five years based on a set of criteria. Most importantly, teachers and others working with ELLs in schooling contexts make judgments about each learner’s English language capability in the four modes of language — listening, speaking, reading, listening. These judgments inform twice yearly reporting to MoE using the English Language Learning Progressions (ELLP).</p> <p>There are no prescribed assessment tools teachers must use to make their judgments in listening, speaking, reading and writing. However, there is extensive support available to teachers, provided by MoE, to select and use effective assessment tools and approaches to inform ESOL funding criteria judgments. Despite that, teachers continue to grapple with what to use and how assessment information gathered aligns with ELLP criteria which places a learner along a continuum of increasing competence.</p> <p>The presentation will journey through a set of informing assessment tools and approaches in the four language modes used in the ELLP context, and how these then can be used to interpret and submit on the ELLP criteria. The ESOL funding application each school submits twice yearly must be evidenced-based on which judgments are made. This challenges many teachers — those new to teaching and those new to a specialised ESOL and language teaching role.</p> <p>Attendees will gain insights into school-based language assessment tools — formative and summative, that inform judgments about ELLs’ english language capabilities. There will be opportunities to converse about what is presented.</p>

**Day 1**  
**1:30PM**  
**Session A**

Research Study

- Reading and incidental vocabulary acquisition
- [D. R. Rahul, National Institute of Technology Tiruchirappalli, India](#)

Studies have constantly reiterated the salience of incidental vocabulary learning through reading over intentional learning. Although intentional learning gains may seem higher immediately after the learning sessions, the learners may not retain what they learned after some time. They do not persist over time, and substantial learning can never be delivered. On the other hand, although it is gradual and slow, incidental learning happens naturally, seemingly with ease. Many studies have confirmed the significant contribution of reading to vocabulary acquisition in an incidental manner. However, only a few studies have demonstrated the extent to which incidental vocabulary learning depends on the pleasure and frequency of reading. Therefore, to fill this gap, we show how the vocabulary competence of 239 Indian learners of English as a second language is associated with their free reading, reading pleasure, and reading frequency. We administered a vocabulary level test and a questionnaire to understand their free reading capacities. We measured the internal reliability of the vocabulary test using Chronbach's alpha. We performed a correlation test to measure the causal relationship between the constructs. We also employed regression analysis to locate the threshold level above which the participants performed better in the vocabulary test. The results show that their reading frequency is significantly associated with their reading pleasure and vocabulary test scores. The findings did not show any significant association of the learners' vocabulary test scores with their mere choice of free reading. We understand that for a sizable incidental acquisition of vocabulary through reading, it is not enough that the learners engage in mere reading, but the frequency and the pleasure derived from the reading activity must be substantial. We also confirm that the findings are consistent with comprehension and pleasure hypotheses.

Day 1 (Thu 2 Dec)	Session B	Intervening in the Early Childhood Education policy process: Guiding principles to support valid language assessment
Session time AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<p><b>Day 1</b> <b>1.30PM</b> <b>Session B</b></p>	<p>Conceptual Paper (Work in Progress)</p> <ul style="list-style-type: none"> <li>Intervening in the Early Childhood Education policy process: Guiding principles to support valid language assessment</li> <li>Denise Angelo, Australian National University/Centre of Excellence for the Dynamics of Language, Canberra, Australia &amp; Catherine Hudson, Australian National University/Centre of Excellence for the Dynamics of Language, Canberra, Australia</li> </ul>	<p>The last decade in Australia has seen an increase in policy and funding initiatives in the Early Childhood Education (ECE) sector. In the policy space, there has been impetus for universal access to ECE facilities (operationalised as availability of preschools, attendance etc), and/or inclusion in education improvement and reform agendas (operationalised as improved child development outcomes). National agreements have supported universal access for the year before commencement of fulltime schooling (known variously as preschool, prep or kindergarten) for children around the age of 4. These policies have also had a focus on improving participation of vulnerable and disadvantaged children. The Indigenous policy, Closing the Gap, mirrors these efforts, with a target for access to ECE and a target assessing children’s developmental outcomes too. Furthermore, a new preschool outcomes measure to be developed by 2025 is proposed, including school readiness. Government financial outlays give impetus to these moves to measurement.</p> <p>All this activity occurs in an ECE teaching, learning and assessment context which is under-specified and under-researched for children who are of non-English Speaking Background (NESB). Furthermore, currently available assessment tools are problematic as they have potential for encouraging misrecognition of NESB children’s development. The attraction of ready-made assessment instruments is an illusory bauble. In this talk we unpack the ECE context and the language assessment ramifications for very young NESB cohorts. Given the dearth of valid assessment resources, we propose a set of principles to guide intervention in the current ECE policy process. This has the potential that the ensuing assessment tools have positive impact for NESB children and their interests are better served.</p>

Day 1 (Thu 2 Dec) Session A		Assessment in English pathway programs: Practices, influences and implications
Day • Time • Session AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<b>Day 1</b> <b>3.30PM</b> <b>Session A</b>	Research Study (Work in Progress) <ul style="list-style-type: none"> <li>• Assessment in English pathway programs: Practices, influences and implications</li> <li>• <a href="#">Kyle Smith, Queensland University of Technology, Brisbane, Australia</a></li> </ul>	<p>Many of Australia's international students complete English pathway programs (EPPs) in order to meet the English language entry requirements of their chosen Higher Education course. Despite growing public interest in the effectiveness of these programs, relatively little research has been conducted and, consequently, little is known about how the students' English language proficiency is assessed. This paper reports on a Bourdieusian multicasestudy of EPPs which investigates the assessment practices in these programs, influences on the assessment practices, and how staff rationalise them. Semi-structured interviews have been conducted with teaching and management staff at two different university English centres; key documents relating to assessment at the two centres have also been analysed. Preliminary findings indicate that there are complex assessment systems in place in the EPPs; and that the various elements of these systems can be explained by a range of influences -- including standardised tests such as IELTS -- and are rationalised in different ways. These findings extend the work of scholars in areas such as quantification, Critical Language Testing and the sociology of assessment.</p>

Day 1 (Thu 2 Dec) Day • Time • Session AEDT (UTC+11)	Session B Event type • Title • Presenters	Formative Assessment in English Language Teaching: Exploring the Enactment Practices of Teachers within Three Primary Schools in Singapore Abstract
<b>Day 1</b> <b>3:30PM</b> <b>Session B</b>	Research Study <ul style="list-style-type: none"> <li>• Formative Assessment in English Language Teaching: Exploring the Enactment Practices of Teachers within Three Primary Schools in Singapore</li> <li>• <a href="#">Kiren Kaur, National Institute of Education, Nanyang Technological University, Singapore</a></li> </ul>	<p>This presentation will share findings based on a research study that explored the formative assessment enactment practices of six lower primary teachers within three primary schools in Singapore. Using a case study approach, data from interviews and lesson observations were examined to determine teachers' formative assessment enactment practices and specifically, influences to their enactment attempts. Findings from the study suggest formative assessment implementation to be complex when the teachers conceptualised and then enacted their practice. The teachers were found to play out processes and practices related to formative assessment in diverse ways based on interpretations they held. The teachers' institutional, social and cultural settings were also found to influence their enactment practices. The presentation will conclude by discussing evidences drawn from the study before outlining avenues for further research and implications for formative assessment implementation within a given school context.</p>

Day 1 (Thu 2 Dec) Session A		Learning-Oriented Assessment: Practical Applications for the Language Classroom
Day • Time • Session AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<b>Day 1</b> <b>4PM</b> <b>Session A</b>	Assessment Practice <ul style="list-style-type: none"> <li>• Learning-Oriented Assessment: Practical Applications for the Language Classroom</li> <li>• <a href="#">Peter Davidson, Zayed University, UAE &amp; Christine Coombe, Dubai Men's College, UAE</a></li> </ul>	<p>While assessment serves many different purposes, traditionally the principle objective of assessment has been to measure students' performance in order to make judgements about placement, progress, achievement, or proficiency. More recently, however, examination boards, test-writers, and teachers have begun to explore ways in which testing can also be used to facilitate learning. In a Learning-Oriented Assessment (LOA) approach or framework, assessment is closely aligned with the curriculum, and all assessment, both formal and informal, contributes to learning. As noted by Cambridge Assessment English (n.d.), LOA emphasizes that all levels of assessment, (macro and micro) should positively impact on both the effectiveness of learning and the evaluation of learning outcomes. But what does this mean for the classroom teacher, and how does LOA impact on the classroom? The purpose of this talk is to discuss how an LOA approach can be adopted so as to maximize opportunities for meaningful student learning in the language classroom, and become an integral part of the learning process. We begin by examining the importance of "assessment for learning" in order to better understand how it laid the foundation for LOA. After defining LOA, we turn to the main focus of this talk, which is how to actually implement LOA in the classroom. We outline what teachers can do within an LOA framework before, during, and after an assessment, in order to create opportunities for student learning. The final part of this talk will address some of the challenges inherent when working within an LOA framework, and outlines the implications that taking an LOA approach has for test writers and for classroom teachers.</p>

Day 1 (Thu 2 Dec) Session time AEDT (UTC+11)	Session B Event type • Title • Presenters	What's Developed in Teacher Continuous Professional Development in Classroom Assessment Literacy? Abstract
<p><b>Day 1</b> <b>4PM</b> <b>Session B</b></p>	<p>Research Study (Work in Progress)</p> <ul style="list-style-type: none"> <li>• What's Developed in Teacher Continuous Professional Development in Classroom Assessment Literacy?</li> <li>• <a href="#">Jiayi Li, Victoria University of Wellington, New Zealand &amp; Peter Gu, Victoria University of Wellington, New Zealand</a></li> </ul>	<p>Formative assessment has become an increasingly popular topic in educational reform since the publication of Black and Wiliam's (1998) highly influential review article. In China, formative assessment has been promoted by the Ministry of Education in its curriculum standards for over two decades. A lack of implementation in classrooms has been attributed to, among other reasons, the lack of assessment literacy among classroom teachers. Research for the enhancement of classroom assessment literacy for formative purposes among teachers of English as a foreign language is still hard to find.</p> <p>This presentation reports on the initial findings of a 12-week professional development program for a group of five secondary school English language teachers in China. Baseline data were collected at the beginning of the program. The program took the form of collaborative action research in which a researcher joined the teachers to form a community of practice for continuous professional development. Readings were shared and discussed during planning and reflection sessions. The implementation sessions in real classrooms were video recorded. The participating teachers were also interviewed before and after the program. The interviews and the video recordings were transcribed and coded using NVivo 12. Patterns of development over the 12 weeks about the teachers' knowledge, beliefs and classroom assessment practices were analysed.</p> <p>Preliminary findings suggest that the teachers' continuing professional development program has had a significant impact on the growth of assessment literacy among the participating teachers. All participants improved on their knowledge and beliefs about formative assessment. However, improvement patterns in formative assessment practices were much more complex.</p> <p>The presentation will highlight a few implications in the development of classroom assessment literacy in the Chinese EFL context. We hope that these implications will be relevant to similar educational contexts as well.</p>

Day 1 (Thu 2 Dec) Session A		The private and public lives of test constructs
Session time AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<b>Day 1</b> <b>4:30PM</b> <b>Session A</b>	Keynote Speech <ul style="list-style-type: none"> <li>• The private and public lives of test constructs</li> <li>• <a href="#">Susy Macqueen, Australian National University</a></li> </ul>	<p>Test constructs are often reified as a fixed set of abilities that are defined and then operationalised. In this talk, I re-interpret the construct as a dynamic phenomenon, emerging from overlapping “spheres of activity” in diverse stakeholder worlds, from the more ‘private’ activities of test developers and test-takers to the more ‘public’ activities of marketers and policy makers. I will focus on the “perceived construct”: how constructs are understood and applied in/to test performances (for example, by test-takers, raters or algorithms). Perceived constructs are activated across a range of test-related events at different proximities to the test experience. For example, from a distance, an employer activates a perceived construct when judging the value of a language test score on an applicant’s CV. At close range, a test-taker’s perceived construct leads to the use of certain strategies and language patterns when preparing for and doing a test. I will examine some of the ways that perceived constructs guide social, political and cognitive activity in and around test-taking, whether the test infrastructure involves humans or machines. I argue that seeing constructs as a dynamic network of activity has the potential to enable more congruence between how people understand and use language assessments, and how they are experienced by test-takers.</p>



Day 2 (Fri 3 Dec)	Session A	Language teachers' views on the future of classroom assessment
Session time AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<p><b>Day 2</b> <b>9am</b> <b>Session A</b></p>	<p>Research Study</p> <ul style="list-style-type: none"> <li>• Language teachers' views on the future of classroom assessment</li> <li>• <a href="#">Ferit Kılıçkaya, Adam Mickiewicz University, Poznan, Poland &amp; Agnieszka Borowiak, University of Lodz, Poland</a></li> </ul>	<p>During the COVID-19 pandemic, language assessment practices have experienced significant disruption. Language teachers, learners, and parents, along with the language assessment sector, had to face new challenges and practices such as how to conduct online formative and summative assessment, while certain benefits emerged during the shift from paper-based to online exams. The current study is a quantitative method study of future of language assessment and views of 182 in-service language teachers working at several institutions ranging from primary to high schools in Turkey. An online survey created via Google Form was shared on several Facebook groups which included pre-service and in-service language teachers, and they were asked to complete the online survey. The data included in-service language teachers' responses to Likert-scale questions, which included five points. The responses provided by the participants of the study suggest that during the pandemic they used a variety of websites that allowed them to assess their students online, such as Quizizz, Kahoot! And Wordwall and believed that they would also use these when the classes would move to face to face. The findings of the study reveal that language teachers would move some of the assessment practices online by creating and using some mini quizzes enhanced with videos and reading activities. However, although the move seemed to be from a pen and paper test taken in a class to at-home test practices, the participants underscored the importance of the security of these exams and therefore suggested having summative assessment in the physical classroom rather than conducting them online. The responses also added that online and frequent low-stakes assessment would be feasible such as mini quizzes and review exercises.</p>

Day 2 (Fri 3 Dec)	Session A	A matter of design: A case study on responding to classroom teacher views whilst maintaining a hard-won system-wide inclusive EAL/D scale
Session time AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<b>Day 2</b> <b>9:30AM</b> <b>Session A</b>	<p>Research Study</p> <ul style="list-style-type: none"> <li>• A matter of design: A case study on responding to classroom teacher views whilst maintaining a hard-won system-wide inclusive EAL/D scale</li> <li>• <a href="#">Denise Angelo, Australian National University/Centre of Excellence for the Dynamics of Language, Canberra, Australia &amp; Catherine Hudson, Australian National University/Centre of Excellence for the Dynamics of Language, Canberra, Australia</a></li> </ul>	<p>It is a feature of the history of the development of proficiency tools for young EAL/D learners in Australia, that these tools have been made more “inclusive” of the full population of EAL/D learners. Cohorts such as Indigenous EAL/D students and refugee students with limited schooling lacked representation in the original proficiency tools of the early 1990s. It can be argued that these newer “inclusive” tools are still not fine grained enough to represent different cohorts, such as the heterogenous Indigenous EAL/D cohort and that further new tools should be produced. There is a tension between this and, in times of continued targeted defunding and the marginality of EAL/D concerns, the policy strength of common jurisdictional tools for the full population of EAL/D learners. This paper is a case study of a way to answer the dilemma through a design which retains the common tool, but adds resources, a training pathway and community connections tailored to the needs of a particular cohort. The design responds to the “views” of non-specialist classroom teachers in a previous research study about their use of a proficiency tool with young Indigenous EAL/D students who are speakers of contact languages. The design translates specialist discipline knowledge of language contact contexts and L2 developmental pathways into classroom practice.</p> <p>Government policies have increasingly emphasised the responsibility of non-specialist classroom teachers to identify and assess all EAL/D learners, provide differentiated pedagogy and monitoring EAL/D progress is implemented. The design in this case study valorises the voices of mainstream teachers as key language assessment users and opens a way to more positive impact on this cohort of young learners, while meeting a policy need for a common tool for the full EAL/D population.</p>

Day 2 (Fri 3 Dec) Session time AEDT (UTC+11)	Session A Event type • Title • Presenters	Oral reflection tasks: Spanish L2 learner insights on Emergency Remote Teaching assessment practices in a higher education context Abstract
<b>Day 2</b> <b>10AM</b> <b>Session A</b>	<p>Research Study</p> <ul style="list-style-type: none"> <li>• Oral reflection tasks: Spanish L2 learner insights on Emergency Remote Teaching assessment practices in a higher education context</li> <li>• <a href="#">Ana Maria Ducasse, RMIT University, Melbourne, Australia</a></li> </ul>	<p>The COVID-19 pandemic has irrevocably altered language teaching and assessment practice. This paper reports on a small-scale study exploring Spanish L2 learners' personal awareness of their language and culture learning through e-assessment tasks in an Emergency Remote Teaching (ERT) context, mediated by task-specific, individual spoken reflections.</p> <p>The value of reflection, for L2 writing and distance learning, has been mediated through different modalities i.e., individual spoken reflection and group spoken reflection. Building on spoken reflection, the aim of the study was to uncover student orientation to e-assessment, through reflective learning during an oral e-assessment task, to understand how innovations to assessment were received in the remote context. In the context of higher education language learning, the reason for asking test takers to reflect on assessment situations arises from the suggestion that for test-takers' thinking about what they did during a previous task might affect their performance in subsequent tasks. (e.g., Huang, 2007, Swain et al., 2009).</p> <p>The study examines L2 learners' (n= 25) reflecting on 5 multi-modal e-assessments, mediated through individually assessed oral audio-recorded post-assessment reflection tasks (n=125). A thematic content analysis within a Self-Determination Theory framework is applied to transcriptions yielding findings on feelings of competence, autonomy, and relatedness (Ryan &amp; Deci, 2000). In these findings the student voices articulate reflections on both frustrations and positive outcomes related to learning and assessment in this context.</p> <p>The language students with no previous experience of remote language learning offer insights on experiences of e-assessment and what they considered conducive to learning (Vogt, K., Tsagari, D., &amp; Csepes, I. 2020) which would be beneficial for course designers and language instructors. The learners' candid and explicit orientations, towards various types of multimodal language-learning e-assessment tasks, offer instructors information on learners' awareness of classroom-based assessment tasks being enablers for individual learning goals.</p>

Day 2 (Fri 3 Dec)

Session A

Fairer for all: innovative approaches to assessing English Language Learners in the mainstream classroom

Day • Time • Session  
AEDT (UTC+11)

Event type • Title • Presenters

Abstract

**Day 2**  
**11AM**  
**Session A**

Assessment Practice

• Fairer for all: innovative approaches to assessing English Language Learners in the mainstream classroom

• [Rosemary Erlam, University of Auckland, New Zealand](#)

A consideration of consequences is central to an understanding of validity (Fulcher, 2015). Testing systems are very unfair for subgroups of the population when they lead to decisions which have long-term implications. Auckland is the most multilingual city in New Zealand and increasingly mainstream classes in schools have large numbers of English language learners (ELLs) from diverse language backgrounds.

This presentation documents three teachers' attempts to adapt either standardised or classroom assessment procedures in their specific teaching contexts in order to provide fairer and more valid outcomes for ELLs. These teachers were completing an in-service professional development programme. A course on assessment required them to critique specific assessment processes used in New Zealand schools and develop appropriate assessment strategies for the assessment of students from diverse linguistic and cultural backgrounds. One teacher gave a fluent te reo Maori speaker access to the text used for an assessment of his English reading ability in Maori, as well as in English. She concluded that in reading assessments bilingual learners need access to texts in both languages. Another teacher innovated an assessment to provide a better indication of learners' oral interpersonal communicative abilities. A third teacher was concerned that a Maths test was being inappropriately used with English language learners and investigated whether her colleagues were making accommodations for these students and if so, how.

The presentation argues that these teachers took innovative approaches in their own teaching contexts in order to ensure fairer outcomes for learners on tests that were either not designed for them or not appropriate for them. However, the bigger issue is whether and how there could be a more systematic and valid approach to both standardised and classroom assessment, for these learners, that is, one that does not rely on the initiative of individual teachers.

Day 2 (Fri 3 Dec) Session time AEDT (UTC+11)	Session A Event type • Title • Presenters	Developing, validating and using a descriptor scale for academic English language proficiency in science and technology Abstract
<b>Day 2</b> <b>11.30AM</b> <b>Session A</b>	Assessment Practice (Work in Progress) <ul style="list-style-type: none"> <li>• Developing, validating and using a descriptor scale for academic English language proficiency in science and technology</li> <li>• <a href="#">Jianqiu Tian, Peking University, Beijing, China</a></li> </ul>	<p>The extensive deployment of Content and Language Integrated Learning (CLIL) English courses in Chinese universities as a bridge for English-Medium instruction (EMI) calls for rigorous framework for assessment of learning, which is usually based on well-devised descriptor scale for the English language proficiency. An effective framework as the Common European Framework of Reference (CEFR) is for assessing language proficiency, the adoption of CEFR in a CLIL course requires a complementary scale of academic language proficiency conceptualized as subject-specific literacies in the content area.</p> <p>This study aims to develop, validate and use such a descriptor, following a multi-method, developmental research design in three broad phases. In the intuitive phase for the initial research and development, the relevant CEFR descriptors would be accumulated and possible descriptors would be collected from sources in science education and other language proficiency scales to complement the CEFR descriptors. Then the qualitative phase for checking and improving the categories and quality of the descriptors involves workshops with teachers to sort descriptors into categories, make qualitative judgements about clarity, accuracy and relevance of the description and sort descriptors into bands of proficiency. Finally, in the quantitative phase, the best descriptors would be calibrated to a mathematical scale and the cut-offs between the levels would be set.</p> <p>Next, the scale will be trialed in the assessment of the students' language in the classroom with tasks calibrated to the scale beforehand, with the feedback employed to identify and eliminate less successful descriptors and scales. Meanwhile, the scale would be integrated into the progressive and sequentially-phased curriculum for both summative and formative purposes in the CLIL course, and be further developed into self-assessment tools for the students to monitor their learning. The trial may be conducted for several rounds to further refine and validate the descriptor scale.</p>

Day 2 (Fri 3 Dec) Session A		Constructive alignment of assessment practices in English language classrooms
Day • Time • Session AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<b>Day 2</b> <b>3PM</b> <b>Session A</b>	Research Study <ul style="list-style-type: none"> <li>• Constructive alignment of assessment practices in English language classrooms</li> <li>• <a href="#">Lynrose Jane D. Genon, Mindanao State University-Iligan Institute of Technology, Iligan City, Mindanao, Philippines &amp; Chezka Bianca P. Torres, Mindanao State University-Iligan Institute of Technology, Iligan City, Mindanao, Philippines</a></li> </ul>	<p>This qualitative study identified the language assessment practices in terms of purpose, type, and timing in four elementary language classes in the Philippines. It then evaluated the constructive alignment and content validity of the assessment and described how the constructive alignment reflects the quality of teaching and learning in these language classrooms. Findings revealed that the assessment practices are entrenched in the teachers' pedagogy serving various purposes (like monitoring, facilitation, and motivation) other than evaluation. Generally, the assessment content and tasks show alignment to targeted learning goals. Several assessments, however, do not target the intended skill but rather tapped on the sub-skill, and they reveal alignment issues as well as teaching-learning conditions in the classrooms. Findings uncover, albeit a snapshot, of the gaps not only in assessment practices but also in the articulation of curriculum goals leaving teachers to interpret and operationalize these on their own. These issues affect the students' achievement since alignment affects achievement and opportunity to learn is at optimum levels when lesson plans are aligned with benchmarks, standards, and assessments, thus increasing academic achievement.</p>

Day 2 (Fri 3 Dec) Session time AEDT (UTC+11)	Session B Event type • Title • Presenters	Student Engagement with Feedback from an Automated Writing Evaluation System for Formative Assessment of EFL Writing Abstract
<b>Day 2</b> <b>3PM</b> <b>Session B</b>	<p>Research Study</p> <ul style="list-style-type: none"> <li>• Student Engagement with Feedback from an Automated Writing Evaluation System for Formative Assessment of EFL Writing</li> <li>• Limei Zhang, Nanyang Technological University, Singapore</li> </ul>	<p>Research shows that feedback can improve English as a Foreign Language (EFL) students' writing abilities. In order to benefit their writing, students need to engage effectively with the feedback on their compositions. Recently, an increasing number of studies investigate the use of tools for automated writing evaluation in EFL writing classrooms, indicating researchers' growing interest in their potential for formative assessment. As an online automated writing evaluation (AWE) system and the most popular one in China, Pigai was used widely in college-level EFL writing courses. The current study examined the accuracy of and students' engagement with the feedback from Pigai for formative assessment of EFL writing courses.</p> <p>89 Chinese college students from an academic writing course were invited to participate in the study. This study addresses the following two research questions: 1) How accurate is Pigai feedback in terms of the errors it can identify in students' writing? 2) How did students engage affectively, behaviourally, and cognitively with the feedback from the AWE system on their writing? Quantitative and qualitative data were collected to answer the two research questions. Results showed that the accuracy of feedback from the AWE system varied across error types. Meanwhile, this study provides an understanding of students' affective, behavioural and cognitive engagement with the feedback from the system. The findings can inform discussion of whether and how to integrate the use of online writing evaluation system into writing classrooms for the purpose of formative assessment.</p>

Day 2 (Fri 3 Dec)	Session A	The CSE-based rubrics and assessment for learning in EFL classroom
Session time AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<p><b>Day 2</b> <b>3.30PM</b> <b>Session A</b></p>	<p>Research Study (Work in Progress )</p> <ul style="list-style-type: none"> <li>• The CSE-based rubrics and assessment for learning in EFL classroom</li> <li>• <a href="#">Jiuliang Li, Beijing Institute of Fashion Technology, China &amp; Noriko Iwashita, The University of Queensland, Brisbane, Australia</a></li> </ul>	<p>Assessment for learning (AfL) serves as a powerful alternative assessment approach that maximizes student learning (Assessment Reform Group, 2002). China's Standards of English Language Ability (CSE) is developed to guide ELT and learning in China (Liu, 2017). They are closely linked as both claim to serve the needs of students and teachers to decide where learners are in their learning, where they need to go, and how best to get there (Assessment Reform Group, 2002; Liu, 2017). AfL practices in China generally lack (Liu &amp; Xu, 2017), while CSE is also in need of more research for use in an educational context (Liu, 2017).</p> <p>This study explores the potential of CSE-based rubrics as a tool for AfL to be used in EFL writing classrooms. We examine how CSE-AfL is implemented in the learning, instruction and assessment in a 10-week English reading and writing class in an urban university in China. Participants are 90 third-semester students and their instructor. The implementation of AfL strategies is interspersed in three phases -- prewriting, during writing, and post-writing. The instructor and the students work collaboratively to set learning goals using the CSE rubrics, with which the students perform peer/self-assessment to assess the scripts produced by students to develop an understanding of their writing ability. Questionnaire surveys and interviews are conducted before and after the course.</p> <p>Preliminary findings suggest that the students have little experience with key AfL strategies. The first round of self/peer assessment shows more active interaction between students and their instructor. Both seem to have reaped benefits from the initial implementation of the program. The paper discusses the perceived and actual influence of CSE-AfL on the educational activities and outcomes where the students and their instructor are engaged toward language development.</p>



Day 2 (Fri 3 Dec) Session time AEDT (UTC+11)	Session B Event type • Title • Presenters	Developing a Computer-Assisted English Language Test for Undergraduates for Diagnostic and Placement Purposes: Challenges and the Way Forward Abstract
<b>Day 2</b> <b>3.30PM</b> <b>Session B</b>	Research Study <ul style="list-style-type: none"> <li>Developing a Computer-Assisted English Language Test for Undergraduates for Diagnostic and Placement Purposes: Challenges and the Way Forward</li> <li>Radhika De Silva, The Open University of Sri Lanka, Nawala, Nugegoda, Sri Lanka &amp; Subhashi Wijesinghe, The Open University of Sri Lanka, Nawala, Nugegoda, Sri Lanka</li> </ul>	<p>Covid-19 pandemic related issues have influenced the whole world and education is one of the fields that has been affected to a great extent by this pandemic. Schools and universities worldwide are currently facing numerous challenges in ensuring quality teaching and fair assessment of their students' skills and achievements. As a solution to these unprecedented challenges, many institutions have utilized virtual platforms for teaching. Even though transition to teaching via online platforms is gaining momentum, assessment online remains a challenge in many institutions, especially in developing countries. Computer-assisted language testing is adopted by language testers worldwide because of its benefits such as speed of delivery, immediate feedback, and test evaluation. This paper attempts to discuss the challenges faced by a team of test developers in the process of developing a computer-assisted English language test in a local context for the purpose of diagnosis, placement, and exemption purposes. The study also revealed the perceptions of the team of test developers on computer-assisted language assessment. The study was conducted at a public university in Sri Lanka using a qualitative research design. An open-ended questionnaire was administered to the sample followed by semi-structured interviews. The perceptions of the team towards computer-assisted language testing were positive. This paper examines a range of challenges encountered by the test developers in the process of computer-assisted English language test development in this pandemic situation.</p>

Day 2 (Fri 3 Dec) Session time AEDT (UTC+11)	Session A Event type • Title • Presenters	Examining the validity of the CSE-based scale for a writing assessment using the argument-based approach to validation Abstract
<b>Day 2</b> <b>4:30PM</b> <b>Session A</b>	<p>Research Study (Work in Progress)</p> <ul style="list-style-type: none"> <li>Examining the validity of the CSE-based scale for a writing assessment using the argument-based approach to validation</li> <li>Noriko Iwashita, The University of Queensland, Brisbane, Australia &amp; Jiuliang Li, Beijing Institute of Fashion Technology, China</li> </ul>	<p>China's Standards of English (CSE) were developed in 2018 as a yardstick for teaching and learning. With a range of criteria of English proficiency to inform teaching, learning and assessment, the CSE is expected to serve as a guideline for English language education initiatives in China (Liu, 2015, 2017; Wang, 2018). While several suggestions on applying the CSE into the writing classroom have been put forward (e.g., He &amp; Xiao, 2019), empirical studies reporting the actual use and impact of CSE and CSE-based tools on teaching and learning are rare to date. Further, it has been reported that CSE scales are not very practical for actual classroom use (e.g., Jin et al., 2018).</p> <p>The current study investigates how CSE-based initiatives could be incorporated into the English classroom to enhance teaching and learning. A new CSE-based scale was developed for a reading and writing course at a large urban university. Employing Bachman and Palmer's (2010) Assessment Use Argument framework, a validation study of the CSE based-scale was conducted in two stages following the method employed by Mendoza and Knoch (2018). In the first stage, 216 writing scripts of three different genres (i.e., 72 each for narrative, argumentative and expository) were rated by four raters. Its quality was analyzed using many-facet Rasch measurement. In addition, the raters were invited to participate in an online questionnaire survey and focus group interviews on sharing their opinions and perceptions of the rating scale. Preliminary findings of the first stage of the study found that the rating scale was of good quality and raters' comments were generally positive. The paper presents suggestions for the second stage of the study and discusses how these findings could be implied to use the CSE-based scale in classroom teaching.</p>

Day 2 (Fri 3 Dec)	Session A	Program Evaluation of University Language Placement Tests: Examining Impact of Placement for French and Chinese
Session time AEDT (UTC+11)	Event type • Title • Presenters	Abstract
<b>Day 2</b> <b>5PM</b> <b>Session A</b>	<p>Research Study</p> <ul style="list-style-type: none"> <li>• Program Evaluation of University Language Placement Tests: Examining Impact of Placement for French and Chinese</li> <li>• <a href="#">Ivy Chen &amp; Ute Knoch, University of Melbourne, Melbourne, Australia</a></li> </ul>	<p>Language placement tests used at university level are ideally connected to the language curricula into which they place learners and are reflective of local needs and student populations. However, once implemented, language placement tests are rarely evaluated to establish how well they are functioning. This study aims to fill this gap by employing a program evaluation framework (Rossi, Lipsey, &amp; Freeman, 2018) to evaluate placement tests for two languages (French and Chinese) developed and implemented at a large Australian university. In this presentation, we focus on the accuracy of placement decisions and how well overriding rules that place students based on their learning background (while overriding their placement test scores) work. These overriding rules were put in place by the language programs to address perceived difficulties of placing particular groups of students.</p> <p>Data collected included placement test responses of 1001 students who took the French and 1319 students who took the Chinese placement test, as well as final subject grades supplied by the language programmes. Survey data was collected from students, and focus groups were conducted with students and academic staff.</p> <p>Overall, we found that the placement test decisions are accurate, though the evaluation did lead to the recommendation of several revisions to the overriding rules, as the data showed some negative impacts of pushing low-scoring test takers with background experience in the language into a higher-level subject. The study has implications for the development of language placement tests and their implementation in specific contexts.</p>