



英国文化教育协会



教育部考试中心  
NATIONAL EDUCATION EXAMINATIONS AUTHORITY

---

## Connecting Assessment with Teaching and Learning

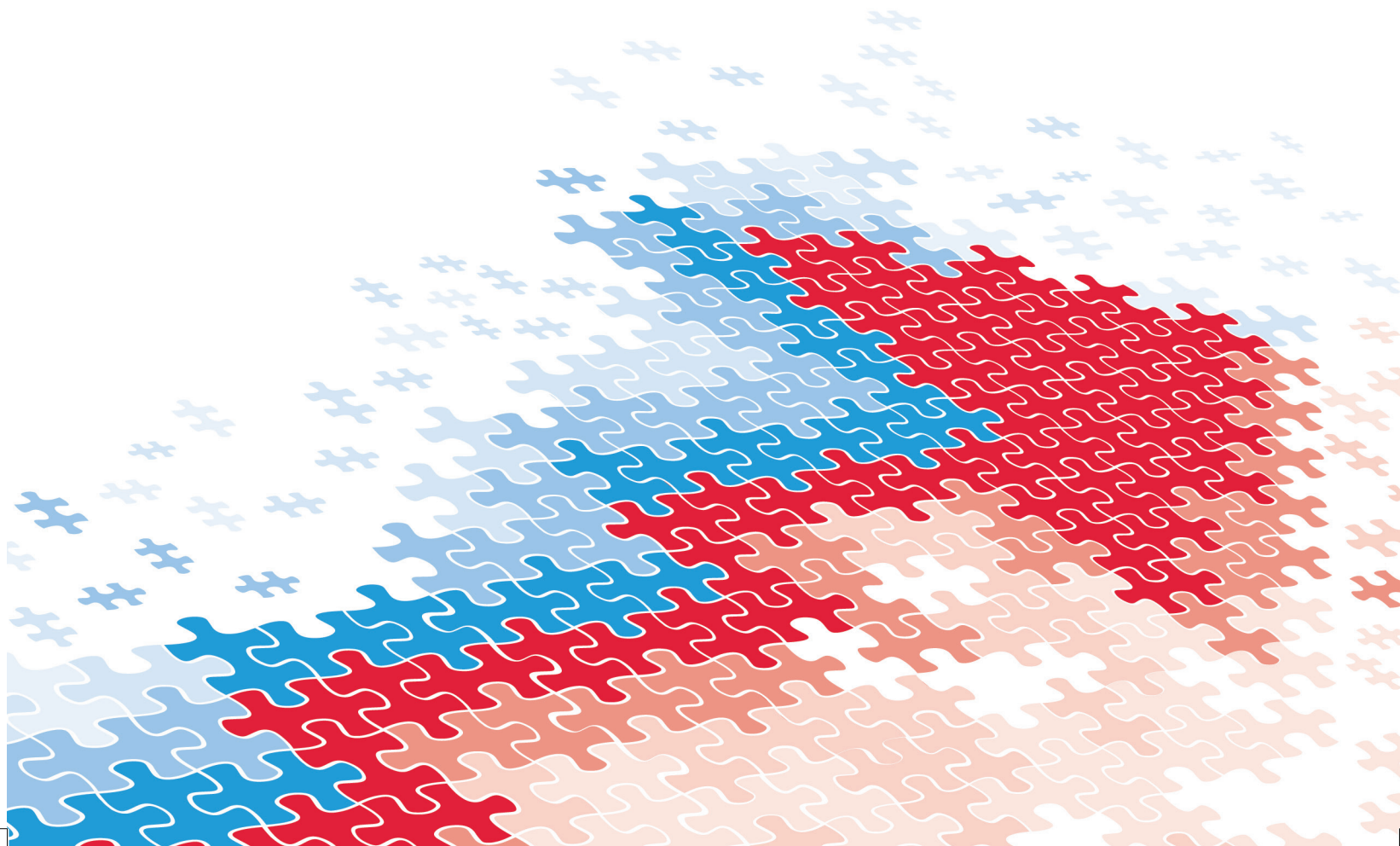
---

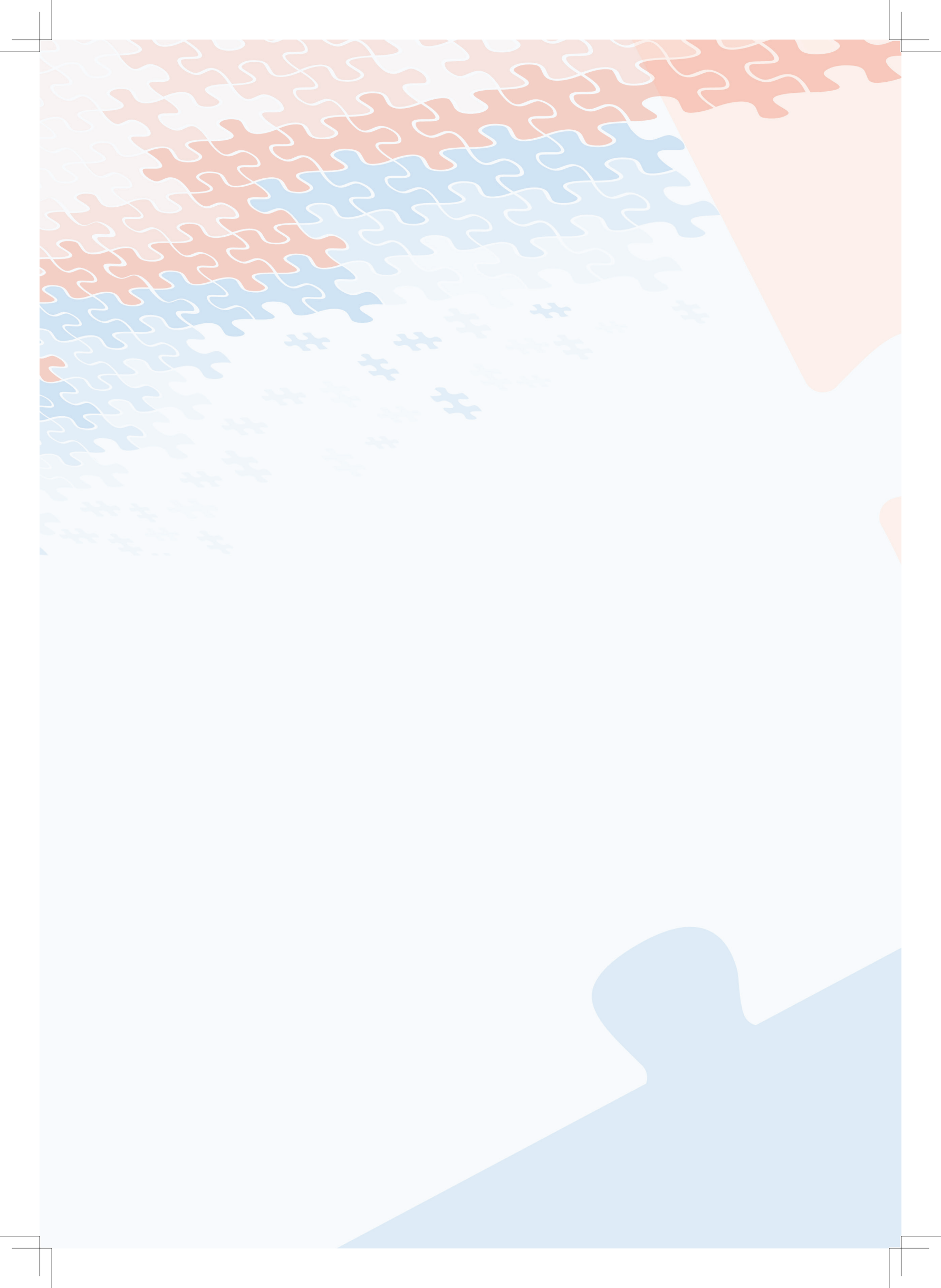
The 3rd International Conference on  
Language Testing and Assessment  
and the 5th British Council  
New Directions in English Language  
Assessment Conference

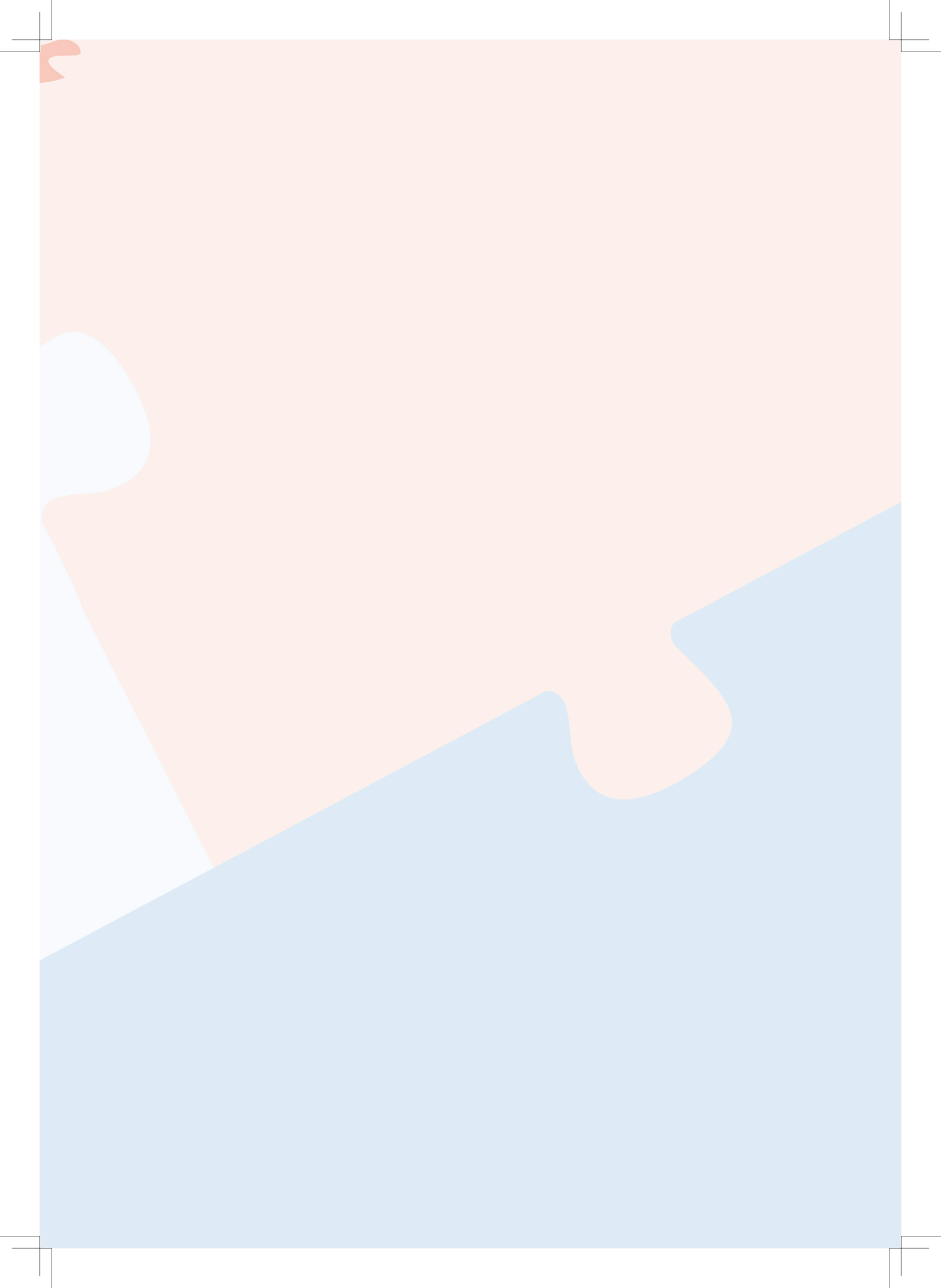
---

2017 Shanghai China 🇨🇳 December 2<sup>nd</sup> — 3<sup>rd</sup>

---











# Content

<hr/>		<hr/>	
Welcome Message	2	Breakout Session on Day 2	
Background	3	· Topic: Assessment Literacy for Language Teachers	35
Previous conferences	5	· Topic: Measuring Teachers' Language Teaching Proficiency	38
Conference theme	7	· Topic: Language Proficiency Frameworks	40
Conference agenda	9	· Topic: Foreign Language Testing for Employability	44
Plenary speakers	13	· Topic: Foreign Language Testing for Admissions and Exit Tests in Education Systems	47
Panel discussion speakers	17	· Topic: Artificial Intelligence and Technology in Language Testing	51
<hr/>		<hr/>	
Breakout Session on Day 1		Sponsors	55
· Topic: Language Proficiency Frameworks	21	Floor plan	57
· Topic: Assessment Literacy for Language Teachers	27		
· Topic: Artificial Intelligence and Technology in Language Testing	31		
· Topic: Foreign Language Testing for Admission and Exit Tests in Education Systems	33		
<hr/>		<hr/>	



# Welcome Message

Welcome to the 3rd International Conference on Language Testing and Assessment and the 5th British Council New Directions in English Language Assessment Conference 2017. We are delighted to welcome guests and speakers from China and across East Asia, the UK, the US and other countries to Shanghai for this momentous occasion.

Today marks the first time that our organisations have joined forces by holding a joint conference on language assessment, putting into practice the memorandum of understanding signed at the 9th China-UK Ministerial Meeting in 2016.

Working together in partnership, we are delighted through this conference, to offer all participants invaluable insights into current research in the field of assessment, as well as setting new directions in both testing and in best practice across East Asia.

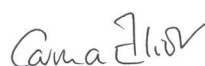
A broad range of stakeholders will benefit from participation - from education and testing practitioners, policy makers and implementers – all those with an interest in the impact of assessment on education from across the region. It provides a unique opportunity to engage our key stakeholders, experts and practitioners in assessment to congregate, network and share best practice, setting the tone for future education reform and internationalisation in assessment.

We have much to celebrate and looking ahead to the future, will continue to work together in collaboration to accelerate international exchange across teaching, learning and assessment and are exciting about what the future holds.

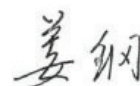
We extend our deepest gratitude to our partner, Shanghai Jiao Tong University for their continued incredible support and advice.

We are very grateful to all of our sponsors, including New Oriental Education & Technology Group, iFLYTEK Co., Ltd, Shanghai Liulishuo Information Technology Co., Ltd., Shenzhen SEASKYLAND Technologies Co., Ltd. and Higher Education Press for their contributions and sustained support.

Finally, we hope that you enjoy and are inspired by the conference, and are able to take home with you valuable new insights into English language assessment globally, and in East Asia in particular, which will benefit your community, your organisation and, most importantly, your students.



**Carma Elliot**  
Director  
British Council China



**Gang JIANG**  
President  
National Education  
Examinations  
Authority China

# Background

**National Education Examinations Authority (NEEA) in China and British Council are proud to announce a joint conference on language assessment:**

## The 3rd International Conference on Language Testing and Assessment and the 5th British Council New Directions in English Language Assessment Conference in Shanghai, 2017

### **NEEA**

The International Conference on Language Testing and Assessment (ICLTA) is an annual academic conference organized by NEEA in China, aiming to provide a platform for Chinese researchers, exams organizations and policy makers in the field of foreign language testing and teaching to exchange ideas and views on foreign language teaching, learning and testing. This year, Shanghai Jiao Tong University will be a supporting partner of this conference.

### **British Council**

New Directions in English Language Assessment conference is an annual event on the East Asian language assessment calendar, hosted by the British Council. The conference provides insight on trends in English language assessment and is a showcase for innovative approaches and research in the field, locally and internationally. New Directions is intended as space in which regional and international policy makers, educational professionals, academics and teachers and assessment practitioners can interact, exchange information, and keep abreast of the leading developments in the field.

## **Theme of Joint Conference: Connecting Assessment with Teaching and Learning**

This conference focuses on the critical role that language assessment plays in education. We will look at how descriptions of proficiency, teaching ability and teachers' assessment knowledge inform teaching and assessment. We will consider how language testing interacts with teaching and learning at the beginning and end of educational cycles and beyond as students are equipped for the working world. As the emphasis on communicative competence in East Asian education systems continues, new technology solutions to large scale testing demands will also be explored. The conference aims to provide a forum for the presentation of innovative approaches and practical solutions, with a focus on real-life case studies and works-in-progress reports.

## Previous conferences



### **1st International Conference on Language Testing and Assessment**

#### Study on Foreign Language Proficiency Assessment System

Dates: 28-29 November 2015  
Venue: Guangdong University of Foreign Studies, Guangzhou, China

The first conference focused on the study of foreign language proficiency assessment systems, including foreign language proficiency scales, foreign language test reform in China, formative assessment in foreign language teaching and standards-based foreign language teaching and evaluation. The conference attracted more than 250 scholars from domestic and overseas universities and research institutions.



### **2nd International Conference on Language Testing and Assessment**

#### Innovation and Development in Foreign Language Testing

Dates: 12-13 November 2016  
Venue: Zhejiang University, Hangzhou, China

The conference in 2016 looked at new developments in foreign language testing, including frameworks for language testing and assessment, language proficiency scales, formative assessment in foreign language teaching, application of artificial intelligence in language testing, and item banking in language testing. The 2016 conference also introduced the development of China's Standards of English. The conference attracted more than 250 scholars from universities and research institutions, both in China and abroad.



## 1st British Council New Directions in English Language Assessment

### Language Assessment in China

Dates: 6-7 November 2013

Venue: National Museum of China, Beijing, China

The first New Directions conference addressed issues associated with assessment for learning, washback and formative assessment, localisation and customisation, adapting tests to Chinese contexts and developing English teacher assessment literacy. The event in Beijing coincided with a number of anniversaries: 35 years since the British Council was established in China; 25 years of IELTS worldwide and 10 years of the British Council delivering IELTS in China in partnership with the National Education Examinations Authority.



## 2nd British Council New Directions in English Language Assessment

### Role of English Assessment in Internationalisation

Dates: 29-30 September 2014

Venue: Meiji Kinenkan, Tokyo, Japan

The second conference in Japan brought together chief educational officers and key officials from across Asia to review issues of localisation and globalisation in English language assessment and of expanding capacity to address current and future education and assessment challenges. It also provided a valuable networking opportunity and set the stage for further cooperation across East Asia and beyond.



## 3rd British Council New Directions in English Language Assessment

### Quality and Consequence

Dates: 15-16 October 2015

Venue: JW Marriott Dongdaemun Square, Seoul, Korea

New Directions 2015 in Korea looked at test quality and the consequences of testing from the point of view of society and education policy. It also considered the impact of testing on pedagogy and curriculum design, examining the challenges in the testing and assessment sector and solutions to be found if countries such as Korea are to address the concerns amongst the wider public about the integrity and quality of testing methodologies.



## 4th British Council New Directions in English Language Assessment

### Standardised Testing and the Development and Use of Proficiency Scales: Addressing Local Challenges, Meeting Global Standards

Dates: 13-14 October 2016

Venue: Sheraton Hanoi Hotel, Hanoi, Vietnam

The focus in Hanoi 2016 was to look at the role and effects of standardised testing and the development and/or use of proficiency scales to describe language ability. We also considered performance-based testing: its benefits and challenges. Previous New Directions have provided an important forum for the interaction of local and global perspectives, reflecting the wide range of challenges facing language education both within the region and internationally. The theme aims to highlight the approaches and solutions to these challenges being proposed in the region and internationally.

## Conference theme

### Language Proficiency Frameworks

This year's conference explores recent developments in the design and use of proficiency descriptor frameworks in East Asia. Papers are exploring the localised use, adaption or development of fit-for-purpose proficiency frameworks.

### Measuring Teachers' Language Teaching Proficiency

Language tests not only assess students' language competence but, more indirectly, the success of the teaching standard and methodology.

### Foreign Language Testing for Employability

A fundamental aim of education systems is to produce employable citizens. Stemming from the interaction between teaching and testing, how should assessment be contributing to sound decisions about candidate's employability?



## Foreign Language Testing for Admissions and Exit Tests in Education Systems

High-level education system goals and aspirations for the language competence of students in East Asia invites further investigation into the complex interaction between testing and teaching at the admissions stage, and teaching and testing at the exit stage.

## Assessment Literacy for Language Teachers

The positive or negative washback of assessment on teaching has generated much discussion and teachers' understanding of how tests work and what they are measuring is a key factor in ensuring that teaching to the construct, or ability, occurs, rather than teaching to the test.

## Artificial Intelligence and Technology in Language Testing

With the focus on education for employability, the ability to actively communicate in a foreign language is key. Yet for teaching of speaking skills, in particular, to be effective, assessment tools need to reflect this component of language use. Large scale testing of speaking has been an elusive goal for most education systems due to human resource requirements, however. Technology has long been put forward as a potential panacea to this challenge.

# Conference agenda

Location	Session	Time	Event/Speakers																	
Foyer	Reception	8:00-9:00	Registration																	
	Opening	9:00-9:30	Opening Speech																	
Grand Ballroom	Plenary 1	9:30-10:15	<b>Jianda LIU</b> , Professor and Dean of the National Key Research Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, China Validating the Levels of the China's Standards of English																	
Lower Lobby	Break	10:15-10:45	Tea break @ lower lobby Break&Learn@TeaRoom: Now Screening! Assessment animated																	
Grand Ballroom	Plenary 2	10:45-11:30	Plenary 2 <b>Barry O'Sullivan</b> , Head of Assessment Research and Development, British Council, United Kingdom Establishing Principles and Procedures for Linking Examinations to the China's Standards of English																	
	Plenary 3	11:30-12:15	Plenary 3 <b>James Purpura</b> , Professor of Linguistics and Education in the Applied Linguistics and TESOL Program at Teachers College, Columbia University, United States Reforusing the Language Assessment Paradigm: Assessing Meaning																	
Hotel Restaurant	Lunch	12:15-13:30	Lunch Break																	
		13:30-15:25	<b>Grandball Room</b> - Special Breakout session-CSE Symposium <b>1.Hua DENG</b> , Lecturer, Hunan Normal University, China Construct validation of the writing strategy competence of the China's Standards of English Co-presenter: <b>Jie DENG</b> , Professor, Hunan Normal University, China <b>Shen ZOU</b> , Professor, Shanghai International Studies University, China <b>2.Linlin CAO</b> , Phd Candidate, Linguistics & Applied Linguistics, Guangdong University of Foreign Studies, China Exploring critical features across levels of CSE reading subscale descriptors Co-presenter: <b>Yongqiang ZENG</b> , President, Guangdong Teachers College of Foreign Language and Arts, China <b>3.Yidong JIA</b> , Dean of English Department, School of Foreign Studies, Shandong University of Finance and Economics, China Developing self-Assessment strategy inventory for acquiring English organizational knowledge																	
Meeting Rooms	Break-out Sessions	12:15-13:30	25 mins																	
			<table border="1"> <thead> <tr> <th>Room 201 - Assessment Literacy for Language Teachers</th> <th>Room 205 - Assessment Literacy for Language Teachers</th> <th>Tearoom - Language Proficiency Frameworks</th> <th>Grandball Room - Special Breakout session-CSE Symposium</th> <th>Tearoom - Language Proficiency Frameworks</th> <th>Room 206+207 - Artificial Intelligence and Technology in Language Testing</th> <th>Room 208 - Foreign Language Testing for Admissions and Exit Tests in Education Systems</th> <th>Boardroom - Language Proficiency Frameworks</th> </tr> </thead> <tbody> <tr> <td><b>Vivien Berry</b>, Senior Researcher, English language Assessment, British Council, United Kingdom What does assessment literacy mean to language teachers? Co-presenter: <b>Susan Sheehan</b>, Senior Lecturer in TESOL, University of Huddersfield, United Kingdom</td> <td><b>Rita Green</b>, Visiting Teaching Fellow, Lancaster University, United Kingdom Establishing baselines for listening and speaking in Bangladesh: Using mixed-methods Co-presenter: <b>Tineke Brunfaut</b>, Senior Lecturer, Lancaster University, United Kingdom</td> <td><b>Jacqueline Martin</b>, Young Learners Project Lead, Pearson Education, United Kingdom Global Scale of English: ensuring global standards meet local requirements</td> <td><b>1.Hua DENG</b>, Lecturer, Hunan Normal University, China Construct validation of the writing strategy competence of the China's Standards of English Co-presenter: <b>Jie DENG</b>, Professor, Hunan Normal University, China <b>Shen ZOU</b>, Professor, Shanghai International Studies University, China <b>2.Linlin CAO</b>, Phd Candidate, Linguistics &amp; Applied Linguistics, Guangdong University of Foreign Studies, China Exploring critical features across levels of CSE reading subscale descriptors Co-presenter: <b>Yongqiang ZENG</b>, President, Guangdong Teachers College of Foreign Language and Arts, China <b>3.Yidong JIA</b>, Dean of English Department, School of Foreign Studies, Shandong University of Finance and Economics, China Developing self-Assessment strategy inventory for acquiring English organizational knowledge</td> <td><b>Spiros Papageorgiou</b>, Managing Senior Research Scientist, Educational Testing Service, United States Issues in aligning assessments to language proficiency frameworks and standards</td> <td><b>Rooms 206+207 - Artificial Intelligence and Technology in Language Testing</b> <b>Pucheng WANG</b>, PhD Candidate, Department of Linguistics and English Language, Lancaster University, United Kingdom Investigating test-takers' cognitive processes while completing a reading-to-write task Co-presenters: <b>Luke Harding</b>, Senior Lecturer, Department of Linguistics and English Language, Lancaster University, United Kingdom <b>LIU Baoquan</b>, Associate Professor, Department of English, Shanghai University of Finance and Economics, China <b>Shuxin ZHANG</b>, University of Chinese Academy of Sciences, China Generic features and coherence of prepared speeches in automated evaluation Co-presenters: <b>Daniel Lam</b>, Postdoctoral research fellow, University of Bedfordshire, United Kingdom <b>Jun Shi</b>, R&amp;D Manager of English Teaching and Assessment, Beijing Shangruitong Educational Technology Co., Ltd, China"</td> <td><b>Room 208 - Foreign Language Testing for Admissions and Exit Tests in Education Systems</b> <b>NI CHEN</b>, Associate Researcher and Vice Dean, School of Education and Training, Renmin University of China, China The effectiveness of admission criteria in predicting collaborative 2+2 program success of Chinese students Co-presenter: <b>Michael Burns</b>, Professor, Memorial University, Canada"</td> <td><b>Boardroom - Language Proficiency Frameworks</b> <b>Heather Koons</b>, Director, Research Services, MetaMetrics, United States Examining graded readers using CEFR levels and text complexity levels Co-presenter: <b>Eleanor Sanford-Moore</b>, Vice President for Research and Development, MetaMetrics, United States</td> <td><b>Elizabeth Shepherd</b>, Senior Researcher, British Council, United Kingdom Implementing an international English language evaluation; learning from PISA, using technology Co-presenters: <b>Martin Murphy</b>, Head of the Sampling Unit, Australian Council for Educational Research, Australia</td> </tr> </tbody> </table>	Room 201 - Assessment Literacy for Language Teachers	Room 205 - Assessment Literacy for Language Teachers	Tearoom - Language Proficiency Frameworks	Grandball Room - Special Breakout session-CSE Symposium	Tearoom - Language Proficiency Frameworks	Room 206+207 - Artificial Intelligence and Technology in Language Testing	Room 208 - Foreign Language Testing for Admissions and Exit Tests in Education Systems	Boardroom - Language Proficiency Frameworks	<b>Vivien Berry</b> , Senior Researcher, English language Assessment, British Council, United Kingdom What does assessment literacy mean to language teachers? Co-presenter: <b>Susan Sheehan</b> , Senior Lecturer in TESOL, University of Huddersfield, United Kingdom	<b>Rita Green</b> , Visiting Teaching Fellow, Lancaster University, United Kingdom Establishing baselines for listening and speaking in Bangladesh: Using mixed-methods Co-presenter: <b>Tineke Brunfaut</b> , Senior Lecturer, Lancaster University, United Kingdom	<b>Jacqueline Martin</b> , Young Learners Project Lead, Pearson Education, United Kingdom Global Scale of English: ensuring global standards meet local requirements	<b>1.Hua DENG</b> , Lecturer, Hunan Normal University, China Construct validation of the writing strategy competence of the China's Standards of English Co-presenter: <b>Jie DENG</b> , Professor, Hunan Normal University, China <b>Shen ZOU</b> , Professor, Shanghai International Studies University, China <b>2.Linlin CAO</b> , Phd Candidate, Linguistics & Applied Linguistics, Guangdong University of Foreign Studies, China Exploring critical features across levels of CSE reading subscale descriptors Co-presenter: <b>Yongqiang ZENG</b> , President, Guangdong Teachers College of Foreign Language and Arts, China <b>3.Yidong JIA</b> , Dean of English Department, School of Foreign Studies, Shandong University of Finance and Economics, China Developing self-Assessment strategy inventory for acquiring English organizational knowledge	<b>Spiros Papageorgiou</b> , Managing Senior Research Scientist, Educational Testing Service, United States Issues in aligning assessments to language proficiency frameworks and standards	<b>Rooms 206+207 - Artificial Intelligence and Technology in Language Testing</b> <b>Pucheng WANG</b> , PhD Candidate, Department of Linguistics and English Language, Lancaster University, United Kingdom Investigating test-takers' cognitive processes while completing a reading-to-write task Co-presenters: <b>Luke Harding</b> , Senior Lecturer, Department of Linguistics and English Language, Lancaster University, United Kingdom <b>LIU Baoquan</b> , Associate Professor, Department of English, Shanghai University of Finance and Economics, China <b>Shuxin ZHANG</b> , University of Chinese Academy of Sciences, China Generic features and coherence of prepared speeches in automated evaluation Co-presenters: <b>Daniel Lam</b> , Postdoctoral research fellow, University of Bedfordshire, United Kingdom <b>Jun Shi</b> , R&D Manager of English Teaching and Assessment, Beijing Shangruitong Educational Technology Co., Ltd, China"	<b>Room 208 - Foreign Language Testing for Admissions and Exit Tests in Education Systems</b> <b>NI CHEN</b> , Associate Researcher and Vice Dean, School of Education and Training, Renmin University of China, China The effectiveness of admission criteria in predicting collaborative 2+2 program success of Chinese students Co-presenter: <b>Michael Burns</b> , Professor, Memorial University, Canada"	<b>Boardroom - Language Proficiency Frameworks</b> <b>Heather Koons</b> , Director, Research Services, MetaMetrics, United States Examining graded readers using CEFR levels and text complexity levels Co-presenter: <b>Eleanor Sanford-Moore</b> , Vice President for Research and Development, MetaMetrics, United States	<b>Elizabeth Shepherd</b> , Senior Researcher, British Council, United Kingdom Implementing an international English language evaluation; learning from PISA, using technology Co-presenters: <b>Martin Murphy</b> , Head of the Sampling Unit, Australian Council for Educational Research, Australia
Room 201 - Assessment Literacy for Language Teachers	Room 205 - Assessment Literacy for Language Teachers	Tearoom - Language Proficiency Frameworks	Grandball Room - Special Breakout session-CSE Symposium	Tearoom - Language Proficiency Frameworks	Room 206+207 - Artificial Intelligence and Technology in Language Testing	Room 208 - Foreign Language Testing for Admissions and Exit Tests in Education Systems	Boardroom - Language Proficiency Frameworks													
<b>Vivien Berry</b> , Senior Researcher, English language Assessment, British Council, United Kingdom What does assessment literacy mean to language teachers? Co-presenter: <b>Susan Sheehan</b> , Senior Lecturer in TESOL, University of Huddersfield, United Kingdom	<b>Rita Green</b> , Visiting Teaching Fellow, Lancaster University, United Kingdom Establishing baselines for listening and speaking in Bangladesh: Using mixed-methods Co-presenter: <b>Tineke Brunfaut</b> , Senior Lecturer, Lancaster University, United Kingdom	<b>Jacqueline Martin</b> , Young Learners Project Lead, Pearson Education, United Kingdom Global Scale of English: ensuring global standards meet local requirements	<b>1.Hua DENG</b> , Lecturer, Hunan Normal University, China Construct validation of the writing strategy competence of the China's Standards of English Co-presenter: <b>Jie DENG</b> , Professor, Hunan Normal University, China <b>Shen ZOU</b> , Professor, Shanghai International Studies University, China <b>2.Linlin CAO</b> , Phd Candidate, Linguistics & Applied Linguistics, Guangdong University of Foreign Studies, China Exploring critical features across levels of CSE reading subscale descriptors Co-presenter: <b>Yongqiang ZENG</b> , President, Guangdong Teachers College of Foreign Language and Arts, China <b>3.Yidong JIA</b> , Dean of English Department, School of Foreign Studies, Shandong University of Finance and Economics, China Developing self-Assessment strategy inventory for acquiring English organizational knowledge	<b>Spiros Papageorgiou</b> , Managing Senior Research Scientist, Educational Testing Service, United States Issues in aligning assessments to language proficiency frameworks and standards	<b>Rooms 206+207 - Artificial Intelligence and Technology in Language Testing</b> <b>Pucheng WANG</b> , PhD Candidate, Department of Linguistics and English Language, Lancaster University, United Kingdom Investigating test-takers' cognitive processes while completing a reading-to-write task Co-presenters: <b>Luke Harding</b> , Senior Lecturer, Department of Linguistics and English Language, Lancaster University, United Kingdom <b>LIU Baoquan</b> , Associate Professor, Department of English, Shanghai University of Finance and Economics, China <b>Shuxin ZHANG</b> , University of Chinese Academy of Sciences, China Generic features and coherence of prepared speeches in automated evaluation Co-presenters: <b>Daniel Lam</b> , Postdoctoral research fellow, University of Bedfordshire, United Kingdom <b>Jun Shi</b> , R&D Manager of English Teaching and Assessment, Beijing Shangruitong Educational Technology Co., Ltd, China"	<b>Room 208 - Foreign Language Testing for Admissions and Exit Tests in Education Systems</b> <b>NI CHEN</b> , Associate Researcher and Vice Dean, School of Education and Training, Renmin University of China, China The effectiveness of admission criteria in predicting collaborative 2+2 program success of Chinese students Co-presenter: <b>Michael Burns</b> , Professor, Memorial University, Canada"	<b>Boardroom - Language Proficiency Frameworks</b> <b>Heather Koons</b> , Director, Research Services, MetaMetrics, United States Examining graded readers using CEFR levels and text complexity levels Co-presenter: <b>Eleanor Sanford-Moore</b> , Vice President for Research and Development, MetaMetrics, United States	<b>Elizabeth Shepherd</b> , Senior Researcher, British Council, United Kingdom Implementing an international English language evaluation; learning from PISA, using technology Co-presenters: <b>Martin Murphy</b> , Head of the Sampling Unit, Australian Council for Educational Research, Australia												

				<p>Co-presenter: <b>Zumin WU</b>, Professor, Beijing Normal University, China</p> <p><b>4.Weiyong HUANG</b>, Lecturer, East China University of Technology, China</p> <p>Constructing an interlanguage pragmatic scale for Chinese tertiary EFL learners</p> <p>Co-presenter: <b>Baocheng HAN</b>, Professor, Beijing Foreign Studies University, China</p> <p><b>5.Li FENG</b>, Associate Professor, Institute of Applied Foreign Languages, Helongjiang University, China</p> <p>The development of descriptors of translation proficiency under the Project of CSE: Problems and solutions</p> <p>Co-presenter: <b>Ming YAN</b>, Professor, Helongjiang University, China</p> <p><b>6. Yi XU</b>   Lecturer, School of Interpreting and Translation Studies, Guangdong University of Foreign Studies, China</p>	<p><b>Nguyen Thi Mai Huu</b>, Executive Deputy Director, National Foreign Language Project 2020, Vice Director of the Center for Language Testing and Assessment, Vietnam</p> <p>Developing the Vietnamese standardized test of English proficiency</p> <p>Co-presenter: <b>Tue Hoang</b>, Dean of Faculty, Foreign Languages, Hanoi University of Industry, Vietnam</p>	<p><b>Duyen Tran</b>, Academic Project Manager, Hanoi University of Industry, Vietnam</p> <p>Teachers' perspectives and practices of EFL assessment at Vietnamese universities</p> <p>Co-presenter: <b>Lawrence Zhang</b>, Professor of Linguistics-in-Education and Associate Dean (AD), International Partnerships, University of Auckland, New Zealand</p> <p><b>Helen Ramsey Dixon</b>, Associate Professor, University of Auckland, New Zealand</p>	<p><b>Qiaozhen YAN</b>, PhD Student, Educational Linguistics, University of Auckland, New Zealand</p> <p>Primary English language teachers' conceptions and practices in classroom assessment</p> <p>Co-presenter: <b>Liyan HUANG</b>, Director of English Education and Assessment Research Center, South China Normal University, China</p> <p>Constructing and validating a conceptual framework of assessment literacy for Chinese EFL teachers</p> <p>Co-presenter: <b>Ming XU</b>, MA Candidate, South China Normal University, China*</p>	<p><b>Dajian CHEN</b>, Doctoral Candidate, School of International Studies, Zhejiang University, China</p> <p>Investigating video-based listening test: the role of visual information in comprehension process</p>	<p><b>Richard Spiby</b>, Test Development Researcher, Receptive Skills, British Council, United Kingdom</p> <p>A Framework for innovation: revising the Aptis listening test</p> <p>Co-presenter: <b>Franz Holzknrecht</b>, Lecturer and Researcher, University of Innsbruck, Austria</p>	<p><b>Li LIANG</b>, Lecturer, Guangdong University of Foreign Studies, China</p> <p>A socio-cultural perspective: EFL teachers' perception in the context of computer-based large-scale oral English test</p>	<p><b>Wenxing ZHANG</b>, PhD in Language Assessment and Evaluation, Jiangxi Science &amp; Technology Normal University, China</p> <p>Exploring the adaptability of the CEFR in developing the TEM writing ability scale</p> <p>Co-presenter: <b>Shen ZOU</b>, Professor in the School of English Studies, Shanghai International Studies University, China*</p>	
	12:15-13:30	25 mins	<p>Co-presenter: <b>Zumin WU</b>, Professor, Beijing Normal University, China</p> <p><b>4.Weiyong HUANG</b>, Lecturer, East China University of Technology, China</p> <p>Constructing an interlanguage pragmatic scale for Chinese tertiary EFL learners</p> <p>Co-presenter: <b>Baocheng HAN</b>, Professor, Beijing Foreign Studies University, China</p> <p><b>5.Li FENG</b>, Associate Professor, Institute of Applied Foreign Languages, Helongjiang University, China</p> <p>The development of descriptors of translation proficiency under the Project of CSE: Problems and solutions</p> <p>Co-presenter: <b>Ming YAN</b>, Professor, Helongjiang University, China</p> <p><b>6. Yi XU</b>   Lecturer, School of Interpreting and Translation Studies, Guangdong University of Foreign Studies, China</p>	<p><b>Nguyen Thi Mai Huu</b>, Executive Deputy Director, National Foreign Language Project 2020, Vice Director of the Center for Language Testing and Assessment, Vietnam</p> <p>Developing the Vietnamese standardized test of English proficiency</p> <p>Co-presenter: <b>Tue Hoang</b>, Dean of Faculty, Foreign Languages, Hanoi University of Industry, Vietnam</p>	<p><b>Duyen Tran</b>, Academic Project Manager, Hanoi University of Industry, Vietnam</p> <p>Teachers' perspectives and practices of EFL assessment at Vietnamese universities</p> <p>Co-presenter: <b>Lawrence Zhang</b>, Professor of Linguistics-in-Education and Associate Dean (AD), International Partnerships, University of Auckland, New Zealand</p> <p><b>Helen Ramsey Dixon</b>, Associate Professor, University of Auckland, New Zealand</p>	<p><b>Qiaozhen YAN</b>, PhD Student, Educational Linguistics, University of Auckland, New Zealand</p> <p>Primary English language teachers' conceptions and practices in classroom assessment</p> <p>Co-presenter: <b>Liyan HUANG</b>, Director of English Education and Assessment Research Center, South China Normal University, China</p> <p>Constructing and validating a conceptual framework of assessment literacy for Chinese EFL teachers</p> <p>Co-presenter: <b>Ming XU</b>, MA Candidate, South China Normal University, China*</p>	<p><b>Dajian CHEN</b>, Doctoral Candidate, School of International Studies, Zhejiang University, China</p> <p>Investigating video-based listening test: the role of visual information in comprehension process</p>	<p><b>Richard Spiby</b>, Test Development Researcher, Receptive Skills, British Council, United Kingdom</p> <p>A Framework for innovation: revising the Aptis listening test</p> <p>Co-presenter: <b>Franz Holzknrecht</b>, Lecturer and Researcher, University of Innsbruck, Austria</p>	<p><b>Li LIANG</b>, Lecturer, Guangdong University of Foreign Studies, China</p> <p>A socio-cultural perspective: EFL teachers' perception in the context of computer-based large-scale oral English test</p>	<p><b>Wenxing ZHANG</b>, PhD in Language Assessment and Evaluation, Jiangxi Science &amp; Technology Normal University, China</p> <p>Exploring the adaptability of the CEFR in developing the TEM writing ability scale</p> <p>Co-presenter: <b>Shen ZOU</b>, Professor in the School of English Studies, Shanghai International Studies University, China*</p>	Tea break@lower lobby Break&Learn@Tea room: Now Screening! Assessment animated	
Lower Lobby	Break	35 mins	<p>Strategic competence in interpreting scale of CSE: A meta-cognitive perspective</p> <p>Co-presenter: <b>Lei Wu</b>, Professor, Guangdong University of Foreign Studies, China</p> <p><b>7.Mei JIE</b>, PhD Candidate, Shanghai Jiao Tong University, China</p> <p>Linking test to the China's Standard of English-Speaking scale: CET-SET4 as a case study</p> <p>Co-presenter: <b>Yan JIN</b>, Professor, Shanghai Jiao Tong University, China</p>	<p><b>Haiping WANG</b>, Postgraduate Supervisor at School of Foreign Studies, East China University of Political Science and Law, China</p> <p>Designing localized CEFR-referenced SCELL for district foreign language schools in Shanghai</p> <p>Co-presenter: <b>Limin ZOU</b>, Dean, Shanghai International Studies University</p> <p>Jing'an Foreign Language-High School, China</p>	<p><b>Quyynh Le</b>, Head of ELT Department, Faculty of Social Sciences and Humanities, Hoa Sen University, Vietnam</p> <p>Apts and teachers' conceptions of assessment and teaching practice</p> <p>Co-presenter: <b>Trang Lam</b>, Program coordinator ELT Department, Hoa Sen University, Vietnam</p>	<p><b>Liyan HUANG</b>, Director of English Education and Assessment Research Center, South China Normal University, China</p> <p>Constructing and validating a conceptual framework of assessment literacy for Chinese EFL teachers</p> <p>Co-presenter: <b>Ming XU</b>, MA Candidate, South China Normal University, China*</p>	<p><b>Richard Spiby</b>, Test Development Researcher, Receptive Skills, British Council, United Kingdom</p> <p>A Framework for innovation: revising the Aptis listening test</p> <p>Co-presenter: <b>Franz Holzknrecht</b>, Lecturer and Researcher, University of Innsbruck, Austria</p>	<p><b>Li LIANG</b>, Lecturer, Guangdong University of Foreign Studies, China</p> <p>A socio-cultural perspective: EFL teachers' perception in the context of computer-based large-scale oral English test</p>	<p><b>Wenxing ZHANG</b>, PhD in Language Assessment and Evaluation, Jiangxi Science &amp; Technology Normal University, China</p> <p>Exploring the adaptability of the CEFR in developing the TEM writing ability scale</p> <p>Co-presenter: <b>Shen ZOU</b>, Professor in the School of English Studies, Shanghai International Studies University, China*</p>			
Grand Ballroom	Panel Discussion	90 mins	<p>Panel Discussion 1: Perspectives on Language Frameworks</p> <p>Panelists: <b>Neus Figueras</b>, Senior Researcher, Regional Ministry of Education in Catalonia, Spain <b>Masashi Negishi</b>, Professor, Tokyo University of Foreign Studies, Japan <b>Lianzhen HE</b>, Professor of School of International Studies, Zhejiang University, China <b>Wenxia ZHANG</b>, Professor of Applied Linguistics in the Department of Foreign Languages and Literatures, Tsinghua University, China <b>Sha Wu</b>, Director of Foreign Language Assessment Department, National Education Examinations Authority, China</p> <p>Chair: <b>Nick Saville</b>, Director of Research &amp; Thought Leadership, Cambridge English, United Kingdom</p>									
	Closing	5 mins	Review of the Day									

# Day 2, Sunday, 3 December 2017-Hyatt on the Bund

Yan JIN, Professor of Linguistics and Applied Linguistics at the School of Foreign Languages, Shanghai Jiao Tong University, China  
The Interface between Language Teaching and Testing: The Role of Local Standards

Luke Harding, Senior Lecturer, Department of Linguistics and English Language, Lancaster University, United Kingdom  
Language Assessment Literacy: Past, present, future

Philip Horne, Assessment Consultant, British Council, China  
A validity case for procedural assessment literacy for EFL teachers in China

Robinson Skusey, Academic Manager, English for Education Systems Japan, British Council, Japan  
Teaching how to rate the productive skills: Case study from Japan  
Co-presenter:  
Judith Fairbairn, Productive Skills Testing Researcher, British Council, United Kingdom

Michelle Raquel, Senior Lecturer, Centre of Applied English Studies, University of Hong Kong, Hong Kong  
Assessment literacy profile of tertiary level course coordinators  
Co-presenters:  
Simon Boynton, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong  
Bernadette Wo, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong

Michelle Raquel, Senior Lecturer, Centre of Applied English Studies, University of Hong Kong, Hong Kong  
Assessment literacy profile of tertiary level course coordinators  
Co-presenters:  
Simon Boynton, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong  
Bernadette Wo, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong

Michelle Raquel, Senior Lecturer, Centre of Applied English Studies, University of Hong Kong, Hong Kong  
Assessment literacy profile of tertiary level course coordinators  
Co-presenters:  
Simon Boynton, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong  
Bernadette Wo, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong

Michelle Raquel, Senior Lecturer, Centre of Applied English Studies, University of Hong Kong, Hong Kong  
Assessment literacy profile of tertiary level course coordinators  
Co-presenters:  
Simon Boynton, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong  
Bernadette Wo, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong

Michelle Raquel, Senior Lecturer, Centre of Applied English Studies, University of Hong Kong, Hong Kong  
Assessment literacy profile of tertiary level course coordinators  
Co-presenters:  
Simon Boynton, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong  
Bernadette Wo, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong

Michelle Raquel, Senior Lecturer, Centre of Applied English Studies, University of Hong Kong, Hong Kong  
Assessment literacy profile of tertiary level course coordinators  
Co-presenters:  
Simon Boynton, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong  
Bernadette Wo, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong

Michelle Raquel, Senior Lecturer, Centre of Applied English Studies, University of Hong Kong, Hong Kong  
Assessment literacy profile of tertiary level course coordinators  
Co-presenters:  
Simon Boynton, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong  
Bernadette Wo, Lecturer, Centre for Applied English Studies, University of Hong Kong, Hong Kong

Location	Session	Time	Event/Speakers
Foyer	Reception	8:00-8:50	Registration
	Opening	8:50-9:00	First Day Recap
Grand Ballroom	Plenary 4	9:00-9:45	Plenary 4 Yan JIN, Professor of Linguistics and Applied Linguistics at the School of Foreign Languages, Shanghai Jiao Tong University, China The Interface between Language Teaching and Testing: The Role of Local Standards
	Plenary 5	9:45-10:30	Plenary 5 Luke Harding, Senior Lecturer, Department of Linguistics and English Language, Lancaster University, United Kingdom Language Assessment Literacy: Past, present, future
Lower Lobby	Break	10:30-11:00	Tea break @lower lobby Break&Learn@Tearoom: Now Screening! Assessment animated
Meeting Rooms	Break-out Sessions	11:00-11:55	<b>Room 201 - Assessment Literacy for Language Teachers</b> Joseph Davies, Business Development Lead, University of Central Lancashire's School of Language and Global Studies, China Grading class participation: Perception and practice of ELTs in China Co-presenter: Laura Davies, Programme Leader, University of Central Lancashire's franchised BA(Hons) International Business Communication Programme, China
		11:00-11:25	25 mins
Meeting Rooms	Break-out Sessions	11:30-11:55	<b>Room 205 - Foreign Language Testing for Employability</b> Johanna Motteram, Teacher, Assessment Specialist, British Council, Singapore Employment in the UK, IELTS writing, and Philippines trained nurses
		25 mins	
Meeting Rooms	Break-out Sessions	11:30-11:55	<b>Room 206+207 - Artificial Intelligence and Technology in Language Testing</b> Hui LIN, Co-Founder and Chief Scientist, Lulishuo, China Large scale adaptive language testing via artificial intelligence
		25 mins	
Meeting Rooms	Break-out Sessions	11:30-11:55	<b>Room 208 - Foreign Language Testing for Admissions and Exit Tests in Education Systems</b> Jason West, Head of English, SHU UTS SILC Business School, China Connecting the Dots with Apts testing insights
		25 mins	
Hotel Restaurant	Lunch	12:00-13:30	Lunch Break
		13:30-14:55	<b>Room 201 - Measuring Teachers' Language teaching Proficiency</b> Le Duy Hai Nguyen, Lecturer, Binh Duong University, Vietnam Measuring teachers' English proficiency: Reviewing a project in Vietnam
Hotel Restaurant	Lunch	13:30-14:55	<b>Room 205 - Foreign Language Testing for Employability</b> Mike Mayor, Director, Global Scale of English, Pearson English, United Kingdom Creating a proficiency framework for the workplace
		25 mins	
Hotel Restaurant	Lunch	13:30-14:55	<b>Room 206+207 - Artificial Intelligence and Technology in Language Testing</b> Xiangdong GU, Professor, PhD Supervisor and Director, Chongqing University, China Analysing and predicting English majors' TEM-4 performance in China Co-presenters: Yu ZHONG, PhD Candidate, Chongqing University, China
		25 mins	
Hotel Restaurant	Lunch	13:30-14:55	<b>Room 208 - Foreign Language Testing for Admissions and Exit Tests in Education Systems</b> Fengping GUO, Lecturer, Southeast University, China An investigation of the use of Apts for EFL placement decisions in tertiary education Co-presenter: Maoxia YANG, Vice Director of the Testing Centre, Southeast University, China
		25 mins	

	Meeting Rooms	14:00-14:25	25 mins	<p><b>Xin QI</b>, Professor of Applied Linguistics, Beijing Foreign Studies University, China</p> <p>Validation of ELT certificate-LORT: A process suited for many-facet Rasch modeling</p> <p>Co-presenter: <b>Baocheng HAN</b>, Deputy Director of the National Research Centre for Foreign Language Education, Beijing Foreign Studies University, China</p>	<p><b>Mohd Sallehudin Abd Aziz</b>, Associate Professor, School of Language Studies and Linguistics, Universiti Kebangsaan Malaysia, Malaysia</p> <p>CEFR in Malaysia: Current issues and challenges in the implementation of the framework</p> <p>Co-presenter: <b>Nurul Farehah Mohamad Uri</b>, Lecturer, Universiti Kuala Lumpur British Malaysian Institute, Malaysia</p>	<p><b>Nguyen Thi Ngoc Quynh</b>, Director of the Center for Language Testing and Assessment, University of Languages and International Studies, Vietnam National University, Vietnam</p> <p>Test quality improvement guidance for VSTEP test writers</p> <p>Co-presenter: <b>Nguyen Thi Quynh Yen</b>, <b>Nguyen Thi Phuong Thao</b>, <b>Thai Ha Lam Thuy Bui Thien Sao</b>, Lecturers, Center for Language Testing and Assessment, University of Languages and International Studies, Vietnam National University, Hanoi, Vietnam</p> <p><b>Nathan T. Carr</b>, Professor, California State University, Fullerton, United States</p>	<p><b>Alan Urmston</b>, Assistant Professor and Assessment Coordinator, English Language Centre, Hong Kong Polytechnic University, Hong Kong</p> <p>Investigating the writing ability of graduates and employees of an Asian university</p> <p>Co-presenter: <b>Michelle Raquel</b>, Senior Lecturer, Centre of Applied English Studies, University of Hong Kong, Hong Kong</p> <p><b>Felicia FANG</b>, Language Instructor, The Hong Kong Polytechnic University, Hong Kong</p>	<p>Co-presenter: <b>Mingwei PAN</b>, Associate Professor, Applied Linguistics, Guangdong University of Foreign Studies, China</p>	<p><b>Yao MENG</b>, Data Scientist, Alibaba Corporation, China</p> <p><b>Qing ZHOU</b>, Professor, College of Computer Science, Chongqing University, China</p>	<p><b>Jing CHEN</b>, Associate Professor, School of Foreign Languages, Sun Yat-sen University, China</p> <p>Re-fitting text complexity for testing reading: Introducing a data-driven approach</p> <p>Co-presenters: <b>Pengpeng FENG</b>, Associate Professor, School of Foreign Languages, Sun Yat-sen University, China</p> <p><b>Kai GUO</b>, MA Student, School of Foreign Languages, Shanghai Jiao Tong University, China</p> <p><b>Tan JIN</b>, Associate Professor, School of Foreign Languages, Sun Yat-sen University, China</p>
	Break-out Sessions	14:30-14:55	25 mins	<p><b>Sayyed Mohammad Alavi</b>, Professor of Applied Linguistics, University of Tehran, Iran</p> <p>Language teachers' listening comprehension diagnostic competence in TBLT</p> <p>Co-presenter: <b>Tina Mirmotahari</b>, Teacher Trainer, University of Tehran, Iran</p>	<p><b>Wen ZHAO</b>, Professor of English, Northeastern University, China</p> <p>Calibrating the CEFR against the China's Standards of English for college English vocabulary education in China</p> <p>Co-presenters: <b>Boran WANG</b>, Professor of English, Northeastern University, China</p> <p><b>Xin Li</b>, Associate Professor of English, Northeastern University, China</p>	<p><b>Daniel Bates</b>, English Lecturer, Visiting Faculty Member, Asia University, Japan</p> <p>The use of language proficiency descriptors for pronunciation assessment</p>	<p><b>Tomoyasu Akiyama</b>, Professor, Bunkyo University, Japan</p> <p>Rater cognition: Impacts of rater differences on teacher employment examinations</p>	<p><b>Shangchao MIN</b>, Associate Professor, School of International Studies, Zhejiang University, China</p> <p>Retrofitting an in-house English listening test for diagnostic purposes</p> <p>Co-presenter: <b>Lidi XIONG</b>, Postgraduate, School of International Studies, Zhejiang University, China</p>	<p><b>Yong XU</b>, Associate Professor, English Language and Literature, Shanghai International Studies University, China</p> <p>Algorithm to predict item difficulty of MCQ grammar assessment</p> <p>Co-presenter: <b>Shanshan SHANG</b>, Associate Professor in School of Business and Management, Shanghai International Studies University, China</p>	<p><b>Jamie Dunlea</b>, Senior Researcher for the Language Assessment Research Group, British Council, United Kingdom</p> <p>A closer look at two tests aligned with the CEFR</p> <p>Co-presenters: <b>Richard Spiby</b>, Test Development Researcher, Receptive Skills, British Council, United Kingdom</p> <p><b>Quynh Nguyen Thi Ngoc</b>, Director of the Center for Language Testing and Assessment, University of Languages and International Studies, Vietnam National University, Vietnam</p> <p><b>Nguyen Thi Quynh Yen</b>, Deputy Director of the Center for Language Testing and Assessment, University of Languages and International Studies, Vietnam National University, Vietnam</p>
Lower lobby	Break	15:00-15:30	30 mins	Tea break@lower lobby Break&Learn@Tearoom: Now Screening! Assessment animated						
Grand Ballroom	Panel Discussion	15:30-16:40	70 mins	Panel Discussion 2: Artificial Intelligence in Language Assessment						
	Closing	16:40-17:00	20 mins	<p>Panel: <b>Helen Yannakoudakis</b>, Senior Research Associate at the Department of Computer Science and Technology, University of Cambridge, United Kingdom</p> <p><b>Alistair Van Moere</b>, Chief Product Officer, MetaMetrics, United States</p> <p><b>SI WEI</b>, Vice Director of FLYTEK Research, FLYTEK, China</p> <p><b>Hui LIN</b>, Co-Founder and Chief Scientist, Lulishuo, China</p> <p>Chair: <b>Barry O'Sullivan</b>, Head of Assessment Research and Development, British Council, United Kingdom</p> <p>Closing Ceremony</p>						

## Plenary Speakers



### Prof Jianda LIU

- Dean of the National Key Research Center for Linguistics and Applied Linguistics
- Guangdong University of Foreign Studies
- China

Jianda LIU is a professor and dean of the National Key Research Center for Linguistics and Applied Linguistics at Guangdong University of Foreign Studies. He is also currently vice-president of Guangdong University of Foreign Studies. He has worked in designing, developing and validating language tests in China. Professor Liu's research interests include language testing, pragmatics, and second language acquisition, and he has published many papers in a variety of academic journals. His book, *Measuring Interlanguage Pragmatic Knowledge of EFL Learners*, was published by Peter Lang. Professor Liu is now leading the development of China's Standards of English (CSE).

#### **Abstract:** Validating the Levels of China's Standards of English

The State Council of the People's Republic of China issued a document in 2014 entitled "The Implementation Opinions of the State Council on Deepening the Reform of the Examination and Enrollment System", designating the construction of a comprehensive assessment system of foreign language proficiency. Among the different tasks, developing China's Standards of English (CSE) is the fundamental task.

CSE describes what Chinese learners and users of English can do with the target language. It comprehensively defines English language proficiency levels in the Chinese EFL context, with specific, accurate and easy-to-understand descriptors in terms of communicative functions and language skills, serving as one of the fundamental strategies for the promotion of English learning, teaching and assessment in China.

This presentation first introduces briefly the CSE, followed by a detailed explanation of the procedures adopted and measures taken to ensure the reliability and validity of the descriptor levels. Quantitative and qualitative data was collected via questionnaires and interviews with language experts, teachers and learners (users). Various statistical procedures were conducted before the final levels were determined for the descriptors. Suggestions on the use of the CSE will be given at the end of this talk.



### Prof Barry O'Sullivan

- Head of Assessment Research and Development
- British Council
- United Kingdom

Professor Barry O'Sullivan is the Head of Assessment Research and Development at the British Council. He has researched factors affecting spoken performance; assessing rater behaviour; assessing speaking and writing; specific purpose assessment; benchmarking and standards; and has worked on the development and refinement of the socio-cognitive model of test development and validation since 2000. He has presented his work at over 150 conferences around the world, while over 90 of his publications have appeared in a range of international journals, and in many books and technical reports. In addition, he has published five books in the area of language testing. He is the founding president of the UK Association of Language Testing and Assessment and holds honorary and visiting chairs at the Universities of Reading and Roehampton (UK) and at the University of Lisbon (Portugal). In 2016 he was awarded fellowship of the Academy of Social Science in the UK.

#### **Abstract:** Establishing Principles and Procedures for Linking Examinations to the China Standards of English

To ensure the appropriate use of a test within an education system, it is necessary to generate evidence of a meaningful link between the test and the underlying standards upon which the system's curriculum is based. This is done in order to create a positive impact on the learning system by ensuring that there is a connection between its three key elements: the curriculum, the delivery of the curriculum and the assessments used within the system.

This year marks the launch of the China's Standards of English (CSE). The CSE, which is expected to have a significant impact on the English language education system in China in the coming years, is underpinned by a clearly defined socio-cognitive language model and as therefore fits extremely well with the socio-cognitive (SC) model for test development and validation (Weir, 2005; O'Sullivan, 2011, 2016) that has been used successfully in projects linking tests to the Common European Framework of Reference for Languages (CEFR) in the past (e.g. O'Sullivan, 2009, 2015).

In this paper I will present and discuss the principles and procedures adopted in a major project to establish an empirical link between two international tests (Aptis and IELTS) and the CSE. Having first established a socio-cognitive theoretical basis to the linking project, NEEA and the British Council worked together to set out detailed procedures to allow for a comprehensive and valid linking claim to be made for each of the tests. The design and preliminary stages of the project are discussed and reflected on in order to demonstrate the comprehensive nature of the work undertaken.





### Prof James Purpura

- Professor of linguistics and education in the Applied Linguistics and TESOL Program
- Teachers College, Columbia University
- United States

James E. Purpura is a Professor of linguistics and education in the Applied Linguistics and TESOL Program at Teachers College, Columbia University. Jim's research has focused on grammar assessment, meaning assessment, and the cognitive underpinnings of assessment. He is also pursuing research in learning-oriented assessment (LOA) as a means of understanding how learning can be incorporated into scenario-based assessments, and how L2 assessments function when embedded in teaching and learning. He is currently Co-editor of *Language Assessment Quarterly*, and series co-editor of *New Perspectives on Language Assessment* and *Innovations in Language and Learning Assessment* at ETS. He served as President of ILTA in 2007 and 2008. He was a Fulbright Scholar in Italy in 2017.

#### **Abstract:** Refocusing the Language Assessment Paradigm: Assessing meaning

Non-native speakers use second or foreign languages (L2) to get/give information at school, to create and maintain relationships online, to get a glimpse into other cultures, or more subtly, to decipher intentions in political discourse. In other words, they use their L2 to express a wide range of meanings within social-interpersonal contexts (e.g., a friend recounting a commuting story), social-transactional contexts (e.g., a client resolving a problem with a bill), academic contexts (e.g., a student writing a term paper), professional contexts (e.g., a scientist giving a talk), and literary or imaginative contexts (e.g., a poet writing/reciting a poem at a poetry slam). Since the ability to effectively express, understand, co-construct, negotiate, and repair meanings is the quintessential quality of communicative success, it stands to reason then that meaning and meaning conveyance should play more than a tangential role in L2 assessment (Purpura, 2004, 2016, 2017).

Notwithstanding, since the 1980s, language testers have focused almost exclusively on functional proficiency (the conveyance of functional meaning—e.g., can-do statements), to the exclusion of the conveyance of propositional meanings (i.e., content) or other implied pragmatic meanings. While the ability to use language to get things done is important, excluding propositional content from the assessment process is like having language ability with nothing to say, and excluding pragmatic meanings guts the heart and soul out of communication.

In this talk, I review how L2 testers have theoretically conceptualized "meaning" in models of L2 proficiency throughout the years, and argue for a reprioritization of meaningfulness over well formedness in L2 assessment. This logically leads to a discussion of the use of language to encode a range of meanings, deriving not only from an examinee's topical knowledge, but also from an understanding of the contextual factors of language use. Throughout the discussion, I highlight how the expression and comprehension of meaning have been operationalized in L2 assessments. Finally, I maintain that despite the complexities of defining and operationalizing meaning, testers need to seriously think about what meanings they want to test, and what meanings they are already assessing implicitly.



### Dr Yan JIN

- Professor of Linguistics and Applied Linguistics at the School of Foreign Languages
- Shanghai Jiao Tong University
- China

Dr. Yan JIN is a professor of Linguistics and Applied Linguistics at the School of Foreign Languages, Shanghai Jiao Tong University, China. She has been involved in the research and practice of language testing and assessment since the late 1980s and has rich experience in the operation of language assessment programs. Dr. Jin's research interest focuses on the development and validation of large-scale and high-stakes language assessments. She is currently chair of the National College English Testing Committee and Vice President of the Asian Association for Language Assessment. She is also co-editor-in-chief of the Springer open-access journal *Language Testing in Asia*. Dr. Jin also sits on the editorial boards of international and Chinese journals including *Language Testing*, *International Journal of Computer-Assisted Language Learning and Teaching*, *Classroom Discourse*, *Foreign Language Testing and Teaching*, *Foreign Languages in China*, *Foreign Language World*, *Foreign Language Education in China* and *Contemporary Foreign Languages Studies*.

#### **Abstract: The Interface between Language Teaching and Testing: the Role of Local Standards**

In recent decades, especially since the publication of the Common European Framework of Reference for Languages (Council of Europe, 2001), language testers are often encouraged or compelled to align locally developed language tests to internationally influential language frameworks. While such an alignment may increase the transparency or enhance the power of local assessments, the question remains as to whether or to what extent language education in a local context benefits from the alignment. In this talk, I will discuss the interface between language teaching and language testing and argue for the significance of contextualized local standards underpinning language education policies.

I will provide the rationale for developing the China's Standards of English, a national framework of English recently developed in China, and emphasize the importance of embedding language standards in the local educational context. I will also discuss the challenges facing the use of national standards for vastly different contexts of English language education in China. Standards for English language education also take the form of national curricula, and high-stakes language assessments are closely aligned to curricular requirements. As an illustration, I will examine the role of the national College English Curricular Requirements in the reform of the College English Test, a national testing system for learners of English in higher education, and discuss the interaction between English language teaching and testing. Finally, I will examine the intricate relationship among the national standards of English, national curricula, and high-stakes assessments, and make suggestions for the way forward.



### Luke Harding

- Senior Lecturer in the Department of Linguistics and English Language
- Lancaster University
- United Kingdom

Luke Harding is Senior Lecturer in the Department of Linguistics and English Language at Lancaster University. His research interests are mainly in language assessment. He has published in the areas of listening assessment, pronunciation assessment, language assessment literacy, diagnostic language assessment, and the challenge of English as a Lingua Franca for language assessment. Luke's work has appeared in *Language Testing*, *Language Assessment Quarterly*, *Applied Linguistics and Language Teaching*. He is the author of a book: *Accent and Listening Assessment* (2011), and is currently the co-editor of the journal *Language Testing*.

#### **Abstract: Language Assessment Literacy: Past, present, future**

Language assessment literacy (LAL) has experienced a surge of interest in the field of language testing over the past five years, with several conferences and a special issue of *Language Testing* devoted to the theme. A number of local and international projects have also been conducted to explore language assessment literacy among teachers, in particular, as well as other stakeholder groups such as policy makers, admissions officers and examiners. Despite this interest, however, there remain conceptual and practical challenges in defining LAL needs across different roles/professions and different contexts, and also in moving beyond understanding needs to fostering more tangible enhancement of LAL.

In this talk, I will discuss the trajectory of language assessment literacy by tracing developments in the recent past, describing a current LAL survey project, and discussing recommended priorities for the future. The talk will begin with some recent examples of LAL-related problems identified in news media reports as a way of framing key issues. Then, a brief survey of different theoretical models of LAL (e.g., Brindley, 2001; Inbar-Lourie, 2008; Fulcher, 2012; Taylor, 2013) will be presented with a view to identifying overlapping themes and unresolved issues. The talk will then turn to focus on a current project – the Language Assessment Literacy Survey (Kremmel & Harding, forthcoming) – a large-scale survey which provides empirical evidence of both the dimensions of LAL, and the needs, lacks and wants of specific stakeholder groups (N > 1000) across 71 countries. The development of the survey will be described, and key findings will be presented comparing the LAL needs of different stakeholder groups, and of teachers, specifically, across different educational contexts.

Finally, the talk will conclude with a discussion of the future direction of LAL, with a focus on the place of language in LAL, and the need to move from language assessment literacy to language assessment communication.





## Panel discussion speakers

# Topic: Perspectives on Language Frameworks

### Discussant/Chair



#### Nicholas Saville

Director of Research & Thought Leadership in Cambridge English and Secretary-General of ALTE, University of Cambridge, United Kingdom

Dr Nicholas Saville is Director of Research & Thought Leadership in Cambridge English (University of Cambridge) and Secretary-General of ALTE. He has a PhD in Language Assessment, an MA in TEFL and a BA in Linguistics. Before joining Cambridge Dr. Saville taught at the University of Cagliari (Italy) and worked in Japan. He is an expert on the Common European Framework of Reference (CEFR) and edits the English Profile Studies series with Dr. Fiona Barker. Dr. Saville is joint editor of Studies in Language Testing (CUP) with Prof. Cyril Weir and has recently completed a volume on Learning Oriented Assessment (LOA) with Neil Jones (SILT 45).

### Panellists



#### Masashi Negishi

Professor, Tokyo University of Foreign Studies, Japan

Dr Masashi Negishi is a professor at Tokyo University of Foreign Studies, Japan. He has been involved in the research and development of a number of English language tests, such as the GTEC (Global Test of English Communication), EIKEN, the National Center Test, and the SST (Standard Speaking Test) in Japan. He has also been on several committees for MEXT's university entrance exam reform and for national surveys of English proficiency among Japanese secondary school students. Since 2016, he has been head of the CEFR-J research team. He has just published a book on testing for Japanese teachers of English.



#### Neus Figueras Casanovas

Senior Researcher, Regional Ministry of Education in Catalonia, Spain

Dr. Neus Figueras Casanovas coordinated the development of foreign language curricula and certificate examinations for adult learners for 20 years in the regional Ministry of Education in Catalonia, Spain. She has been involved in a number of international research and development projects, and has given courses and presented in universities in Spain, in various European countries, in Asia and in the USA.

She collaborates regularly with the Council of Europe in the uses and the dissemination of the CEFR in testing and assessment and is one of the authors of the Manual for Relating examinations to the CEFR (Council of Europe, 2009). Her current interests also include assessment literacy and the development of context relevant assessments.

She was the first President (2004-7) of EALTA (European Association for Language Testing and Assessment, [www.ealta.eu.org](http://www.ealta.eu.org)) and received the Third International British Council International Assessment Award in 2015.



## Lianzhen HE

Professor, School of International Studies, Zhejiang University, China

Dr HE's main research interests are language testing and English language teaching. She got her Master's degree from the University of Birmingham (1992) and her PhD degree in language testing from Guangdong Foreign Studies University, China (1998). Dr HE was a senior visiting scholar at University of California at Los Angeles in 2004, and was local chair of the 2008 Language Testing Research Colloquium (LTRC) held at Zhejiang University. She was the Benjamin Meaker Visiting Professor at University of Bristol in 2014. She has also been a key-note speaker at several international conferences.

Dr HE has directed more than 10 large-scale research projects on language testing and language teaching. She is also Deputy Chair of the Advisory Board of Foreign Language Teaching and Learning in Higher Education Institutions in China and has been on the editorial board of a number of journals including *Language Assessment Quarterly*.

Dr HE has published widely in applied linguistics and language testing, including 28 English textbooks which are used nationwide in Chinese universities, a monograph on cognitive computer adaptive language tests and a number of journal articles on language assessment, discourse analysis and language teaching.



## Wenxia ZHANG

Professor of Applied Linguistics in the Department of Foreign Languages and Literatures, Tsinghua University, China

Dr ZHANG Wenxia is professor of Applied Linguistics in the Dept. of Foreign Languages and Literatures of Tsinghua University in Beijing where she has been teaching various courses since 1987. She received her Ph.D. in EFL writing assessment in 1999 from Hong Kong Polytechnic University, under the supervision of Professor Liz Hamp-Lyons. Her current research interests include: EFL teaching and learning in the Chinese context, automated writing evaluation, and language testing and assessment. Her major research publications include "World Englishes: Issues in and from writing assessment" (with L. Hamp-Lyons, 2001), *The rhetorical patterns found in Chinese EFL student writers' examination essays in English and the influence of these patterns on rater response* (Tsinghua University Press, 2004), and *Understanding the Learning Variables in Chinese EFL Contexts* (with M. Liu 2014, Beijing University Press).



## Sha WU

Director of Foreign Language Assessment Department, National Education Examinations Authority, China

Ms. Sha WU is now serving as Director of the Foreign Language Assessment Department at the National Education Examinations Authority (NEEA), under the Chinese Ministry of Education. NEEA is responsible for test development and administration for many national examination programs in China, including the Gaokao. NEEA is now undertaking several reforms to the Gaokao, particularly in the area of foreign language tests.

Ms. WU has been deeply involved in the development and administration of a number of language test programs and is now engaged in the development of China's Standards of English (CSE), along with a new English language test product.

## Panel discussion speakers

# Topic: Artificial Intelligence in Language Assessment

### Discussant/Chair



#### **Prof Barry O'Sullivan**

Head of Assessment Research and Development, British Council, United Kingdom

Professor Barry O'Sullivan is the Head of Assessment Research and Development at the British Council. He has researched factors affecting spoken performance; assessing rater behaviour; assessing speaking and writing; specific purpose assessment; benchmarking and standards; and has worked on the development and refinement of the socio-cognitive model of test development and validation since 2000. He has presented his work at over 150 conferences around the world, while over 90 of his publications have appeared in a range of international journals, and in many books and technical reports. In addition, he has published five books in the area of language testing. He is the founding president of the UK Association of Language Testing and Assessment and holds honorary and visiting chairs at the Universities of Reading and Roehampton (UK) and at the University of Lisbon (Portugal). In 2016 he was awarded fellowship of the Academy of Social Science in the UK.

## Panellists



### Alistair Van Moere

Chief Product Officer, MetaMetrics, United States

Dr. Alistair Van Moere is Chief Product Officer at MetaMetrics Inc. and is focused on improving education for learners of all ages. Previously, he was President of Pearson's Knowledge Technologies group, where he managed automated scoring services for writing, speaking and math, as well as delivering tens of millions of assessments in the US each year. Dr. Van Moere's PhD from Lancaster University won the Jacqueline Ross TOEFL Award for best dissertation in language testing, and he holds an MBA from Warwick Business School. He has authored over 20 publications with a focus on educational technology and language assessments.

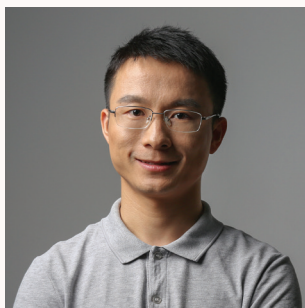


### Helen Yannakoudakis

Senior Research Associate at the department of Computer Science and Technology, University of Cambridge, United Kingdom

Dr Helen Yannakoudakis is a Senior Research Associate at the Department of Computer Science and Technology of the University of Cambridge, specialising in the development of Automated Language Teaching and Assessment technology, and the use of AI in Education. Helen is also a Fellow of Girton College, Cambridge; a Research and Development specialist at iLexIR, where she is working on the development of viable commercial applications in AI; and a committee member of Women@CL, an organisation that supports and mentors women engaged in computing research. Helen's research interests lie at the intersection of Computational Linguistics, Machine Learning and Visualisation, and particularly revolve around the areas of automated language teaching and assessment, self-assessment and tutoring systems, error detection and correction, and second language acquisition. Inter alia, Helen's work has contributed to two applications: Write & Improve, an automated writing assessment system; and the English Profile Visualiser, a system that enhances human cognition of AI-model characteristics.

Helen holds a PhD in Natural Language and Information Processing from the University of Cambridge, during which she also worked on the English Profile Programme in collaboration with Cambridge English Language Assessment; an MPhil in Computer Speech, Text and Internet Technology; and a BSc in Computer Science.



### Si WEI

Vice Director of iFLYTEK Research, iFLYTEK, China

Dr. Wei received a BE degree in Electronic Information Engineering (2003) and a Ph.D. in Signal and Information Processing (2008) from the University of Science and Technology of China, Hefei. He joined iFLYTEK Research in 2008 and is currently the vice director of iFLYTEK Research. His research interests include pronunciation quality assessment, speech recognition, speaker and language recognition, speech synthesis, hand written recognition and natural language processing.



### Hui LIN

Co-Founder and Chief Scientist, Liulishuo, China

Dr Lin received his Bachelor's and Master's degrees from Tsinghua University, and his Ph.D. from University of Washington. He is the Co-Founder and Chief Scientist of Liulishuo. Before founding Liulishuo, Dr Lin was a research scientist at Google headquarter. His research area includes speech recognition, natural language processing, machine learning and big data mining. He has published over 30 papers in leading journals and conferences, and has been actively involved in various technical committees. Dr Lin holds multiple patents in US and China.



## Breakout Session on Day 1



## Topic: Language Proficiency Frameworks

**Hua DENG**

- Lecture
- Hunan Normal University
- China

Hua DENG is a college English teacher for almost 10 years in China, now in pursuit of her doctoral degree majoring in language testing. She is a member of the China's Standards of English, mainly in charge of the development of writing strategy sub-scale.

**Co-presenter:****Jie DENG**

- Professor
- Hunan Normal University
- China

Jie DENG has been an English teacher in Hunan Normal University for almost 30 years, during which she has accomplished her doctoral degree and has developed a unique perspective in language testing, especially in validation research. As a member of the China's Standards of English, she is the leader who designs the overall framework of initial writing strategy sub-scale and conducts its validation test in various aspects.

**Co-presenter:****Shen ZOU**

- Professor
- Shanghai International Studies University
- China

**Abstract: Construct Validation of the Writing Strategy Competence of the China's Standards of English**

It is necessary to conduct validity verification of language proficiency scales in order to ensure the scientificity and practicability of the scales. As a sub-scale of the China's Standards of English, the English writing strategy competence scale is developed on a basis of the writing strategic framework, which proposes a cognitive model of writing comprising three processes of planning, formulation and revision, with strategies typical of each process extracted. The present study, based on a large-scale questionnaire, aims to examine the construct validity of the writing strategic framework through the use of structural equation modeling (SEM). The data were collected in more than 100 schools and colleges nationwide during 2016 by the National Education Examinations Authority. It was found that, firstly, the writing strategy competence sub-scale mirrors the cognitive process of planning, formulation and revision; secondly, the writing strategy competence can be measured through typical writing activities or special behaviours occurring in the writing process, and it concludes that cognitive strategies and regulatory strategies are main contributors to the writing strategy competence; thirdly, the employment of writing strategies in different writing competence grades is remarkably differentiated, and strategies in the process of planning usually occur above A3, which means that planning is a high-level cognitive activity.

**Linlin CAO**

- Phd Candidate in Linguistics & Applied Linguistics
- Guangdong University of Foreign Studies
- China

Linlin CAO is currently a Phd candidate in the national key research centre for linguistics & applied linguistics in Guangdong University of Foreign Studies. Her research orientation is rater effect, validity of language proficiency scales and learner corpora in language testing and assessment. In the last few years, she has participated in several projects like the Construction of Reading Proficiency Scale in China's Standards of English, the National

English Tests Project, etc. Papers presented in conferences include the Automatic and human rating differences in computerized listening and speaking test in Guangdong NMET, in the 1st ICLTA in 2015, etc.

**Co-presenter:****Yongqiang ZENG**

- President of Guangdong Teachers College of Foreign Language and Arts
- China

Yongqiang ZENG is currently the president of Guangdong Teachers College of Foreign Language and Arts, doctoral supervisor in Guangdong University of Foreign Studies, CPPCC member of Guangdong province. Professor Zeng's research orientation includes language testing, computational linguistics and long-distance English education and research. He's also director of more than 10 national and provincial level research projects and has published several academic monographs. Over 50 academic papers have been published in international and domestic core journals like the Journal of Education Measurement, Language Testing in Asia, Modern Foreign languages, Foreign Language World, Media in Foreign Language Teaching, etc.

**Abstract: Exploring Criterial Features Across Levels of CSE Reading Subscale Descriptors**

As influential language proficiency scales under construction, China's Standards of English (CSE) employ the 'can-do statements' to comprehensively describe proficiency levels.

However, the description is too general and the quantitative justification takes each level of descriptors as a unidimensional competence unit. The internal diversity or criterial features within levels of descriptors remain unexplored. This study takes a corpus-based approach to investigate the criterial features of reading subscale descriptors from CSE A2 to B2 levels, in particular the cognitive competence. The components constitute the descriptors are

coded and classified according to Bloom's revised taxonomy of educational objectives. Frequency and type measures are employed to obtain trends in the difference in each level. As empirical support, a corpus is set up, containing the tagged reading part of representative textbooks and the recent ten years of examination papers in each CSE levels. Statistical methods like ANOVA and multiple regressions are used to examine whether there exist statistically significant patterns. Four criterial features, more concretely topic, genres, cognitive verbs and cognitive targets are observed in the corpus. Furthermore, student interview data are analysed to validate the results. Finally, this study also sketches out some avenues for future research.



### Yidong JIA

- Dean of English Department, School of Foreign Studies
- Shandong University of Finance and Economics
- China

Dr Jia obtained his PHD in Language Testing and Assessment from the Beijing Normal University (BNU). His research interest includes English testing and assessment as well as English language teaching and education. He joined the development group of English organizational knowledge, part of the project of China's Standards of English (CSE) launched in October, 2014, the most comprehensive English proficiency project hitherto in China. Currently he is an associate professor at Shandong University of Finance and Economics.

#### Co-presenter:

### Zunmin WU

- Professor
- Beijing Normal University
- China

#### Abstract: Developing Self-Assessment Strategy Inventory for Acquiring English Organizational Knowledge

China's Standards of English (CSE), the most large-scale English proficiency project hitherto in China, involves horizontally six language skills (listening, speaking, reading, writing, translation, interpretation) and knowledge (organization and pragmatics), ranging vertically from the beginning level of learning to the most academically advanced native-like level. This paper addresses self-assessment strategy for acquiring organizational knowledge. Due to the consensus that the distance between L1 and L2 influences the acquisition of L2, it is vital for English learners to focus on organizational knowledge since Chinese and English belong

to two quite different language families. More importantly, the less similar L1 and L2 are, the more strategic learners will become. The self-assessment strategy checklist for organizational knowledge attempts to illustrate what organizational knowledge learners should possess by developing can-do statements. The descriptive statements are indicators of learners' acquiring organizational knowledge by employing cognitive, metacognitive, social and affective strategies. They are hoped to serve some purposes for stakeholders. For example, they provide a way for learners to chart their progress through incremental steps and foster their critical thinking and intrinsic motivation. Besides close-ended surveys of strategy use, the author proposes the inclusion of open-ended items by asking learners to write down their own strategies, allowing them freedom to develop their own preferred ones. One size doesn't fit all—while a general strategy checklist is developed, personalized ones also required. Besides, the self-assessment strategy inventory needs further validating in the future.



### Weiying HUANG

- Lecturer
- East China University of Technology
- China

Weiying HUANG, Ph.D of Applied Language Studies, is a lecturer at East China University of Technology, where she teaches Chinese tertiary EFL learners courses of English reading and writing. She also gives lectures on Language Testing to MA students of Applied Linguistics. Her research interests include Language Testing, and Interlanguage Pragmatics. She has conducted research into Chinese students' pragmatic development and test development to measure their pragmatic competence.

#### Co-presenter:

### Baocheng HAN

- Professor
- Beijing Foreign Studies University
- China

#### Abstract: Constructing an Interlanguage Pragmatic Scale for Chinese Tertiary EFL Learners

The past two decades have witnessed a rapid growth of language proficiency scales, for instance, the Canadian Language Benchmarks (CLB), the International Second Language Proficiency Ratings (ISLPR), and the Common European Framework of Reference (CEFR). However, although pragmatics plays an essential role in communicative competence (Bachman, 1990; Bachman and Palmer, 1996), descriptors

of pragmatic competence in existing language proficiency scales are rare.

The current study aims to explore a data-based approach to construct an ILP competence scale for Chinese tertiary EFL learners. ILP competence, in this exploratory study, is operationalized as learners' ability to successfully perform various speech acts in oral communication. 60 scenarios were formulated as a test battery of Oral Discourse Completion Tasks (ODCT) covering 8 speech acts. 80 Chinese university students from 3 universities participated in the ODCT pragmatic test, and their pragmatic performance were rated and then analyzed to reveal distinctive features of language use at three levels. Representative linguistic features were formulated as descriptors of pragmatic competence to form a 3-level scale. The paper also discusses the implications of the scale of ILP competence for pragmatic learning, teaching and testing.



### Li FENG

- Associate Professor of Institute of Applied Foreign Languages
- Heilongjiang University
- China

Li FENG, associate professor of Institute of Applied Foreign Languages in Heilongjiang University. She got her MA in English linguistics and is a current PhD candidate in Heilongjiang University.

#### Co-presenter:

### Ming YAN

- Professor
- Heilongjiang University
- China

#### Abstract: The Development of Descriptors of Translation Proficiency under the Project of CSE: Problems and Solutions

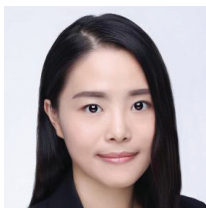
As a part of the project of China's Standards of English, the construction of translation Proficiency scales has drawn great attention, in which the development of descriptors of translation proficiency has played a crucial role. This paper will report the solutions to the major problems and difficulties encountered in the three-year construction process. The following content will be covered:

1. Premises of the research: The unified parameter system designed for the CSE project serves as the guideline and reference framework. Since some parameters have not been readily studied with regards to translation proficiency, we had to do our own research accordingly, which is yet to be argued and empirically testified.

2. Constructionist approach: Language scales are worlds constructed by descriptors in linguistic form as cornerstones. The quality of language scales is heavily decided by the quality of descriptors. A constructionist view of scales has helped us identify all the stake-holders and analyze their roles and potentialities of making contributions and being confined.

3. Research process: The development of translation proficiency descriptors is a process of clarifying the descriptive elements of the system of translation proficiency, including large-scale pooling and on-going polishing of raw descriptors. The sources of collecting raw descriptors and the stages and methods of polishing them will be introduced here.

4. A semantic-formal analysis of the descriptors: A comprehensive semantic-formal analysis of the structure of the descriptors has been done. We have discovered that the problems of linguistic descriptions have been caused by the vagueness in understanding the concept of translation proficiency.



### Yi XU

- Lecturer of School of Interpreting and Translation Studies
- Guangdong University of Foreign Studies
- China

Yi XU is a Lecturer in the School of Interpreting and Translation Studies at Guangdong University of Foreign Studies (GDUFS) in China. Xu received her B. A. degree in English Language and Literature in 2010, M. A. in Language Testing in 2013 and PhD in Translation Studies in 2016. Xu's research interests include interpreter's competence and testing and assessment in interpreting.

**Co-presenter:**

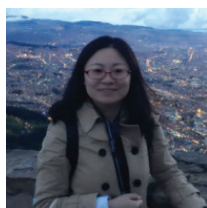
### Lei MU

- Professor
- Guangdong University of Foreign Studies
- China

#### **Abstract: Strategic Competence in Interpreting Scale of CSE: A Meta-cognitive Perspective**

Strategic competence is a core component of interpreting competence. This paper first defined the construct of strategic competence in interpreting competence from the perspective of meta-cognitive theory. Focused on consecutive interpreting (CI), this paper then introduced the methods and procedures of quantifying the parameters of meta-cognitive strategic competence in interpreting. The study employed

bibliometric analysis of documents, interpreting experiments and immediate retrospective interviews to collect descriptors and explore potential factors of meta-cognitive strategic competence in CI. Drawing on the results of interviews and Exploratory Factor Analysis, the paper then elaborated on the descriptor framework and descriptors strategic competence in interpreting scale of CSE.



### Wei JIE

- PhD Candidate
- Shanghai Jiao Tong University
- China

**Co-presenter:**

### Yan JIN

- Professor
- Shanghai Jiao Tong University
- China

#### **Abstract: Linking Test to the China's Standards of English-Speaking Scale: CET-SET4 as a Case Study**

This study presents the procedures and process of linking the Spoken English test band 4 of the College English Test (CET-SET4) to the China Standards of English-Speaking Scales (CSE-SS). The implementation of the linking process is based on the suggestions of the Common European Framework of Reference for Languages (CEFR) linking manual (Council of Europe, 2009). Familiarization tasks are adapted for the self-study of judges for familiarization. The test centered, basket method is used in the first session of standardization where judges need to place each test item to a specific CSE-SS level. The examinee centered, contrasting-groups method is employed in the second session of standardization where judges need to rate examinees' performance to corresponding CSE-SS scale to set up a cut-score. Ten judges from different professional fields (teachers, test-developers, scale-developers, psychometricians) circumscribe these two multidisciplinary decision-processes. Evidence is provided for the linking both in terms of quantitative as well as qualitative analyses. In the end, reflections and discussion on the standard setting process are reviewed by looking at the quality of the linking and the judges' report at the end of each session.



### Jacqueline Martin

- Young Learners Project Lead
- Pearson Education London
- United Kingdom

Jacqueline Martin is Young Learners Project Lead, Global Scale of English at Pearson. In this role, Jacque leads on the development of creating Young Learner specific learning objectives aligned to the Global Scale of English and works with Content teams to ensure that these learning objectives underpin all new products and services. Formerly Jacque worked as Primary Teacher in the UK before entering the world of publishing as an editor. Jacque joined Pearson in 2009 working in and then leading the ELT Primary Publishing team until her move to the Global Scale of English in Jan 2017.

#### **Abstract: Global Scale of English: Ensuring Global Standards Meet Local Requirements**

The Common European Framework of Reference for Languages (CEFR) has become a standard framework of reference for English Language teaching and assessment around the world even though, as the name suggests, it was originally created for a European context. Its influence has become global for many reasons, not least of which is the lack of alternative standards to support ELT teachers and testers. Few would dispute the benefits that standards in education can bring-but equally there needs to be evidence that such standards meet the needs of teachers and learners in their specific contexts. How can we be sure that a global standard such as the CEFR meets the local requirements in East Asia?

This presentation looks at the development of the Global Scale of English (GSE), a proficiency framework that builds on and extends the CEFR to account for the needs of more learners: those studying English for academic or professional purposes as well as young learners. The presenter will look at the research carried out with teachers from over 50 countries-including many in Asia-into defining a proficiency framework that addresses some of the limitations of the CEFR and takes into account the needs of local curricula.

The presenter will also report on a second research phase aimed at aligning other language frameworks and standards used in Japan to the GSE. Such studies are mutually beneficial: confirming that global standards can be applied to the local context and that local standards are globally relevant.





### Spiros Papageorgiou

- Managing Senior Research Scientist
- Educational Testing Service, Princeton
- United States

Spiros Papageorgiou is a Managing Senior Research Scientist in the Center for English Language Learning and Assessment at Educational Testing Service (ETS) in Princeton, NJ. Spiros received his doctoral degree from Lancaster University, UK, and worked as a language testing researcher at the University of Michigan before moving to ETS, where he conducts research on language assessment, primarily supporting the TOEFL family of assessments. He has published widely on standard setting, in particular in relation to the CEFR levels, score reporting and interpretation, and listening assessment. Spiros has served in the Executive Board of the International Language Testing Association (ILTA).

#### Abstract: Issues in Aligning Assessments to Language Proficiency Frameworks and Standards

The demand for alignment of assessments to language proficiency frameworks and standards has increased worldwide, because of educational reforms which push for accountability, including close monitoring of students' progress and use of standardized tests (Deville and Chalhoub-Deville, 2011). In the field of second language assessment, the alignment of tests to proficiency levels of the Common European Framework of Reference (CEFR; Council of Europe, 2001) has dominated relevant research for more than a decade. This research has implications for efforts to create and successfully implement localized language proficiency frameworks such as China's Standards of English (CSE). Aligning test content and scores to the levels of a language proficiency framework has been advocated as a way to facilitate score interpretations across different educational contexts (Council of Europe, 2009). However, researchers have also noted several issues when embarking on the highly technical alignment process, given the unavoidably generic description of proficiency contained in language frameworks (Alderson et al., 2006; Weir, 2005). Drawing in particular on research related to the CEFR, this paper discusses several important issues related to the alignment of assessments to language proficiency frameworks. The paper will examine the different methodological approaches adopted by researchers and test providers and consider the implications of these methodological approaches for the alignment of different assessments to the same language proficiency framework. The paper will conclude with recommendations for alignment in the context of implementing new, localized language proficiency frameworks.



### Nguyen Thi Mai Huu

- Executive Deputy Director
- National Foreign Languages 2020 Project Vice Director of the Center for Language Testing and Assessment
- University of Languages and International Studies, Vietnam National University
- Vietnam

Nguyen Thi Mai Huu is currently a PhD candidate in Applied Linguistics at the University of Languages and International Studies - VNU and a visiting scholar at the Institute of Education - Massey University. Her research interest includes language test development and validation, test specifications, classroom assessment, and assessment for learning. She was involved in the development of the Vietnamese, which is now a national test format of Vietnam.

#### Abstract: Developing the Vietnamese Standardized Test of English Proficiency

In 2008, Vietnam launched the National Foreign Languages 2020 Project aimed at enhancing foreign language education, especially English education in Vietnam to prepare for the economic growth of the country. Part of the Project's mission is to design an instrument to measure learners' English language proficiency. In response to the call, a team of researchers from ULIS - VNU with some assistance from the US Embassy in Hanoi in the form of expertise provided by Fred Davison was created to develop what is now known across Vietnam as VSTEP. B1-C1 (the Vietnamese Standardized Test of English Proficiency (VSTEP). This is a high-stake four-skill test, which aims at B1 to C1 (including B1, B2, and C1) English language proficiency range. B1 to C1 levels are set as expected learning outcomes for university graduates, post graduates, and English language teacher trainees, and many others who want to get a job. This paper will provide an account of how this team got there in terms of planning, conceptualization, design and development, try-outs, and informing the stake-holders. This paper also discusses how the team adapted and localised the CEFR. The VSTEP is now an established fixture in the assessment domain, and recognized nationally. Hundreds of thousands of people take the test annually. Furthermore, it is generating some positive washback on ELT in the country.

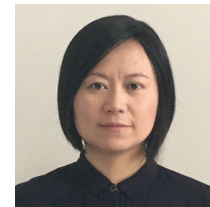


### Haiping WANG

- Postgraduate Supervisor at School of Foreign Studies
- East China University of Political Science and Law
- China

Dr. Wang is a post doctor and postgraduate supervisor at School of Foreign Studies in East China University of Political Science and Law, Shanghai. She holds a PhD in the field of Language Assessment from Shanghai International Studies University and was a visiting scholar at University of Bristol's Graduate School of Education. Her recent publications include: Wang, H. (in Press) *Testing Lecture Comprehension through Listening-to-Summarize Cloze Tasks* (monograph). Springer International Publishing; Wang, H. (2017). A Review of Studies on Cognitive Processes in Language Testing, *Foreign Language Testing and Teaching* (the only journal specialized in language testing in mainland China).

#### Co-presenter:



### Limin ZOU

- Dean of Jing'an Foreign Language High School
- Shanghai International Studies University
- China

Ms Zou, Ed.M., has been teaching English in SJFH ever since the foundation of the school. She has participated in the reforming of school English curriculum and has been the leader of Individualized English Reading course project. She is now working as dean of SJFH, who is responsible for the school course management and evaluation.

#### Abstract: Designing Localized CEFR-referenced SCELL for District Foreign Language Schools in Shanghai

District Foreign language schools have special needs on English teaching and learning standards to highlight the international characteristics of district education. In order to cater for those requirements, we have developed a set of localized CEFR-referenced English proficiency scales (SCELL: Shanghai Criteria of English for Language Learners) from Levels 1 to 6 to be aligned with grades 1 to 12 in foreign language schools, thus streamlining their English

teaching and assessment in a more systematic and scientific manner.

In the process of designing SCELL, we have localized CEFR by using bottom-up approach. First, the scale designing is synchronized with classroom teaching and assessment, which means teachers in the district foreign language schools involved themselves in rounds of discussions on anchoring descriptors of SCELL scales. Second, they implemented SCELL in the classroom to adjust their actual teaching practices. Hence designing of SCELL starts from classroom and serves classroom.

Validity of the can-do statements as ability descriptors across listening, speaking, reading and writing skills on SCELL is guaranteed by triangulating empirical data of students' actual language performance, teacher and expert opinion through focus group interviews and judgements by test developers. Validity arguments are being made through rounds of empirical data collection and this process is iterative. Scale-anchoring of SCELL lasted two years, which aims at providing descriptive information about the abilities that students need to attain the corresponding level and making recommendations for teachers to adjust their teaching methods and optimize the teaching effect based on our interpretation of SCELL.

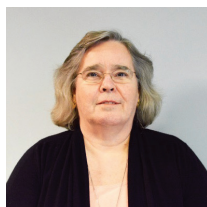


### Heather Koons

- Director, Research Services
- MetaMetrics
- United States

Dr. Heather H. Koons is Director, Research Services at MetaMetrics, where she works with clients to link their reading and mathematics assessments with the Lexile Framework for Reading and Quantile Framework for Mathematics. She also coordinates research projects related to text complexity, English learners, and vocabulary development. Koons taught English in East Asia and high school English in the U.S. before working in test development at the North Carolina Department of Public Instruction. Koons earned her M.A. in Education at Stanford University and her Ph.D. in Educational Psychology, Measurement, and Evaluation at the University of North Carolina at Chapel Hill.

#### Co-presenter:



### Eleanor Sanford-Moore

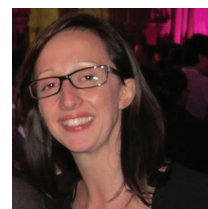
- Vice President for Research and Development
- MetaMetrics
- United States

Dr. Eleanor Sanford-Moore oversees MetaMetrics's research and development projects, linking studies, and other external consulting. For the past eight years, she has also been involved with studies related to English language learner assessments. Prior to joining MetaMetrics in 1998, she was with the North Carolina Department of Public Instruction and responsible for the development of many public school assessments. She has written test reviews for "Buro's Mental Measurements Yearbook," and served on various national committees related to large-scale assessment. Dr Sanford-Moore earned her undergraduate degree in statistics and master's and doctorate of philosophy in psychology from North Carolina State University.

#### Abstract: Examining Graded Readers Using CEFR Levels and Text Complexity Levels

Graded readers are texts utilized by EFL learners to build reading fluency and achieve varied academic and professional goals. Publishers typically assign a level (e.g., "Intermediate") to each text to inform readers of its relative difficulty. Each level, from the lowest to the highest, is expected to provide increasing challenge and present individuals with opportunities to enhance their English reading skills. Many publishers also assign Common European Framework of Reference for Languages (CEFR) levels to their texts. Educators and learners often select texts based on CEFR framework levels, but how well do these levels represent the texts' reading challenge?

To address this question, researchers identified graded readers (N = 463) from four series produced by globally prominent English language teaching publishers and a subset that had specific CEFR designations (N = 332). The extent to which their text complexity, denoted by Lexile measure, varied across graded reader/CEFR levels within and across series was examined. Results indicated that for both graded reader and CEFR, as the levels increased, their median text complexity measure also increased, but within the same graded reader or CEFR level, and across publishers, the text complexity varied greatly. These results suggest that learners using publisher-designated graded reader or CEFR levels to select texts may experience substantially different reading challenges from texts within the same level. The overlap in complexity across levels may limit learners' ability to select appropriately challenging reading material. The presentation will report the results of the study and address implications for EFL learners.



### Elizabeth Shepherd

- Senior Researcher
- British Council
- United Kingdom

Elizabeth joined the British Council in 2006 as part of the UK Education research team. She then worked for the British Council in Washington DC, USA where she managed the UK - US Higher Education Policy Dialogues and dissemination of multi-lateral partnership grant funding. Following this role she was posted to Hong Kong where she led the global research team Education Intelligence, developing their research and engagement strategy. Since 2015 she has been working as part of the UK based assessment research team developing a new global evaluation of English language capability, whilst undertaking a PhD in Applied Linguistics, looking specifically at language policy and planning.

#### Co-presenter:



### Martin Murphy

- Head of Sampling
- The Australian Council for Educational Research
- Australia

Martin Murphy is head of the Sampling Unit at the Australian Council for Educational Research (ACER). He has been involved with the design and selection of many major national and international educational surveys. He leads a team of staff that has drawn school samples for countries participating in the OECD Programme for International Student Assessment (PISA) survey in all six survey periods conducted since its inception in 2000, including the latest survey which will be in the field in 2015 and involves over 70 countries. He has developed and implemented sample designs for many of the surveys conducted under Australia's National Assessment Program, including national assessments in Civics and Citizenship, Science Literacy and ICT Literacy, as well as calibration and equating samples connected to the national census assessments of literacy and numeracy (NAPLAN)

**Abstract: Implementing an International English Language Evaluation; Learning from PISA, using Technology**

Cross national comparison of students' performance across different education systems has become an important reference point for educational policy development. Few are better known and more influential than the OECD's PISA programme. Creating increased debate around the validity and use of data from internationally comparative surveys has been the inclusion of sub-national units of comparison, most notably Shanghai within the 2012 round of PISA, and Beijing, Shanghai, Jiangsu and Guangdong (B-S-J-G) in 2015.

Since 2015 the British Council has piloted a global study with the aim of evaluating international English language capability to inform education system development and to diagnose the impact policy decisions can have on the development of learning systems over time.

Working in partnership with the Australian Council for Educational Research, the research has been conducted using an adaptation of PISA's two-stage cluster sampling process. 150 schools in each unit of comparison were randomly selected to take part, with 12 sampled students per school completing an English language assessment. The target student population also mirrors that of PISA, defined as students in grade 10 who receive a minimum of 90 minutes formal English as a foreign language study per week as part of the school programme.

Piloted across 4 education systems reflecting both national and sub-national units of comparison, this presentation will look in details at the sampling, assessment implementation and comparison of assessment outcomes across all 4 language learning skill areas in; Sri Lanka, Bangladesh, the Madrid region of Spain, and the Bogota metropolitan area of Colombia.



**Mingwei PAN**

- Associate Professor of Applied Linguistics
- Guangdong University of Foreign Studies, Guangzhou
- China

Dr Mingwei Pan earned his PhD from Shanghai International Studies University, and is now Associate Professor of applied linguistics at Guangdong University of Foreign Studies, China. His main research interests include test development and validation, language proficiency scale development, language assessment literacy and assessing young learners. He has published extensively on international journals and peer-reviewed books.

**Abstract: Validating the Writing Scales of the CSE: A Developmental Perspective**

The China's Standards of English (CSE) are intended to describe the extent to which language learners or users across different proficiency levels in the Chinese EFL context can use the target language in real-life performances. It is generally assumed that writing proficiency grows contingent upon the overall development of L2 proficiency. However, whether the developmental pace presents a linear regression in relation to learners' exposure to English, and how the development is instantiated in different aspects remain to be explored. Following Norris and Manchon's (2012) framework of writing proficiency development, this study sets out to validate the CSE writing scales by mapping the descriptors from various perspectives. Regarding learners' linguistic development, it was found from the IRT difficulty values that the CSE writing scales generally display a steady progress with an even interval between adjacent levels. The contexts where learners perform writing activities largely sit within an educational setting, though a shift can be felt from English for general purposes to English for academic purposes. However, learners' control over textual output seems to develop at a slower pace when they approximate a higher proficiency level. Similarly, their writing strategy development displays a rather intriguing scenario, where beginners are distinguished in terms of whether or not certain strategies are used, whilst advanced counterparts seem to have almost no significant differences in employing writing strategies. It is expected that this study will articulate more validity arguments for the CSE writing scales and enrich the existing body of literature on L2 writing development.



**Wenxing ZHANG**

- PhD in Language Assessment and Evaluation
- Jiangxi Science & Technology Normal University
- China

Dr Zhang is currently an associate professor in Jiangxi Science & Technology Normal University. He received his PhD degree from Shanghai International Studies University in 2016. His published work was mainly on second language assessment, language teaching, and his research interests include second language assessment, language proficiency scale development, cognitive diagnosis, and language teaching. At present, he is a key member of the writing sub-project (hosted by Prof. Shen ZOU) of the China Standards of English (CSE), sponsored by the National Educational Examinations Authority (NEEA).

**Co-presenter:**



**Shen ZOU**

- Professor in School of English Studies
- Shanghai International Studies, Shanghai
- China

Prof Zou is a professor in the School of English Studies, Shanghai International Studies University. Currently he is a member of the National Foreign Language Teaching Advisory Committee under the Ministry of Education in China, and he is also the chief test developer for Test for English Majors (TEM). His main research interests are: language assessment, program evaluation, and teaching of writing. At present, he is head of the writing sub-project of the China Standards of English (CSE), sponsored by the National Educational Examinations Authority (NEEA).

**Abstract: Exploring the Adaptability of the CEFR in Developing the TEM Writing Ability Scale**

The CEFR has had profound impact on language teaching, learning and assessment not only in Europe but also in other parts of the world. This study focuses on the adaptability of some CEFR writing descriptors in developing a can-do writing scale for Test for English Majors (TEM) in China. First, we developed an initial pool of writing descriptors, consisting of descriptors of written interaction from the CEFR and those extracted from the current teaching and testing syllabuses for English majors in China. Then, a questionnaire was constructed on the basis of the collected descriptors in order to elicit university teachers' views on the importance of these descriptors. In order to further investigate which specific level the remaining descriptors would fall into, 35 university teachers of English were invited to complete the revised questionnaire while rating 36 TEM scripts. Then, band-setting of the descriptors was initially determined on the basis of the statistical analysis of the questionnaire data. In order to collect further evidence for our calibration of the descriptors, 8 university teachers of English were interviewed. Our study results have showed that the investigated CEFR writing descriptors can be used in the description of the writing proficiency of our English majors, but most of the CEFR descriptors surveyed have had their levels altered on our writing scale. And the biggest changes for the CEFR descriptors have occurred among descriptors between logits 0 and 1 on the newly constructed writing scale. The findings have implications for further research in CEFR's adaptability in other Chinese contexts.

## Breakout Session on Day 1



# Topic: Assessment Literacy for Language Teachers

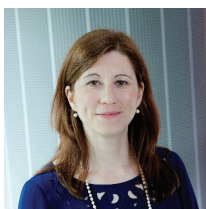


### Rita Green

- Visiting Teaching Fellow
- Lancaster University
- United Kingdom

Dr Rita Green has worked in the field of language testing for over 30 years acting as a trainer, consultant and researcher for a wide range of clients, including the British Council, DfID, Thai Ministry of Foreign Affairs, FLTRP and Beijing Foreign Studies University. She has carried out numerous workshops internationally, most recently as Director of the Hornby Regional Training Programmes in Dubai, Dhaka and Tashkent. She is a Visiting Teaching Fellow at Lancaster University as well as an EALTA Expert Member. She is the author of *Statistical Analyses for Language Testers and Designing Listening Tests: A Practical Approach*.

#### Co-presenter:



### Tineke Brunfaut

- Senior lecturer
- Lancaster University
- United Kingdom

Dr Tineke Brunfaut is a senior lecturer at Lancaster University, UK. Her main research interests are in language testing, and reading and listening in a second language. She is a recipient of the ILTA Best Article Award for a study on diagnostic assessment (Alderson, Brunfaut, & Harding, 2015) and of the TOEFL Outstanding Young Scholar Award for her professional activities and contributions to the field of language assessment. Her work has been published in journals such as *Applied Linguistics*, *Language Assessment Quarterly*, *Language Testing*, *Studies in Second Language Acquisition* and *TESOL Quarterly*.

#### **Abstract: Establishing Baselines for Listening and Speaking in Bangladesh: Using mixed-methods**

In light of proposals by the Bangladesh Ministry of Education for changes to the assessment of English listening and speaking skills in the country, three baseline studies were commissioned - each focusing on a different level of schooling - to ascertain the current state of play in the teaching and assessment of these skills in the Bangladesh school system.

In this presentation, we will focus on the research conducted at the Higher Secondary school level. More specifically, we will describe in detail the methodological approach we adopted that enabled us to gain comprehensive insights, in a relatively short period of time, into current classroom practices and stakeholder perceptions on English listening and speaking in Higher Secondary schools in Bangladesh. We will discuss how the use of four different research methods (expert judgements, classroom observations, interviews, and focus groups) with four different stakeholder groups (teachers, students, school principals, and community members) ensured the recording of multiple perspectives and triangulation of the data. In our talk, we will share with the audience our research instruments and procedures, and describe their advantages as well as how methodological challenges were offset.

We will argue how the highly structured mixed-methods research design ensured the practical feasibility of the study and enhanced confidence in the baseline findings, which include insights into teachers' and students' English listening and speaking proficiency, exposure to and practice in oral English, teachers' language assessment literacy, and issues regarding the level of readiness to implement continuous assessment in Bangladesh Higher Secondary schools.



### Hyoshin LEE

- Director of Foreign Language Centre / Director of the Dept. of English Language
- Konkuk University Glocal Campus
- Korea

Hyoshin Lee is a director of the Dept. of English Language and Foreign Language Center, Konkuk University Glocal Campus, Chungju, Korea. Before moving to the university, she taught English at state secondary schools for about 20 years. She has worked on teacher/trainer training since she received her doctorate degree in teacher development in 2003 from The University of Manchester, UK. She has been involved in a wide range of English education including teacher education and curriculum development. Her major interests include continuous professional development, curriculum alignment, learner autonomy and ESP.

#### **Abstract: Korean English Teachers' Voices on Aligning National Curriculum, Teaching and Assessment and its Implications for Teacher Education**

In recent years, Korea has emphasized more than ever the reform of the national curriculum for the preparation of the future society. However, in reality, the link between the national curriculum and the teaching and assessment at the actual school site is considered to be very poor. With critical understanding of this context, the study aims to raise issues regarding the discrepancy between national curriculum, teaching practice and school assessment at secondary schools in Korea and make suggestions to narrow the gap between those, with particular reference to language teacher education. The study employed document analysis, questionnaire survey, in-depth interview and classroom observation. The findings of the study showed that teachers had low level of understanding of achievement assessment with low motivation of implementing it leading to the low level of connection between achievement standards suggested by the national curriculum and teaching practice. It was also found that there was Conflict within assessment policies (i.e. ideal vision of achievement assessment policy vs. reality-focused school assessment policy for high school or university entrance). Low level of teachers' professionalism on assessment with too much reliance on textbooks was also observed. Model development to apply the achievement assessment suggested by the national curriculum, development of textbooks based on assessment standards, classroom-based teacher education and creation of teachers' collaborative working culture were strongly suggested.





### Duyen Tran

- Academic Project Manager
- Hanoi University of Industry
- Vietnam

Dr Duyen Tran is the Manager of the language education project and the Head of Language Practice Division, Faculty of Foreign Languages, Hanoi University of Industry, Vietnam. She has many years of experience working as an English teacher, examiner and test item reviewers at her university. She has been actively involved in designing and conducting teacher professional training courses, especially in English language teaching and assessment. She received her PhD degree from Queensland University of Technology, Australia in 2015. Her research areas include language teaching and assessment, the employment of ICT in language assessment, and English for Special Purposes.

#### Co-presenter:



### Tue Hoang

- Dean of Faculty of Foreign Languages
- Hanoi University of Industry
- Vietnam

Dr Tue Hoang is the Dean of the Faculty of Foreign Languages, Hanoi University of Industry, Vietnam. He has been working as an English teacher and an educational administrator at his university for more than 15 years now. He has conducted many studies and projects to improve the quality of English language education at tertiary level. He received his Doctor of Education degree from Queensland University of Technology, Australia in 2015. His research interests include ICT in education, blended learning, English for Occupational Purposes and Teacher Professional Development.

#### Abstract: Teachers' Perspectives and Practices of EFL Assessment at Vietnamese Universities

Assessment literacy has recently been identified as an important area for EFL teachers' professional development in Vietnam. While assessment literacy encompasses a wide range of skills and knowledge with priority varies according to the educational contexts, few studies have been conducted to identify specific areas of assessment literacy that EFL teachers in Vietnam need to develop. This

paper, therefore, reports on a qualitative case study which investigated teacher's perspectives and practices of EFL assessment at one public and one private university in Vietnam. Data were collected from 14 one-to-one interviews with six teachers and two executives, over-50-hour classroom observations, six focus group interview with students and analysis of related documents. Findings pinpointed important elements of assessment literacy that the teachers need to develop to improve the effectiveness of their teaching and assessment practices. The findings also revealed significant differences between early careers and mid-career EFL teachers in terms of their perspectives and practices of EFL assessment, which seems to be resulted from recent changes in EFL teacher training programs in Vietnam. The study carries important implications for EFL teacher training and professional development in Vietnam and in other countries with similar educational contexts.



### Quynh Le

- Head of ELT Department, Faculty of Social Sciences and Humanities
- Hoa Sen University
- Vietnam

Quynh Le has been teaching English for more than 10 years. He obtained an MA in English Language Teacher Development in 2005 and a PhD in Education in 2013, both from the University of Nottingham, UK. His research interests include learner autonomy, learning strategy, language assessment, and vocabulary learning. At present, he is Head of ELT Department, Faculty of Social Sciences and Humanities, Hoa Sen University, Vietnam.

#### Co-presenter:



### Trang Lam

- Program coordinator – ELT Dept.
- Hoa Sen University
- Vietnam

Trang Lam is currently an English lecturer and also a program coordinator of the English Language Teaching Department at Hoa Sen University. She finished a Master's Degree in TESOL. She has been teaching English at the university level for over 8 years. Her many interests include professional development, curriculum development, language teaching and language assessment.

#### Abstract: Aptis and Teachers' Conceptions of Assessment and Teaching Practice

Assessment of students' learning is one of the most important responsibilities of teachers in the English language classroom. The use of effective and appropriate assessment tools allows teachers to gauge students' learning outcomes and make informed decisions to improve their quality of teaching. Therefore, improving teachers' assessment literacy has a pivotal role in enhancing students' learning. In exam-oriented contexts, language assessment literacy is especially vital to the promotion of actual acquisition of proficiency as opposed to the commonly used "teaching to the test" approach. This paper presents the results of a study related to teachers' conceptions of assessment and teaching practice resulting from the implementation of a new language proficiency test. It focuses on investigating how teachers' conceptions of assessment develop when the Aptis test is introduced to measure their students' proficiency at a Vietnamese higher education institution. Using a mixed method approach, the study also seeks to examine how this influence their teaching practice when the test is used as a final exam for students completing the English programme offered by the institution. Findings confirm that the introduction of the new proficiency test has positive effects on teachers' conceptions of assessment as they become more aware of the features of the proficiency test and their implications for teaching. The study also finds that the transparency and availability of materials related to the specifications of the proficiency test are conducive to enhancing teachers' assessment literacy and encouraging them to "teach to the construct" of the test.



### Vivien Berry

- Senior Researcher, English Language Assessment
- British Council
- United Kingdom

Vivien Berry is a Senior Researcher, English Language Assessment at the British Council, where she leads an assessment literacy project to promote understanding of basic issues in language assessment, including the development of a series of video animations, with accompanying materials. With external colleagues, she also investigated the speaking construct in face-to-face and video-conferencing delivered modes of an international speaking test. She has published extensively on many aspects of language assessment and regularly presents research findings at international conferences. Vivien has also worked as an educator and educational measurement/assessment specialist in tertiary contexts in Europe, Asia and the Middle East.

**Co-presenter:****Susan Sheehan**

- Senior Lecturer in TESOL
- University of Huddersfield
- United Kingdom

Susan Sheehan is a Senior Lecturer in TESOL at the University of Huddersfield where she teaches on both undergraduate and postgraduate courses. She has a long-standing interest in testing and assessment and has presented on the subject at various national and international conferences. Before joining the University of Huddersfield Susan worked for the British Council, where she oversaw the creation of a new placement test and co-wrote the British Council / EAQUALS Core Inventory of General English. Her latest project investigated teacher perceptions of language assessment literacy with Vivien Berry and Sonia Munro.

**Abstract: What Does Assessment Literacy Mean to Language Teachers?**

Numerous researchers have suggested that language teachers' initial training in assessment is insufficient and therefore their level of knowledge about assessment is inadequate. Considerable research has been conducted in East Asia into the symbiosis between teachers and assessment. Many of these studies have collected data from surveys, with the results subjected to various sophisticated statistical analyses. A notable exception to this is the study by Xu and Carless (2017) who conducted classroom observations and follow-up interviews with a teacher from a well-respected foreign language teaching university in South China. However, the focus of their study was on teacher feedback, not on teachers' beliefs and practices about assessment.

This presentation discusses a primarily qualitative study designed to investigate the practices, beliefs and attitudes towards assessment of language teachers, in order to identify their needs and requirements for further training. Base line interviews were conducted, in which most of the participants stated that assessment had not featured in initial teacher training programmes. Classroom observations then revealed that teachers were successfully deploying a range of assessment techniques. Focus group discussions indicated that participants wanted access to assessment training materials but not to engage with the theoretical underpinnings of assessment.

The presentation will end with an example of materials developed for one of the topics they stated they were most interested in, namely the Common European Framework of Reference (CEFR) and Levels. Teachers' understanding of this topic should prove especially relevant in China, where China's Standards of English (CSE) are currently under development.

**Sathena Chan**

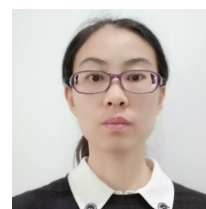
- Senior Lecturer in Language Assessment
- University of Bedfordshire
- United Kingdom

Dr Sathena Chan is a Senior Lecturer in Language Assessment at the Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire. Her research interests include integrated reading-into-writing assessment, cognitive processing of language use, task design and rating scale development. She has been leading test development and validation projects for examination boards and educational organisations in the UK and worldwide. Her recent publications include the book, *Defining Integrated Reading-into-Writing Constructs* (2018, CUP), journal articles in *Assessing Writing* (2015) and *Language Testing in Asia* (2017), and open-access research reports. She also conducts professional training and workshops on assessment literacy and statistical analyses for language testing.

**Co-presenter:****Daniel Lam**

- Postdoctoral Research Fellow
- University of Bedfordshire
- United Kingdom

Daniel Lam holds a PhD in Linguistics and English Language from the University of Edinburgh (2015), an MSc in Language Teaching from the University of Edinburgh (2010), and an MPhil in Linguistics from the University of Hong Kong (2010). Daniel's main research interest currently is in the complexity of discourse in speaking tests and assessments. His doctoral research examined the Group Interaction task in the School-based Assessment (SBA) component of the Hong Kong Diploma of Secondary Education Examination (HKDSE), combining Conversation Analysis of the test discourse with stimulated recall with student-candidates and teachers.

**Co-presenter:****Jun SHI**

- R&D Manager of English Teaching and Assessment
- Beijing Shangruitong Educational Technology Co., Ltd
- China

Jun SHI is the R&D Manager of English Teaching and Assessment. She graduated from Tsinghua University as a master majoring in English Language Testing in 2005. After that, she has been working in Beijing Normal University Press Group as English teaching and learning materials editor for ten years. Now she and her colleagues are focusing on developing a mobile English assessment platform for primary school students, aiming at creating a reliable and valid evaluating system to test primary school students' English proficiency, and improving teacher language assessment literacy.

**Abstract: The Impact of Developing an Assessment Platform on Teacher Assessment Literacy**

Although learning English is becoming more and more important in China, there is currently no compulsory exit English exams at primary level. As long as the national curriculum is followed, there is no standardised way of teaching and assessing English in the classrooms – it very much depends on the teachers. There is a need to support the teachers to develop and conduct learning-oriented assessment which can provide timely and useful feedback to improve students' learning of English. The presentation reports on a current project to develop a mobile assessment platform which can generate useful assessment data about students' learning of English in classrooms. The project also aims to establish better links among the curriculum, textbooks, and classroom assessment, and to improve teacher language assessment literacy. The project consists of the following key stages: Feasibility study, CEFR familiarisation, Test specification and Item writing guidelines development, Development of the assessment platform, Item writers training, Item production and revisions, Pilot and validation. The project involves the assessment provider, local and international consultancy teams, the British Council in China, teacher researchers and local teacher representatives trained as item writers. In the presentation, we will report some results of the pilot, followed by a discussion of the impact of this project on teachers' language assessment literacy and their teaching in classrooms. Implications for classroom-based assessment and teacher language assessment literacy will also be provided.



### Qiaozhen YAN

- PhD in Education Linguistics
- University of Auckland
- New Zealand

Qiaozhen is currently a PhD student in Educational Linguistics from the Faculty of Education and Social Work, the University of Auckland, New Zealand. She received both her B.A. in English Language and Literature and M.A. in Foreign Linguistics and Applied Linguistics from Chongqing University, China. Her research interests include classroom assessment, large-scale language testing, and the assessment of young language learners.

#### Co-presenter:



### Lawrence ZHANG

- Professor of Linguistics –in-Education
- University of Auckland
- New Zealand

Lawrence Zhang is a Professor of Linguistics-in-Education (Applied Linguistics & TESOL) and Associate Dean (AD) International Partnerships for the Faculty of Education and Social Work. His research programme spans cognitive, linguistic, sociocultural and developmental factors in reading/ biliteracy development, critical reading awareness in language education, metacognition, self-regulated learning (SRL) and reading development in L1 and L2 contexts, bilingual/biliteracy acquisition and bilingual/biliteracy education in primary and secondary schools, and learning and teaching English as a second/foreign language at university settings, the effects of self-regulated reading and writing instruction on bilingual/biliteracy development, and teacher identity and cognition in language teacher education.

#### Co-presenter:

### Helen Ramsey Dixon

- Associate Professor
- University of Auckland
- New Zealand

Helen Ramsey Dixon is an Associate Professor and Deputy Head of School in the School of Learning, Development and Professional Practice. Her research interests are in the areas of assessment for learning, self-regulation, self-

efficacy and teachers' beliefs. Feedback and self and peer assessment are of particular interest to her. How assessment for learning may play out in practice both within the compulsory schooling sector and within Higher Education are questions of interest. Her interest in Higher Education and self-efficacy has now informed a second platform of research related to academic self-efficacy as it pertains to both undergraduate and postgraduate study.

#### Abstract: Primary English Language Teachers' Conceptions and Practices in Classroom Assessment

It is refreshing that the field of language testing and assessment has recognized the importance of classroom assessment, a paradigm that is gradually coming into its own and with a focus on learning (Turner & Purpura, 2016). Although researchers in general education have sought to investigate the complexity of formative classroom assessment (e.g., McMillan, 2007; McMillan, Saefurrohman & Balinas, 2016), classroom assessment in L2 educational contexts seems incomplete. A few studies have focused on the characteristics of language classroom assessment and on the alignment of teachers' conceptions of classroom assessment and their assessment practices. Despite encouraging results from classroom assessment research at secondary and tertiary levels, young EFL learners as a special learner group have gained little influence. Against this background, this study explores EFL teachers' conceptions and practices in classroom assessment of young learners, within the Chinese testing culture context.

Specifically, this study looks into how EFL teachers conceptualize classroom assessment and implement assessment practices in the primary school context in China. It also examines how teachers' conceptions play out and align with their assessment practices using a mixed methods research design, which incorporates two phases. This presentation reports on the first phase of the research where quantitative data were collected through a survey study with 850 primary school English teachers from two major provinces in China. The preliminary findings will be discussed and implications for the interrelationships across instruction, assessment and learning for young EFL learners in the Chinese context and beyond will be considered.



### Liyan HUANG

- Director of English Education and Assessment Research Centre
- South China Normal University
- China

Dr Huang works currently as a professor of applied linguistics and PhD supervisor at the School of Foreign Studies, South China Normal University. She also works as an official trainer for Cambridge English Language Assessment. She holds a MSc from the University of Oxford and a PhD from Lancaster University, UK. Her major research interests include language testing, second language acquisition, teacher education and educational policy. She has published widely in these areas in international journals such as *Annals of Dyslexia* and *Research Notes* and domestic journals such as *Foreign Language Teaching & Research*.

#### Co-presenter:



### Ming XU

- MA Candidate
- South China Normal University
- China

Mr. Xu is an MA candidate at South China Normal University. His major research interests lie in language testing and second language acquisition.

#### Abstract: Constructing and Validating a Conceptual Framework of Assessment Literacy for Chinese EFL Teachers

Assessment plays a crucial role in the process of language teaching and learning. So developing assessment literacy is necessitated for teacher professional development and students' language learning. Assessment literacy includes the ability to collect dependable information about student achievement, and to use that information to maximize student achievement (Stiggins, 1998). Despite its significance, the discussion of the "what", "why" and "how" about assessment literacy is still in its infancy in China. A conceptual framework of assessment literacy which is in line with the contemporary assessment literature and local educational context is absent. Assessment training for both pre-service and in-service EFL teachers appears to be inadequate to some extent. With the issue of the Framework of Student Core Competencies by the Minister of Education, the assessment literacy for EFL teachers is strongly called for developing and reformulation, especially in the view of the transformation from testing culture to assessment culture. In the present study, the data derived from questionnaire surveys with 468 EFL primary and secondary EFL teachers and interviews with 44 EFL teachers from secondary and tertiary education was firstly analyzed and a conceptual framework consisting of 13 components was initially constructed. A question survey with 600 EFL teachers from primary, secondary and tertiary education was then conducted and analyzed with the statistical tool of exploratory factor analysis with the purpose of validating the proposed framework, which is expected to exert influence on the design and implementation of assessment courses or programs for teacher education.

## Breakout Session on Day 1



# Topic: Artificial Intelligence and Technology in Language Testing



**Pucheng WANG**

- PhD Candidate in the Department of Linguistics and English Language
- Lancaster University
- United Kingdom

Pucheng WANG is a PhD Candidate in the Department of Linguistics and English Language at Lancaster University. His main research interests are in the area of language testing and validation. He is particularly interested in writing assessment – specifically test-takers' thought processes during task completion. He is currently working on a PhD project which looks at Chinese test-takers' cognitive processing while completing an English reading-to-write task.

**Co-presenter:**



**Luke Harding**

- Senior Lecturer in the Department of Linguistics and English Language
- Lancaster University
- United Kingdom

**Co-presenter:**



**Baoquan LIU**

- Associate Professor in the Department of English
- Shanghai University of Finance and Economics
- China

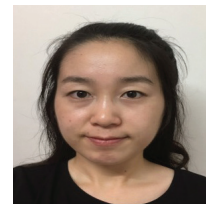
Baoquan LIU is an associate professor in the Department of English at Shanghai University of Finance and Economics. His research interests are mainly in the area of language assessment, particularly: writing, the assessment of intercultural competence, and language assessment for business purposes. His recent research projects have investigated teachers' perceptions of high-stakes assessment reforms in China.

**Abstract: Investigating Test-takers' Cognitive Processes While Completing a Reading-to-write Task**

While research into integrated tasks has been increasing in recent years (Cumming et al., 2005; Plakans & Gebriel, 2017; Weigle & Parker, 2012), there are few studies which have combined eye-tracking technology with introspective methods to explore cognitive processing during task completion. This study investigates test-takers' thought processes while completing an integrated reading-to-write task—the Test for Business English Majors-Band 8 (TBEM-8) Writing Component—through a triangulated approach in order to provide evidence of cognitive validity (O'Sullivan & Weir, 2011).

16 Chinese participants completed the TBEM-8 reading-to-write task while their eye movements were tracked using a Tobii TX300 eye-tracker. These eye traces provided data on test-takers' looking behaviour, and also formed the stimuli for a stimulated recall to elicit cognitive processes. The results showed that the participants approached the task idiosyncratically, but all engaged in a wide range of cognitive processes with varying frequencies. Overall, the most reported processes were text interpretation (36%), selecting, connecting and generating (20%) and monitoring (15%). Participants spent, on average, over a quarter of their time in reading the source texts (including the instructions). Participants who attained higher scores referred more frequently to the source texts, while the lower-scoring test-takers consulted a set of 'key concepts and expressions' for reference during writing.

Implications are discussed for the complexity of cognitive processing during integrated reading-to-write tasks and the usefulness of diverse methods for looking at test-takers' thought processes.



**Shuxin ZHANG**

- University of Chinese Academy of Sciences
- China

Shuxin ZHANG is a third-year graduate student, under the supervision of Professor Qun ZHENG. Her master's thesis explores the potential of the text analysis tool Coh-Metrix for automated assessing of coherence across different genres in speech samples. She is interested in coherence relations in academic spoken discourse, particularly those in such a specific genre as narrative.

**Abstract: Generic Features and Coherence of Prepared Speeches in Automated Evaluation**

As coherence is widely investigated in ESL written discourse, studies of coherence in Chinese students' spoken discourse across different genres are definitely rare. The present study investigates the potential of the text analysis tool Coh-Metrix for automated assessing of coherence across different genres in speech samples. For this study, 28 Coh-Metrix automated indices in 5 categories related to coherence were selected to further model human assessing of speaking proficiency of second language speech samples. Materials being analyzed were 44 TOEFL speech samples which were taken from 44 individual Chinese learners. The samples were limited to their responses to 2 independent tasks, i.e., tasks across narrative and persuasive genres. Pearson correlation and step-wise linear regression were used to determine the indices that best correlate and predict human judgments of speaking proficiency. Results showed that: (1) Automated indices predicting human judgments varied across genres. (2) Indices related to latent semantic analysis (LSA), situation model predicted 32.9% of human ratings for the narrative discourse. Indices related to causality, coreferential cohesion, and situation model might predict 50.3% of human ratings for the persuasive discourse. These findings demonstrate important implications for automated scoring in university speaking examinations. Those indices that are predictors of human assessment might be possible



features in automated scoring models. Besides, assessment of the predictive ability of automated indices related to coherence might promote the accuracy of the present scoring model, given the fact that the present model covers exclusively features of pronunciation, fluency, grammar and vocabulary dimension.



### Dajian CHEN

- Doctoral Candidate in the School of International Studies
- Zhejiang University
- China

His research top of interest is language testing. His research topic of interest is language testing, the focus of his doctoral dissertation is video-based listening testing.

#### **Abstract:** Investigating Video-based Listening Test: the Role of Visual Information in Comprehension Process

Along with the acoustic information, visual information do exist in most real-life listening comprehension activities. Thus implementing videos in listening comprehension test is considered more authentic than the traditional audio input. Many researches have focused on the topic of Video-based listening test over the past few decades, yet the role of visual information in comprehension process is still unclear due to the conflicting results of previous studies. This study attempts to investigate the role of visual information at two different comprehension levels (local & global) by grouping test items into two types (detail questions & gist questions). Repeated measures design with order and materials counter-balanced was adopted to test the effect of visual information on different comprehension levels. 313 first-year college students participated in this listening test which contains three video lecture recordings and three audio lecture recordings attached with 26 comprehension questions in total. Preliminary findings indicate that: 1) in general, video listening test performance are not significantly different from audio listening test; 2) for gist questions, video listening comprehension scores are significantly higher than audio listening comprehension; for detail questions, however, video listening test scores are significantly lower than audio listening test scores. The qualitative methodology "listen & recall" will be used to further investigate the listener's comprehension process with & without the help of visual information.



### Richard Spiby

- Test Development Researcher, Receptive Skills
- British Council
- Unite Kingdom

Richard Spiby has been the Test Development Researcher for the receptive skills (reading and listening, together with grammar and vocabulary) since June 2016. His main responsibilities involve analysing operational data and revising and developing the receptive skills components of the Aptis test.

Richard has previously worked in the UK and Turkey, mainly in the university sector, in test production, management and research. His particular areas of interest are cognitive processing in language assessment, strategy use in reading and listening and methods of testing vocabulary.

#### **Co-presenter:**



### Franz Holzknrecht

- Lecturer and Researcher
- University of Innsbruck
- Austria

Franz Holzknrecht holds an MA in Language Testing from Lancaster University and an MA in English Language Teaching and Sports Education from the University of Innsbruck. He works as a lecturer and researcher at the University of Innsbruck's Language Testing Research Group. His research interests are mainly in language assessment, particularly in listening assessment, writing assessment, and the use of eye-tracking. Franz has presented his research at numerous international conferences and has published his work in Papers in Language Testing and Assessment and in several book chapters.

#### **Abstract:** A Framework for Innovation: Revising the Aptis Listening Test

The revision process is an important but sometimes poorly defined aspect of the test development cycle. This presentation describes the model of test revision used for the Aptis listening test. At its launch in 2012, the British Council's Aptis testing system was designed to be a flexible and dynamic system. Since then, it has expanded globally in the fields of education and beyond. The Assessment Research Group,

which seeks to provide validation evidence for the evolving uses and interpretations of the Aptis system, aims to review and revise the test in an ongoing 5-year cycle. As such, it has created a multidimensional framework of reference to guide the revision process. This involves identifying and balancing the requirements of important stakeholders, and establishing a set of key principles to inform potential innovations. In determining the scope of revisions, a large body of research from a range of sources is utilised, drawing on internal validation carried out by the Assessment Research Group itself, projects undertaken as part of the Assessment Research Awards and Grants scheme, and research commissioned on particular aspects of the listening test and construct, including analyses of eye-tracking data on the cognitive validity of the Aptis listening test done at the University of Innsbruck, Austria. As a result, the proposed revised test is to contain a greater variety of task types, with particular attention to broadening the construct assessed at B2 level of the CEFR. The revision framework is evaluated and implications for future test revision cycles are discussed.

## Breakout Session on Day 1



# Topic: Foreign Language Testing for Admissions and Exit Tests in Education Systems



## Ni CHEN

- Associate Researcher and Vice Dean of the School of Education and Training
- Renmin University of China
- China

Dr Chen worked for years in Renmin University of China as associate researcher and vice dean of the school of Education and Training. Her publication includes the book *Influencing Factors of Chinese Students' Cross-Border Mobility for Higher Education* (2014) and other articles in the field of cross border education and cross culture communication.

### Co-presenter:



## Michael Burns

- Professor
- Memorial University
- Canada

Professor Burns recently retired from Canada's Memorial University after completing 35 years in senior academic positions, primarily in economics but latterly in international business, at leading universities worldwide. Over his career Professor Burns achieved distinction through academic and professional leadership, as a widely published researcher in international journals (including the prestigious *American Economic Review* and *Journal of Political Economy*), as policy adviser to various government agencies and through the development and management of collaborative international education projects in over 30 countries. He retains a role as CEO of International Educational Management & Development (IEMD), an international consultancy based in Australia.

### Abstract: The Effectiveness of Admission Criteria in Predicting Collaborative 2+2 Program Success of Chinese Students

Collaborative 2+2 programs have become one of the most popular pathways enabling Chinese students to proceed overseas to complete a western undergraduate degree. This paper examines the effectiveness of admission criteria including English Entrance Exam (EEE), English Entrance Interview (EEI), Gaokao total (GKT), Gaokao English (GKE), and Gaokao Maths (GKM) in predicting the success of academic and English language preparation for Chinese students in 2+2 international programs. Included are the admissions and enrollment data for 242 Chinese students admitted to a 2+2 program between a Chinese university and a North American university over the period 2006-2009. Regression analysis was applied to the data. Results show that all the admission criteria are correlated with IELTS and the academic outcomes of Chinese students two years after the admission when they are going to transfer to the North American campus. EEE is the most predictive of the IELTS score at the end of the first two years where EEE, EEI, GKE and GKM together explained almost 50% of the variance of the IELTS score. GKE, GKM and EEE are the most important predictors for the students' 2 year academic average and these three admission measures explained almost 30% of the variance in the 2 year academic average. These findings offer support as to the effectiveness of admissions criteria of 2+2 international programs and also underline the important role that English language assessment plays in globalized education system such as the double campus international collaboration program.



## Adrian Raper

- Technical Director
- ClarityEnglish
- Hong Kong

Dr Raper is the technical director at Clarity English and is responsible for product development and teacher support. For more than

20 years he has led the teams developing some of the most popular English teaching programs in the world. Throughout that time he has worked with teachers in workshops, training sessions and conferences to help integrate ICT into the classroom and learning space.

He graduated with a PhD in Artificial Intelligence from Southampton University in the UK and is currently developing software platforms for delivering language tests using mobile devices.

### Abstract: Splitting a Placement Test CEFR Output by Skill

A placement test is usually quick to take and often completed without writing and speaking sections to make it easy to mark by computer. An overall CEFR level, or other overall score, is given based on reading, listening and language elements. Is this the most useful information to give an English teacher tasked with raising the student's performance to a required level?

Clarity has a newly released placement test, and one of the main requests has been to split the result by skill.

What do institutions plan to do with this information? Does it turn the test into a diagnostic one with the results revealing different teaching paths available to the student? Or is the extra information going to be irrelevant and lost the moment the student joins a course?

This paper will present the results of a survey asking those institutions who request this skill split what will be done with the answers — as well as data from their test-takers showing how the split works in practice.



## Jeanne O'Connell

- Senior Tutor in Assessment (Centre for English Language Education)
- The University of Nottingham Ningbo China
- China

Jeanne is responsible for assessment in the Centre for English Language Education (CELE) at University of Nottingham Ningbo China (UNNC), which currently supports over 6,000 students on its Preliminary Year, Pre-Sessional and Academic Support Unit programmes. Before joining UNNC, she worked in University College Dublin (UCD), overseeing the development of English language tests for various levels and purposes and also conducting associated research, particularly in the areas of reporting test results and the Common European Framework of Reference (CEFR).



## Li LIANG

- Lecturer
- Guangdong University of Foreign Studies
- China

Cantonese, working in GDUFS, specializes in socio-cultural theory research and language education and assessment.

### **Abstract: Tri-unity scheme Test: A collaborative alternative entry pathway**

This paper describes a new test providing an alternative entry pathway into university in China. It first gives background to the HE teaching context in China and the goals of the test. More than 9 million students take Gaokao (The National Higher Education Entrance Examination) each year and a student's score is essential in determining university admission. The test described in this paper, known as the Tri-Unity Scheme Test (TUST), evaluates potential domestic undergraduate candidates not only on the basis of their Gaokao scores, but also considers the high school leaving exams (Huikao) and university screening.

The paper then goes on to describe the assessment process in detail as administered at a Sino-foreign university in China, where, for the past 3 years, TUST has replaced the direct entry admissions examination. University staff across various departments work together to plan, develop, deliver, mark and report the results of the test. This collaborative cross-departmental approach to TUST presents challenges specifically in the speaking component of the test, which takes the form of a group discussion (VanMoere, 2006). Four of the five raters of this test component are lecturers and professors with little or no background in language assessment. The paper discusses how these raters perceive and interpret student performance, and then describes the standardisation procedure currently used to ensure an acceptable level of inter-rater reliability.

The paper concludes with an outline of how findings from the current set of standardisations will feed into a revised standardisation procedure in future years. who taught the course of College English and 87 students who had taken the EPT participated in the study.

### **Abstract: A Socio-cultural Perspective: EFL Teachers' Perception in the Context of Computer-based Large-scale Oral English Test**

Within the worldwide trend of computer application in the educational arena, there have been few attempts to examine the teachers' perception of face-to-computer high-stakes matriculation language testing although college acceptance issues have been one of the top priorities all around the world. This article reports a case study on the face-to-computer matriculation EFL oral proficiency test reform in Guangdong(Canton), China. The goal of this research is to investigate the high school EFL teachers' perception of the testing reform on teaching and learning and provide the face-to-computer language testing designers some practical implications and suggestions for future development. Two hundred and twenty-three high school teachers in twelve high schools took a survey with closed and open-ended questions. It is found that the testing reform received a major acknowledgment for its necessity and impacts on teaching. Six principal obstacles in this reform are identified. Furthermore, change in tools alone will not suffice in our educational reforms assisted by computers. This article discusses the often-overlooked construct--culture at macro and micro levels in the testing context. The impact of cultural artifacts on the face-to-computer testing is modeled by connecting the ideas of culture to the framework of activity theory.

## Breakout Session on Day 2



# Topic: Assessment Literacy for Language Teachers



### Joseph Davies

- Business Development Lead
- University of Central Lancashire's School of Language and Global Studies
- China

Joseph Davies is Business Development Lead for The University of Central Lancashire's School of Language and Global Studies in China and also lectures on their franchised BA (Hons) International Business Communication Programme in Shanghai. Over the past eight years Joe has designed, managed and delivered a range of academic and professional English Language and Communication courses across China. Joe regularly attends and presents original research at International academic conferences and has had previous research published. His research interests include second language classroom participation, second language willingness to communicate, classroom assessment and socio-pragmatics.

#### Co-presenter:



### Laura Davies

- Programme Leader
- University of Central Lancashire's franchised BC(Hons) International Business Communication Program
- China

Laura Davies is the Programme Leader for the University of Central Lancashire's franchised BA(Hons) International Business Communication Programme based in Shanghai, P.R. China. Laura has been managing and teaching within various English Language Teaching contexts throughout China for over eight years. She regularly attends and presents original research at International academic conferences. Her current

research interests include employability within ELT curricula, second language assessment and validity, socio-pragmatics and second language acquisition.

#### Co-presenter:

### Yueting XU

- Guangdong University of Foreign Studies
- China

#### Abstract: Grading class participation: Perception and practice of ELTs in China

As a form of active learning, class participation has become a common requirement of many university courses, including English language courses in China. The benefits and proliferation of graded class participation notwithstanding, our informal conversations with university English instructors in China suggest that most of them determine grades of class participation impressionistically. Given that empirical evidence about how English language instructors grade class participation in higher education (HE) contexts in China is generally lacking, this study addresses this gap by exploring teachers' perceptions and practice of grading class participation through two research projects. In particular, two research questions are posed:

RQ1: What are the teachers' perceptions of assessing class participation (ACP)?

RQ2: What are the teachers' practices of ACP?

Project 1 is a survey study in which 41 English-language teachers (both native and non-native English speakers) participated in an online survey, which was comprised of three open-ended questions and nine closed-ended questions to elicit both qualitative and quantitative data.

Project 2 consists of three case studies of university English teachers (Linda, Rosa, and May, all pseudonyms) working in three Chinese universities of different tiers. Based upon the results from Project 1 about teachers' perceptions of ACP, this project sought an in-depth understanding about what teachers think, and actually do, when assessing class participation.

Based on findings of the teachers' perceptions and practice of ACP, discussions are centered on challenges for valid assessment of class participation, potentially effective solutions for ACP and implications for teacher assessment literacy.



### Jiming ZHOU

- English Language Teacher
- Fudan University
- China

Jiming ZHOU is an English Language teacher at Fudan University, China. Her research interests include assessment and learning, educational innovation and English language teaching. Her works about assessment for learning in Chinese college English classrooms have been published in *Assessment and Evaluation in Higher Education*.

#### Abstract: Assessing Oral Presentations in English Classrooms: Teachers' and Students' Criteria

English oral presentation is a common classroom assessment strategy in multiple disciplines in higher education. The ability to present a topic via English in a clear and eloquent way is one of the most valuable skills aspired by higher education and employment markets. However, teachers may understand and use the assessment criteria in very different ways. There is a dearth of research exploring how in-service teachers develop their classroom assessment literacy in terms of developing and using the oral presentation rubrics.

This paper draws upon the principles of grounded theory. The datasets of this study include: different versions of oral presentation rubrics collected from college English teachers in a university in Shanghai; individual interviews with five teachers about how they understood and developed their assessment criteria; and the field notes during the observation of how the five teachers evaluated their students' oral presentations. Three classes of students (n=90) were invited to write down their peer feedback instantly after their peers' English oral presentations throughout a whole semester. The corpus of students' feedback was analysed and students' perceived criteria were elicited.

This paper compares teachers' and students' self-generated criteria for English oral presentations. The pedagogical implication is that teachers



need to be aware of the teacher-student perceptual gap with regard to the targeted competence, and exemplars of presentations help students to understand the abstract criteria. The paper also proposes the constituents of teacher assessment literacy in terms of oral presentation, and connects it with the broader teacher assessment literacy.



### Philip Horne

- Assessment Consultant
- British Council
- China

Philip Horne works as an Assessment Consultant for British Council China. He has worked on numerous assessment-related projects in the Middle East, East Asia and South-East Asia. These relate to exam preparation, assessment literacy and item writing. He is pursuing an MA in Language Testing through Lancaster University where his research interests include assessing pronunciation, Global English and Language Assessment Literacy.

#### Abstract: A Validity Case for Procedural Assessment Literacy for EFL Teachers in China

Pill and Harding (2013) published a cline identifying five stages of Language Assessment Literacy (LAL) from total inexperience (1) to complete expert (5). O'Sullivan (2016) suggests most EFL practitioners, including teachers, are around the third (functional) stage, meaning key concepts are understood in broad strokes, but moving all professionals towards more principled understanding (the fourth, or procedural stage) of LAL is both an important and neglected aspect of validation.

This session is based upon a replication study of Fulcher (2012), who conducted a self-evaluative questionnaire on perceived awareness of LAL among various practitioners. My study replicated that questionnaire, conducted in this instance among 50 EFL teachers and high-stakes examiners in China. It also addresses the limitation in the original study of only using self-evaluative data, by additionally asking participants to write brief definitions of terminology such as "construct validity". The results were analysed using summary item statistics and achieved a Cronbach's alpha of 88.

This paper empirically consolidates the aforementioned conclusions of O'Sullivan (2016), by demonstrating i) that many EFL professionals are indeed at the functional stage, ii) there are key areas of LAL within which practitioners are not even aware of the nature of the knowledge gap, and iii) that addressing these gaps through LAL training would lead to an overall increase

in test validity, including for instance improved classroom testing, or better communication with stakeholders. EFL in China, as a "testing-oriented" culture, is used as a context to suggest practical frameworks for LAL development.



### Robin Skipsey

- Academic Manager, English for Education Systems Japan
- British Council
- Japan

Robin Skipsey is an academic manager responsible for the design and delivery of a major national teacher-training scheme on behalf of the Japanese Ministry of Education. This scheme aims to help Japanese school teachers of English to foster productive skills (speaking and writing) and teach students to use English in an interactive and communicative way, rather than simply as a body of linguistic knowledge to be memorized for exam purposes. Currently he is engaged in a project to increase teachers' awareness of how to assess the productive skills.

#### Co-presenter:



### Judith Fairbairn

- Productive Skills Testing Researcher
- British Council
- United Kingdom

Judit Fairbairn is a language-testing researcher at the British Council who leads on speaking and writing testing projects, including developing and revising British Council tests and marking scales. Her areas of expertise include: rating scales, online rater training, rater reliability, test data statistical analysis, language assessment literacy training for teachers and equality, diversity and inclusion in testing. She has developed rater training, standardisation and monitoring systems for Aptis and IELTS.

#### Co-presenter:



### Gordon Allan

- Trainer of English
- British Council
- Japan

Gordon has been working for British Council Japan since 2003, and has extensive experience teaching junior high school and university classes. He now works in teacher training, helping to plan and deliver the Leaders of English Education Project (LEEP), a cascade training programme for teachers of English at elementary, junior high and high school level. Gordon is CELTA and DELTA qualified, and has recently completed an MA in Language Testing with Lancaster University. His dissertation research focused on the teaching and testing of speaking in Japanese high schools, and the potential impact of testing speaking in university entrance exams.

#### Abstract: Teaching How to Rate the Productive Skills: Case Study from Japan

Over the past few decades, the Japanese government has recognised the need for English language skills and there has been an increasing focus on the teaching and assessment of the language at all levels of education. There has been a tendency, however, to neglect the development of spoken and written English. This is most likely due to two reasons: the lack of a speaking and writing component in the high-stakes university entrance exams and, secondly, the lack of expertise in teaching and testing speaking and writing skills.

Educational reforms in Japan demonstrate recognition of this lack of productive skills learning, as reflected in the MEXT decision to mandate the use of commercial 4-skills tests for university entrance from 2022, and a recent recommendation from the Japanese Language Testing Association to improve teachers' assessment literacy in order to ensure adequate washback from this change to the University Entrance exam system.

Building on the teacher training work already in progress, the British Council is addressing these language assessment literacy training needs with the development of a programme which can train teachers on how to assess the productive skills. This presentation will describe the methodology and delivery model for the proposed training, the data collection process, and some of the findings so far in setting up this programme.



### Qiuxian CHEN

- Associate Professor
- Shanxi University
- China

CHEN Qiuxian is an Associate Professor in Shanxi University, China. She received her PhD in Queensland University of Technology, Australia. Her expertise is in the field of English as Foreign Language teaching and assessment, and assessment policy.

#### Co-presenter:



### Jiajin ZHANG

- Associate Professor
- Shanxi University
- China

Jiajin ZHANG is an Associate Professor in Shanxi University. She has taught College English for an extended period of time and developed an expertise in English pedagogy over the years.

#### Abstract: Teacher & the Implementation of Formative Assessment in Local Settings: A Multiple Case Study

This paper presents findings of a multiple-cases study which examined, through a sociocultural perspective, the implementation of formative assessment in eight universities located in a developing mid-western province of China. Particularly, practitioners' voice is highlighted, which focuses on the challenges that these local universities were faced up with to bring into practice the formative assessment policy initiated via College English Curriculum Requirements (CECR)(Ministry of Education, 2007).

Data set includes interview data which were conducted to English teachers (focus group) and senior administrators (individual) from the eight universities respectively. Analysis by way of constant comparison and grounded theory reveals that changes did happen to the College English assessment regime in response to the CECR in some of the universities, but not in others. In addition, despite some teachers' willingness to enact formative assessment in their classrooms, they were hindered to do so by multiple problems such as logistic limitations, misinterpretation of formative assessment at the institutional level, limited assessment literacy, lack of proper training and support, and etc.

These issues in a way indicates that, apart from the social, cultural, and historical issues that formative assessment is met with in China overall, these less developed regions are in an especially urgent need of support in terms of teacher assessment literacy so as to achieve the policy intent to a satisfactory extent. This is particularly important for a balanced development in the area of education throughout the country.



### Michelle Raquel

- Senior Lecturer
- Centre of Applied English Studies
- University of Hong Kong
- Hong Kong

Dr Michelle Raquel is a Senior Lecturer at the Centre of Applied English Studies, University of Hong Kong. She has been involved in the design and administration of several large-scale assessments for tertiary institutions in Hong Kong and specializes in the use of Rasch measurement for language testing and assessment. She has also been involved in the research and development of a computerised diagnostic assessment of academic English and dynamic assessment.

#### Co-presenter:



### Simon Boynton

- Lecturer
- Centre for Applied English Studies
- University of Hong Kong
- Hong Kong

Simon has worked in Poland, Spain, Argentina, Thailand, Vietnam, Hong Kong, and the UK as a teacher, manager, and Cambridge ESOL teacher trainer with organizations as International House, the British Council, the Study Group (ILA Vietnam), the University of Leeds, and City University of Hong Kong. He currently works in the Centre for Applied English Studies at the University of Hong Kong. He was the Programme Coordinator for Science and is currently the Programme Coordinator for Medicine, responsible for 6 undergraduate courses and 1 postgraduate course. His interests include teacher training, materials design, assessment, and undergraduate ESP (technical writing).

#### Co-presenter:



### Bernadette Wo

- Lecturer
- Centre for Applied English Studies
- University of Hong Kong
- Hong Kong

Bernadette resided in the U.S. for 31 years and completed an Ed.D. in Curriculum & Teaching at Teachers College, Columbia University. At the Centre for Applied English Studies at HKU, she has taught English to Arts, Business, Dentistry, Education, Engineering, Medical, Science, and Social Science students. She is a member of the Centre's Assessment SIG. Other research interests include qualitative research and research on writing. Bernadette's teaching experiences also include elementary school and adult education.

#### Abstract: Assessment Literacy Profile of Tertiary Level Course Coordinators

Assessment literacy has been identified as a critical factor in improving student learning (William & Thompson, 2007; Hattie, 2009; Popham, 2009; Hill, 2017) and thus the construct of assessment literacy required by various stakeholders has been of keen interest recently. In particular, Taylor (2013) outlined the difference between the assessment literacy profile of a university language teacher and a university administrator. However, her framework does not include a key stakeholder in a tertiary language context – the course coordinator. Within a tertiary language context, course coordinators are considered middle managers where they hold teaching and administrative duties; apart from teaching, they are also responsible for preparation of course materials and assessments, training and management of course teachers, handling course level student inquiries, grade moderation and decision making, and preparation of internal and external reports. Usually, they hold this position due to seniority, years of experience, and/or qualification.

This presentation outlines the results of a case study that explores the aspects of the assessment literacy construct required by course coordinators. Course coordinators working in a Hong Kong tertiary institution were asked to complete an assessment literacy survey, review a language assessment code of practice document, and participate in in-depth semi-structured interviews that solicit their suggestions on content and implementation of the code of practice document and probe their assessment literacy knowledge. The presentation concludes with the assessment literacy profile of a university course coordinator that could help in the hiring and professional development of current tertiary level course coordinators.



## Breakout Session on Day 2



# Topic: Measuring Teachers' Language Teaching Proficiency



## Nguyen Le Duy Hai

- Lecturer
- Binh Duong University
- Vietnam

Nguyen Le Duy Hai is currently working as a lecturer at Binh Duong University. Nguyen holds a Bachelor's degree in Teaching English as a Foreign Language (TEFL) from Ho Chi Minh City University of Education and a Certificate in Teaching English to Speakers of Other Languages (CELTA). His research interests include language assessment and second language acquisition.

### Abstract: Measuring Teachers' English Proficiency: Reviewing a Project in Vietnam

Increase in teachers' English language capacity is universally assumed to lead to improved teaching quality in English-medium classrooms. This reasoning is reflected in a large-scale project initiated by Ho Chi Minh city (HCMC) DOET in 2011 as a part of Vietnam's National Foreign Language 2020 Project to measure local high school teachers of English language. However, this project has received mixed reactions. This research aims to explore Vietnam high school teachers' perspectives on the above-mentioned project and thereby to identify its impacts on those teachers. This qualitative research involves 2 stages. Firstly, a questionnaire was sent to 83 high school teachers of English in HCMC to collect their opinions about the HCMC DOET's project. Secondly, the results of this survey were analysed by applying Bachman and Palmer's (1996) framework of test usefulness. The preliminary data reveals that this ambitious project failed to foster Vietnamese teachers' command of English. Furthermore, the use of proficiency tests such as IELTS and Cambridge English exams is viewed to cause negative washback and impacts on teachers' professional development and be irrelevant to their classroom practices due to the lack of construct English-for-Teaching coined by Freeman, Katz, Gomez and Burns (2015). These tentative findings suggest a paramount need to designate a more appropriate test for assessing Vietnamese teachers' English language proficiency and "Aptis

for Teachers" is possibly a promising test to fulfil the ultimate objective of promoting English language competence of those teachers.



## Xin QU

- Professor of Applied Linguistics,
- Beijing Foreign Studies University
- China

Dr Qu, Professor of Applied Linguistics, is now serving as a postdoc researcher in the National Research Centre for Foreign Language Education at Beijing Foreign Studies University, China. Her main research and publication interests are in language assessment, language education, and relationships between assessment, teaching and learning. Xin has conducted research into performance assessment, test validation, and test-data management and analysis. She is currently working (supervised academically by Prof. Baocheng Han) on a major project of ELT Certificate development and validation for EFL teacher certifying in China, which is funded by National Planning Office of Philosophy and Social Science.

### Co-presenter:



## Baocheng HAN

- Deputy Director of National Research Centre for Foreign Language Education
- Beijing Foreign Studies University
- China

Professor HAN is Deputy Director of the National Research Centre for Foreign Language Education at Beijing Foreign Studies University, China. He also serves on the Executive Board

of China Language Assessment (CLA) and the China Council for Self-taught Higher Education Examinations. As a professor of Applied Linguistics, his main interests lie in the field of language assessment and language education. He has specific expertise in the nature of language ability, and extensive experience of working on test design and development projects. Professor Han has published widely in the field, and advised institutions on aspects of language assessment.

### Abstract: Validation of ELT Certificate-LORT: A Process Suited for Many-Facet Rasch Modeling

The purpose of this study was to validate ELT Certificate Lesson Observation and Report Task (ELTC-LORT), developed by China Language Assessment in the design of performance-based testing for China's EFL teachers' certifying. The validity of task construct and rating rubric were supported using Many-Facet Rasch Model. Participants (N=40) were provided with a ten-minute video excerpt from a real EFL lesson, and required to deliver a three-minute report on the teacher's performance. Two raters graded the records of the candidates' reports using scales of the rating rubric developed to measure EFL teacher candidates' oral English proficiency and ability to analyze and evaluate teaching.

Many-Facet Rasch analysis demonstrated a successful estimation, with Resd (.00) and StRes (.01) as expected and 57.23% variance explained by the measures. There existed a noticeable spread among the participants (-4.82 to 5.70 logits) and their traits (-.70 to 1.16 logits), proving the task functioned well in measuring candidates' performance and reflecting the difference of their ability. The raters were found to have good internal self-consistency (Infit MnSq 1.00, ZStd .00 for rater 1; .98, -.10 for rater 2), but not the same leniency (separation 5.78, fixed (all same) chi-square 34.4, significance .00). Meanwhile, the rating scales worked well, with the average measures advancing much in line with Rasch expectations.

The findings provide evidence of validity for ELTC-LORT, and also implications for further research and practice of the task. We hope the study could provide insight into measuring EFL teachers' competencies and encouraging their professional development with valid measurement.



### Sayyed Mohammad Alavi

- Professor of Applied Linguistics
- University of Tehran
- Iran

Dr. Sayyed Mohammad Alavi received his PhD from Lancaster University in 1997. Currently, he is a professor of applied linguistics in the University of Tehran, Islamic Republic of IRAN. He is teaching language teaching, language testing and assessment, and materials development to BA, MA, and PhD students. His research interests include language testing/assessment, task-based language teaching and assessment, materials development, and teacher education. He published and presented in both national and international circles. His recent papers were published in Language testing Journal, Language Assessment Quarterly, System, RELC Journal, Educational Assessment, and Language Learning Journal.

formed the bases for developing a unique listening comprehension scenario test that aimed to assess listening comprehension teachers' diagnostic competences. Thirdly, to establish the priori validity of the test, 10 experienced language teachers were asked to rate the scenario test items in terms of their occurrences in listening comprehension practices. Finally, to detect the listening comprehension diagnostic competence in pre, while, and post task cycles, the listening comprehension scenario test was administered to 65 language teachers. Both qualitative and quantitative techniques were employed to analyze the data sets. The results showed (1) a significant difference in the frequency of diagnostic behaviours of teachers at any of these stages, (2) listening comprehension teachers need various diagnostic competences for pre, while, and post task cycles of teaching listening comprehension. This study suggests language teacher educators to design appropriate diagnostic competence models for any of the cycles of listening comprehension teaching. (246)

#### Co-presenter:



### Tina Mirmotahari

- Teacher Trainer
- University of Tehran
- Iran

Tina Mirmotahari is an MA graduate from University of Tehran (Alborz campus). Her areas of interest include teacher education and teacher assessment. She has been teaching English for 12 years and she is currently a teacher trainer.

#### **Abstract: Language Teachers' Listening Comprehension Diagnostic Competence in TBLT**

This study examined diagnostic competence of listening comprehension teachers for pre-task, while-task, and post-task cycles of Task Based Language Teaching of listening comprehension. To achieve this, firstly, four listening comprehension courses of four teachers with almost ten years of teaching and three years of teacher training experiences were observed and 55 hours of their practices were video-taped. Secondly, two qualified language researchers were asked to analyze the video-taped data and extract the points related to the diagnostic competence of teaching listening comprehension. These points and the suggestions made in the literature about teaching and assessing listening comprehension

## Breakout Session on Day 2



## Topic: Language Proficiency Frameworks

**Judith Fairbairn**

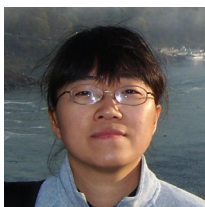
- Productive Skills Testing Researcher
- British Council
- United Kingdom

**Abstract: Developing Rating Scales for the Productive Skills Using the CEFR**

Rating scales are normally used to mark the productive speaking and writing skills and represent the test construct, or the components of the written or spoken performance to be measured. The difficulty lies in balancing the desire to test everything using detailed descriptors for all features of the performance against the practical aspects of rating. The need for reliable and accurate scoring that adequately reflects the performance of the test taker means that rating scales must be carefully considered.

An important element of the Aptis test design has been the explicit incorporation of the Common European Framework of Reference (CEFR) into the test system. Tasks are targeted and results are reported at CEFR levels. For the productive skills, the performance descriptors from the CEFR form a central part of the rating scale design.

This presentation describes a project to revise the Aptis writing and speaking rating scales. The research methodology, the steps taken and the data analysis is discussed. Benefits and issues encountered in applying the CEFR to the marking of the Aptis speaking and writing tasks are highlighted.

**Hang Li**

- Associate Professor
- School of International Studies
- Zhejiang University
- China

Professor Li has been teaching college-level English to EFL learners since 2002. She got her PhD in 2012, and her thesis title is “*Effects of Rater-scale Interaction on EFL Essay Rating Outcomes and Processes*”. She has published several research articles in peer-reviewed journals, including “*A comparison of EFL raters’ essay rating processes across two types of rating scales*” (Language Assessment Quarterly, 2015, April).

**Abstract: Developing a Rating Scale for the Cognitive Diagnostic Assessment of Chinese College EFL Learners’ Writing Proficiency**

By identifying deficiencies, problems or areas where students would need help with (Alderson, 2005), diagnostic language assessment can be of great help to second language learners. However, despite renewed interest in diagnostic language assessment, little work has been done on diagnostic assessment of writing (Knoch, 2011). The only instrument developed for the diagnostic assessment of ESL writing proficiency was by Kim (2011), who developed the Empirically-derived Descriptor-based Diagnostic (EDD) checklist from essay raters’ think-aloud protocols and applied Reduced Reparameterized Unified Model (Reduced RUM), a type of cognitive diagnostic assessment (CDA) model, on the ratings thus acquired. Xie (2017) validated a revised version of the checklist, using again Reduced RUM.

To further study on this particular approach to writing assessment, and to create a rating scale to gauge Chinese college-level EFL learners’ writing proficiency, the present study tried to create a new rating scale in the checklist form. Based on theoretical models concerning writing processes (Deane, et. al, 2008), and analysis of student essays, a scale comprising a total of 21 items was created, one third of which are gauged by statistics provided by computational linguistic tools such as Coh-Metrix, Lexical Complexity Analyzer (Lu, 2012), and the L2 Syntactic Complexity Analyzer (Lu, 2010).

To validate this checklist, a FACETS analysis was done to investigate the reliability of the 14 human-rated items; meanwhile, correlations between automated scored items and the overall writing ability were also looked into. Results indicate that the checklist does show promise as a valid instrument for diagnostic purposes.

**Shaoyan ZOU**

- Lecturer
- Qingdao Agricultural University
- China

Shaoyan ZOU is now a lecturer at Qingdao Agricultural University, Qingdao. She is about to obtain a PhD degree in Applied linguistics at Shanghai Jiao Tong University, Shanghai under the supervision of Professor Yan JIN. Her research interest mainly lies in Language Testing and Evaluation.

**Co-presenter:****Jun WANG**

- Doctoral student
- Shanghai Jiao Tong University
- China

Jun WANG is now a doctoral student of Shanghai Jiao Tong University, Shanghai. She is under the supervision of Professor Yan JIN. Her research interest is Language Testing and Evaluation.

**Abstract: Rating Scale Impact on the Scoring of CET4 Writing: A Mixed-Method Study**

In assessing EFL writing performances, two types of rating scales were commonly used by raters: holistic rating scale and analytic rating scale. Given the significant differences between the two types of scales, their assumptions about writing components and development, as well as the implications for essay marking processes and scores (see Barkaoui, 2007a; Goulden 1992, 1994; Weigle, 2002), there have long been controversies about which type is more superior (Hamp-Lyons, 1995; Weigle, 2002; White, 1984, 1995). However, to date quite a few empirical

studies have been conducted to compare the impact of the two types of scales.

In view of this lack of research, this study adopted a mixed-method approach to compare the impact of two scales on the rating of the CET4 writing scripts: one a priori holistic rating scale designed intuitively, and one analytic rating scale developed empirically. The quantitative comparison involved 21 trained raters applying the two rating scales consecutively to the rating of 30 CET4 writing scripts. Whereas, the qualitative comparison was undertaken through in-depth interviews with 10 raters.

Rasch model analysis of the quantitative data indicated that 1) the two scales were both effective in terms of scale discrimination and rating reliability; however, 2) the analytic rating scale was more effective in controlling rating variations as the raters' behaviors were more consistent both at group level and at individual level. Rater interviews also revealed a preference for the analytic rating scale. Finally, the findings were discussed in terms of their implications for rating scale development and validation in EFL writing.



### Mohd Sallehudin Abd Aziz

- Associate Professor of School of Language Studies and Linguistics
- Universiti Kebangsaan Malaysia
- Malaysia

Mohd Sallehudin Abd Aziz is Associate Professor at the School of Language Studies and Linguistics, Faculty of Social Sciences and Humanities, UKM. He started his career as a language teacher in 1986 after graduating from Western Kentucky and Northern Arizona University, USA. He obtained his PhD in Language Testing and Evaluation from Universiti Kebangsaan Malaysia. His main research interests include Language Testing & Evaluation and English For Academic Purposes. He was Co-editor of GEMA and currently a co-editor (2013-2014) for Akademika -the Journal of Southeast Asia Social Sciences and Humanities.

#### Co-presenter:



### Nurul Farehah Mohamad Uri

- Lecturer
- Universiti Kuala Lumpur British Malaysian Institute
- Malaysia

Nurul Farehah Mohamad Uri is a lecturer at Universiti Kuala Lumpur British Malaysian Institute (UniKL BMI). She started her career as a language teacher in 2010 after graduating from Universiti Teknologi MARA (UiTM) with a TESL degree. She obtained her M.A (English Language studies from Universiti Kebangsaan Malaysia in 2013. Currently, she is pursuing her PhD in Language Testing and Evaluation from UKM. Her doctoral research focuses on the implementation of CEFR in Malaysia. Her main research interests include Language Testing & Evaluation and English For Academic Purposes.

#### Abstract: CEFR in Malaysia: Current Issues and Challenges in the Implementation of the Framework

The Common European Framework of Reference or CEFR is now acknowledged internationally as the standard language proficiency framework to adopt. The framework which was first intended to serve as a shared language framework in the EU to regulate its language education system across several countries within the union has been embraced countries outside Europe such as Vietnam, Japan, and Indonesia. The framework has also made its way to Malaysia and was officially introduced in 2013. Recent important documents such as Malaysian Education Blueprint 2013 -2025 and English Language Education Reforms 2015-2025 have shown that the government has agreed not only to incorporate and align the framework into the present education system but accelerate its implementation. Unfortunately, there are a number of issues that need to be dealt with first before CEFR is fully implemented in the country. This paper therefore intends to highlight some of these important issues that Malaysian ministry of education faces in implementing CEFR. In order to understand the issues thoroughly, this paper will first provide some historical background of the Malaysian education system and relevant issues associated to it. Next, some studies on CEFR that have been conducted in Europe, Asia, and in Malaysia will be thoroughly discussed. Then, it will highlight some of the challenges that may hinder the implementation of the framework. Some possible solutions are also deliberated. This paper will also suggests the need to critically evaluate these studies before the framework is fully incorporated into the education system.

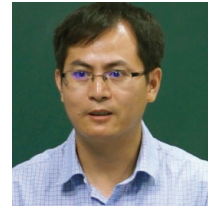


### Wen ZHAO

- Professor of English
- Northeastern University
- China

Dr Zhao is professor of English at Northeastern University. Her research efforts have focused on curriculum and instruction, corpus linguistics, and vocabulary study. She is currently the dean of Foreign Studies College.

#### Co-presenter:



### Boran WANG

- Professor of English
- Northeastern University
- China

Dr Wang is professor of English at Northeastern University. His research efforts have focused on cognitive linguistics, corpus linguistics, and computer-assisted language learning. He is currently the vice dean of Foreign Studies College.

#### Co-presenter:



### Xin LI

- Associate Professor of English
- Northeastern University
- China

Xin LI is associate professor of English at Northeastern University. Her research efforts have focused on language testing, assessment and computer-assisted language learning. She is currently the vice dean of Foreign Studies College.

#### Abstract: Calibrating the CEFR against the China's Standards of English for College English Vocabulary Education in China

The Common European Framework of Reference (CEFR) (Council of Europe 2001) has, over the past decade, come to be widely used as a reference tool for teaching, learning and assessment (Alderson, 2002; North, 2014). The focus of the current study is on scaling the China Standards of English (CSE) vocabulary descriptors for College English education in Mainland China, where College English education refers to English language education for non-English major students at tertiary level. A review of the CEFR and the College English Curriculum Requirements (CECR) (Ministry of Education, 2007) indicated that the vocabulary descriptors in both documents were inadequate to describe vocabulary knowledge for College English education in China.

On the basis of the CEFR and CECR descriptors, a pool of 39 descriptors were collected and categorized. Twenty-two English teachers from a Mainland China university were invited to participate in the study. They were first given the



CEFR and CECR vocabulary descriptors, after which they were asked to scale the descriptors with the CEFR as the reference point. Multi-Faceted Rasch Measurement (MFRM) was used to validate teachers' scaling of the vocabulary descriptors. The MFRM analysis showed that while the descriptors at C1 level were generally ranked as expected, teachers had difficulties in ranking descriptors at the CEFR B1, B1+ and B2 levels. The study indicates that teacher judgement of the scales provides evidence of the CSE scales, and can be a source of valuable information for the future improvement of the CSE.



### Thomas Garside

- Director of Teacher Training,
- Trinity College London
- Hong Kong

Tom Garside has 19 years of teaching and training experience in Europe, New Zealand and China, and is Director of Teacher Training at a busy Trinity CertTESOL and DipTESOL provider in Hong Kong. He also leads in-service training for native and non-native-speaker teachers in a wide range of teaching settings around Asia. Tom's experience in the Chinese EAP context, along with his work investigating the role of cognitive levels of output in student and teacher interaction in Chinese pre-university classrooms, has deepened his perspective into the ways that assessment settings relate to learning environments and vice versa.

#### Abstract: Application of the Sociocognitive Framework in Contemporary Testing

**Aims:** To outline the key features of the sociocognitive framework, and to examine the notion of cognitive validity as a driving force for the development of language assessments

To align aspects of the framework to local and contemporary assessment activities, with a special focus on the Trinity ISE and the 2017 Gaokao examinations

The idea of a holistic, sociocognitive framework has developed from the early work of Weir (1990, 1993) in China, Weir and O'Sullivan (2011) in more international contexts, and then by additional research by the assessment thinktank at CRELLA. The framework brings together the 'social, cognitive and evaluative dimensions of language use' to inform assessment practices. Incorporating the assessment context, the outcomes and final use of test results enables a progressive view of assessment. This is particularly relevant in China following the changes to the 2017 Gaokao examinations, and for testing organisations such as Trinity College London.



### Nguyen Thi Ngoc Quynh

- Director of the Center for Language Testing and Assessment
- University of Language and International Studies
- Vietnam National University Hanoi
- Vietnam

Nguyen Thi Ngoc Quynh holds a PhD in Applied Linguistics from the University of Melbourne, Australia. She is currently the Director of the Center for Language Testing and Assessment at the University of Languages and International Studies, Vietnam National University Hanoi. She plays a leading role in several projects on language assessment and teacher development of the National Foreign Language Project 2020. She has also taught applied linguistics, teacher education and language assessment courses. She reviews for some journals on second language acquisition and teacher education, and has presented and published on second language education and assessment.

#### Co-presenter:

### Nguyen Thi Quynh Yen, Nguyen Thi Phuong Thao, Thai Ha Lam Thuy, Bui Thien Sao

- Lecturers, Center for Language Testing and Assessment
- University of Languages and International Studies
- Vietnam National University
- Hanoi, Vietnam

#### Co-presenter:

### Nathan T. Carr, Professor,

- California State University
- Fullerton
- United States

#### Abstract: Test Quality Improvement Guidance for VSTEP Test Writers

The Vietnamese Standardized Test of English Proficiency (VSTEP) is a test of general English proficiency developed as part of Vietnam's National Foreign Language Project 2020. The test is based on the Common European Framework of Reference, and includes a multi-level test targeting levels B1-C1. The VSTEP consists of sections assessing reading, writing, speaking, and listening, with all four sections taken by all test takers. This inclusion of performance-based tasks in a large-scale language proficiency test is intended to promote positive washback from the test, and to shift the focus of English instruction in Vietnam toward a more communicative orientation.

The proposed presentation outlines the procedures followed to assure the quality of the VSTEP tests by identifying and reviewing problematic reading and listening test items used on the VSTEP B1-C1 test following initial administrations. Item analyses focusing on B-index values at each cut score, as well as the point-biserial for an overall estimate of discrimination, were combined with distractor analyses, which included both response frequencies and option point-biserial values. For problematic items and distractors, trends were identified; these are discussed with reference to principles of good item writing (see, e.g., Cangelosi, 1990; Haladyna & Downing, 1989a, 1989b; Haladyna, Downing, & Rodriguez, 2002), and in terms of the revised guidance to be provided to VSTEP task developers at the University of Languages and International Studies, Vietnam National University Hanoi. These insights can be used to provide improved guidance and training to item writers.



### Daniel Bates

- English Lecturer, Visiting Faculty Member
- Asia University
- Japan

Daniel Bates has been teaching English across Asia for over ten years in a number of universities across South Korea, Thailand and now Japan. He has developed a keen interest in language assessment over that time and is focusing my research on language testing techniques and models. In addition, he has a growing interest in pronunciation pedagogy and assessment, and is researching educator's views and techniques for teaching and assessing pronunciation in the classroom.

#### Abstract: The Use of Language Proficiency Descriptors for Pronunciation Assessment

While pronunciation generally accounts for 25% of the total score in most of the international speaking proficiency frameworks, it is frequently ignored or under-represented in ESL classrooms by test creators, materials developers and teachers alike. This presentation will look at how these pronunciation frameworks can be modified and used practically as assessment tools on ESL courses, with a particular focus on low-level learners.

This presentation will begin by outlining some of the current rubrics in place for pronunciation in the major language proficiency speaking tests, such as IELTS, TOEFL and TOEIC. Next, in order to assess how these rubrics can be adapted for assessment in localized settings, the presentation focuses on how frequently the reoccurring elements from these rubrics (such as certain segmental and suprasegmental features) are

used in common ESL textbooks, and whether an adequate amount of time and space is given to pronunciation.

Finally, the presentation will move on to look at how these rubrics can be developed and adapted to specific learning environments, in this case Japanese university classrooms, and how these pronunciation frameworks can be adjusted for both formative and summative assessment in ESL university courses.

---



## Breakout Session on Day 2



# Topic: Foreign Language Testing for Employability



### Johanna Motteram

- Teacher, Assessment Specialist
- British Council
- Singapore

Dr Johanna Motteram is an experienced language teacher with a special interest in high stakes language testing. Her teaching work includes candidate preparation for language tests and writing for specific contexts. Her research interests include further expanding the construct of writing by addressing questions related to the author/audience relationship as expressed in text, and test washback and impact.

Johanna is Australian but lives and works in Singapore. She has interests in language teaching and language test preparation in the region.

#### **Abstract: Employment in the UK, IELTS Writing, and Philippines-trained Nurses**

Philippine-trained nurses Nurses play an important role in the health systems of many countries. Demand for their services is high. One barrier to individuals joining this pool of valued professionals in the UK is the requirement that Nurses achieve an overall IELTS score of 7, with no band score less than 7, for registration for employment. Anecdotally, Philippines trained Nurses have trouble with the Writing subtest. Indeed, reported mean band scores for Filipino speakers in the Academic test in 2015 listed Writing as lowest mean band score (6.3) by half a band score (ielts.org).

This project addresses the specific challenges faced by Philippines trained Nurses attempting to achieve a 7 in Writing. The data includes an online survey (n=750) and a dataset of 100 marked responses to one IELTS Writing subtest (50 Task Ones, 50 Task Twos).

Analysis of the survey responses revealed social and geographic obstacles to accessing quality test preparation assistance. Quantitative and qualitative analysis of the dataset found

achieving elements of the scoring criteria Task Achievement /Response and Coherence and Cohesion posed difficulties for participants. These research findings informed logistical and instructional choices made in the development of an online IELTS Writing course specifically targeted to this candidate population.

This paper discusses the unique challenges of the regional project (the Philippines, Hong Kong, and Singapore), and illustrates the connections between research findings and instructional choices. It presents an application of mixed-methods language testing research work to propose a solution to an international workforce supply problem.



### Lan LUO

- PhD student
- School of International Studies
- Zhangjiang University
- China

Lan LUO is currently a full-time PhD student at the School of International Studies of Zhejiang University. She received her bachelor degree in French from Communication University of China in 2009. Then she furthered her studies in Durham University and received her master degree in business management in 2011. Later, she has taught English at Hangzhou Normal University for over four years. Last September, she resigned from the university and became a full-time PhD student at Zhejiang University. Her research interests lie in the field of language testing and ESP.

#### **Abstract: A Study on the Effect of Classroom-based Assessments on Medical Students' EMP Learning**

The present study focuses on the effect of classroom-based assessments on medical students' EMP (English for Medical Purposes) learning. Five classroom-based EMP assessments

were designed based on the framework of task characteristics in LSP (Language for Specific Purposes) tests and the conceptual framework of classroom-based assessment. After pilot testing, they were implemented in a university-level EMP course to explore their effectiveness in promoting medical students' EMP skills and the extent to which they have in fostering medical students' motivation and self-reflection in EMP learning. A-100-student class was divided into two groups (the experimental group and the control group) based on the ranking of an English proficiency test at the beginning of the course. T-test indicated that there were no significant differences between the English proficiency of the two groups, and the students' GPA scores in medicine were also taken into consideration at this stage. Both groups were under the same condition of teaching and testing, but only the experimental group participated in the classroom-based assessments. Feedbacks were given back to them through Seesaw app (an e-portfolio tool) to help them track their learning process. All the students were invited to complete a pre-study and post-study questionnaire to identify if there is any enhancement in their learning motivation and self-reflection. Qualitative data were collected through videotaped classroom observation and in-depth interviews which provided further interpretation of students' classroom behaviours and their perspectives towards EMP learning. The final results revealed that the experimental group gained more motivation and self-reflection in EMP learning, and had significantly improved its EMP skills compared with the control group. This study not only highlights the importance and necessity in LSP testing, but also provides rich insights into the reform of EMP assessments and EMP learning.



### Mike Mayor

- Director, Global Scale of English
- Pearson Education
- United Kingdom

Mike Mayor is Director, Global Scale of English at Pearson. In this role, Mike heads up research into creating audience-specific learning objectives aligned to the Global Scale of English and works with Content teams to ensure that these learning objectives underpin all new products and services created by Pearson English. On leaving university, Mike worked as a teacher of English in France before entering the world of publishing as a lexicographer. Mike joined Pearson in 2003 and headed up the dictionaries until his move to the Global Scale of English in 2013.

**Abstract: Creating a Proficiency Framework for the Workplace**

As English increasingly becomes the language of communication for international companies, so the pressure on governments and institutions grows to ensure that employers have linguistically skilled workforce to draw on. At the same time, there is disillusionment with existing assessments that are used to demonstrate work-readiness, a disillusionment shared by both employers and employees. In research carried out by Pearson and LinkedIn, only 7% of employees felt they were able to communicate effectively at work. Could one reason for this dissatisfaction be the fact that current assessments have been based on proficiency frameworks that were not designed to measure English as it is used at work?

This presentation looks at the development of a proficiency framework for the workplace: the Global Scale of English Learning Objectives for Professional English. The presenter will give an overview of the project to develop business-specific learning objectives aligned to both the Global Scale of English (GSE) AND THE Common European Framework of Reference for Languages (CEFR) before reporting on a second phase of the project: identifying the key language skills required for specific jobs. This phase involved aligning GSE Learning Objectives for Professional English to the job tasks specified for over 200 roles in the O\*NET database, a jobs database sponsored by the US Department of Labor, Employment and Training Administration.



**Alan Urmston**

- Assistant Professor and Assessment Coordinator in the English Language Centre
- The Hong Kong Polytechnic University
- Hong Kong

Alan is Assistant Professor and Assessment Coordinator in the English Language Centre at the Hong Kong Polytechnic University, responsible for assessment policy, design and implementation. He also teaches on undergraduate and postgraduate programmes within the University. Alan holds a PhD in Educational Linguistics from the University of

Bedfordshire and an MA in Teaching English as a Second Language. His main research interests are in language testing and assessment, in which he publishes and presents at international conferences. Alan has coordinated numerous assessment research and development projects and worked as a consultant to educational institutions and government bodies.

**Co-presenter:**



**Michelle Raquel**

- Senior Lecturer
- Centre of Applied English Studies
- University of Hong Kong
- Hong Kong

**Co-presenter:**



**Felicia Fang**

- Language Instructor
- The Hong Kong Polytechnic University
- Hong Kong

Felicia is Language Instructor in the English Language Centre at the Hong Kong Polytechnic University. She teaches undergraduate courses in EAP and Business English within the University. She has also been involved in the operation and development of large-scale assessments for the University. She is interested in the assessment of writing, diagnostic assessment of academic English and educational measurement models (e.g. MFRM). She is currently pursuing her EdD (part-time) in Education with Graduate School of Education of the University of Bristol.

**Abstract: Investigating the Writing Ability of Graduates and Employees of an Asian University**

The need for university students to take an exit test of their language proficiency is a question that has had particular relevance in societies such as Hong Kong where proficiency in the world language of English is so important, even though it may not necessarily be the dominant language in society. Concerns about the English proficiency of university graduates led to the need for some kind of measure of their proficiency and from this the Graduating Students' Language Proficiency Assessment (GSLPA), an assessment of written and spoken English, was born in the late 1990s, perhaps the only English language proficiency assessment

designed specifically for this purpose anywhere in the world. The GSLPA was relaunched in 2005 utilising the best of available technology for assessment delivery, scoring and reporting and has been in operation since, as a graduation requirement of students for the University. In 2013, the Human Resources Office of the University required a test of written English proficiency for both recruitment and promotion for non-academic staffing and as the GSLPA was designed as a workplace assessment, it was utilised for this purpose as well.

This paper reports on an investigation into the characteristics of writing ability that have been demonstrated by both final-year students and employees/potential employees of the University, utilizing methods of Many-Facet Rasch Analysis and Structural Equation Modelling, to gain insights into whether the University is adequately preparing its students for the workplace.



**Tomoyasu Akiyama**

- Professor
- Bunkyo University
- Japan

Tomoyasu AKIYAMA is a professor at Bunkyo University in Japan. He obtained his M.A. at Ohio State University in 1992 and his Ph.D. at the University of Melbourne in 2004. He has taught language testing and teaching of English to his students who want to become English teachers at junior and senior high schools. He won an ELEC (English Language Education Council) award in 2000. His research interests include validity investigations of high school, university and teacher employment examinations, and applications of the Rasch measurement to large scale speaking tests.

**Abstract: Rater Cognition: Impacts of Rater Differences on Teacher Employment Examinations**

This paper explores rater differences which have occurred in employing prospective English teachers mainly qualitatively (think-aloud method) using NVivo, which reveals how raters assess English teacher candidates. In order to become an English teacher at public schools in Japan, all candidates are required to undertake teacher employment examinations. One part of this examination requires candidates to show their teaching skills and ability via microteaching. This paper focuses on rater's rating process of candidates' teaching performance via microteaching because raters, as previous studies (e.g., Lumley, 2005; Taylor, 2011; Bejar, 2012) indicated, may have a large impact on giving scores to test-takers.

Thirty candidates undertook microteaching tests,

which were videotaped for a later assessment. Seventeen raters rated all candidates using six assessment criteria (e.g., instruction ability, lesson flow and personality, delivery, overall impression and expertise). Raters were required to give comments to every rating assessment criterion while assessing. Data were analysed using NVivo in order to explore rater cognition when they were assessing. Results of the think-aloud method show that raters rated candidates not only with different interpretations of and different attention to aspects of assessment criteria but also according to different core teaching values and ideal images of English teachers or English classes possessed by the raters themselves. These could lead to divergence between lenient and harsh ratings of the same candidates by different raters. Finally, this paper discusses issues with rating in performance assessment contexts and implications for cognition research based on implications drawn from this study.

---

## Breakout Session on Day 2



# Topic: Foreign Language Testing for Admissions and Exit Tests in Education Systems



## Jufang KONG

- Lecturer
- Zhejiang Normal University
- China

Dr Kong earned her PhD. from Shanghai International Studies University, and is now a full-time teacher at Foreign Languages College of Zhejiang Normal University, China. Her main research interests include test validation, assessing reading, language assessment literacy and assessing young learners.

followed by a cross-sectional review of the existing primary school exit or secondary school entrance English test papers in China. Discrepancies are identified between cognitive and psychological characteristics of young language learners and the design of current tests assessing young EFL learners in China. The juxtaposition of young learners' responses to the tests elicited from the interviews together with the review of the test papers reconfirms that verbal input alone somehow poses a threat to the cognitive validity of a language test intended for young learners. It is also found that pictorial/non-verbal prompts, even though just playing a supplementary role of information input, are conducive to enhancing the cognitive validity of tests for young learners. This study also attempts to shed more light on the cognitive processing of young EFL learners involved in test taking and on the improvements for test design.

The contexts in which English is not only the subject of teaching and learning, but is also the medium through which other subject areas are taught, are increasing in traditional EFL environments. Approaches to English-medium instruction can range from full immersion programs based on international school models, to partial offering of some subjects in the curriculum. Reading ability is a key component of the academic English proficiency students will need to succeed in these environments as it underpins all subjects across the curriculum. Building an understanding of the reading demands of the different subjects, along with an estimate of the ability of students to cope with those demands would provide an invaluable element of feedback for both learners and teachers.

This study employs the Lexile Framework for Reading, developed to match readers with text at a level that provides challenge but not frustration, to investigate the reading demands in one international school in China. In addition to textbooks in key subject areas, the examinations which students study towards were benchmarked using the same framework.

In this presentation, the picture that emerges will be contextualized against data for a range of similar education levels in English speaking countries and EFL environments. This allows for the development of a profile for the reading needs of learners in this English-medium context and will offer insights into the development of similar profiles for other contexts.

### Co-presenter:



## Mingwei PAN

- Associate Professor of Applied Linguistics
- Guangdong University of Foreign Studies, Guangzhou
- China

### Abstract: Assessing young English Learners in China: Cognitive Validity Re-examined

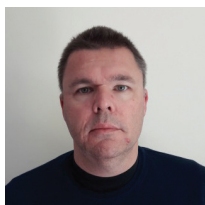
China has witnessed a considerable number of young children learning English as a foreign language in the past decades, particularly after the implementation of the two-child policy nationwide. High-stakes decisions are thus made on admission to secondary education largely based on the scores of provincial or municipality-level standardized English tests. However, it remains to be explored as to what extent these tests are valid. This study focuses on the cognitive validity of assessments of young English learners in China through a survey of the extant literature on the characteristics of the young language learners in age, interest and motivation, social and personal growth,



## Ali Van Moere

- Chief Product Office
- MetaMetrics
- United States

### Co-presenter:



## Richard Spiby

- Test Development Researcher, Receptive Skills
- British Council
- United Kingdom

### Abstract: Understanding the Reading Demands in an English Medium Instruction Program



## Shangchao MIN

- Associate Professor at School of International Studies
- Zhejiang University
- China

Dr Min is an Associate Professor at the School of International Studies, Zhejiang University, Hangzhou, China. Her main research interests are in language testing and assessment, especially computerized language testing and

item response theory. She has been involved in several research projects funded by the National Social Science Foundation of China on topics such as the development of computerized adaptive language tests and has undertaken several research projects funded by the Provincial Department of Education as a principal investigator. She has publications on Language Testing and Language Assessment Quarterly.

**Co-presenter:**



**Lidi XIONG**

- Postgraduate at School of International Studies
- Zhejiang University
- China

Lidi XIONG is a postgraduate at the School of International Studies, Zhejiang University, Hangzhou, China. Her main research interests are in language testing and assessment.

**Abstract: Retrofitting an In-house English Listening Test for Diagnostic Purposes**

Cognitive diagnostic modeling has gained great momentum in the field of language assessment in recent years because it can help researchers identify individual learners' strengths and weaknesses in multiple attributes. Despite the surge of interest, empirical studies remain rare, mainly focusing on reading ability (e.g., Chen & Chen, 2016; Jang et al., 2015; Li et al., 2016), with only a handful on listening ability, using large-scale standardized tests such as TOEFL (Lee & Sawaki, 2009) and MELAB (Sawaki et al., 2009). By adopting the Generalized-DINA model, this study attempted to retrofit Zhejiang University English Listening Test, a high-stakes in-house test, for diagnostic purposes. This test is used to make pass or fail decisions for granting bachelor's degree for non-English majors at this university. The data used are 917 test-takers' responses to 30 items of 1 test form in 2016. It was found that of the 6 attributes measured in the test, namely, recognizing sounds (A1), understanding words and syntactical structures (A2), extracting explicit information (A3), summarizing main ideas (A4), making inferences (A5) and identifying speakers' attitudes and purposes (A6), A2 was the best mastered and A4 the worst mastered by test-takers across all grades. In addition, this study reported detailed information regarding the attribute mastery probability of individual test takers as well as the three listening proficiency groups (i.e., low-level, intermediate, and high-level). Such information can be provided to the program administrators and teachers for enhancing the EFL listening curriculum and designing teaching materials.



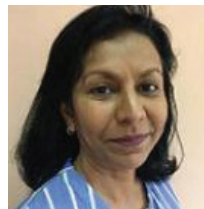
**Jason West**

- Head of English Department
- SILC / Shanghai University
- China

Jason W. BSc. MLM is the Head of English at SHU UTS SILC Business School, where he's been teaching for 10 years. Jason has a strong ethos for supporting student-mentoring initiatives such as speech contests and academic counseling. In addition to holding a CELTA and English teaching, Jason has held a number of teaching and learning, management, and, corporate and operational HR roles in the telecom's industry, in Australia.

**Abstract: Connecting the Dots with Aptis Testing Insights**

How do you know if you are measuring the right things and what do you do with the data? As China's longest running education joint venture, we've applied the Aptis test results to our business, to improve quality initiatives and outcomes for our students. Examples of this application include recruitment and streaming. One challenge has been how to measure and manage our entry-level Pathway students' English levels across all skill areas, to optimise their educational efficacy. So, using Aptis has helped to solve this problem. Language testing is often viewed as a discrete one-off test, but having useful data is much more powerful as an enabler, so that you can connect the dots. In this session, we'll present our insights into how Aptis testing has unlocked potential for SILC / Shanghai University. If you can't measure it, you can't manage it!



**Suchithra Nair**

- Lecturer
- Universiti Teknologi MARA
- Malaysia

Suchithra Nair is a lecturer at Universiti Teknologi MARA. She holds a Master's degree in Instructional Design from Aberdeen University, UK. She has considerable experience in teaching and testing proficiency level English. Also, she has been involved in the design, development and implementation of an exit test which has been use in use since 2016 at Malaysia's biggest public university.

**Co-presenter:**



**Saidatul Akmar Zainal Abidin**

- Associate Professor
- Universiti Teknologi MARA
- Malaysia

Saidatul Akmar is a lecturer at Universiti Teknologi MARA, Malaysia. She holds a PhD in Testing & Evaluation from UK. She also has practical knowledge of teaching and testing proficiency level English. She has been involved in the design, development and implementation of an exit test which has been use in use since 2016 at Malaysia's biggest public university.

**Abstract: Design, Development and Implementation of a University Exit Test**

Within the context of recent concerns that graduating students are unprepared for the English language demands of the workplace in Malaysia, this report describes the development and implementation of the English Exit Test (EET) at a Malaysian institution. The EET was solely designed to gauge graduating students' readiness for the workplace. The inception of the EET marked the beginning of the unspoken dialogue between the workplace players and the university course designers in so far as assessing their English readiness for the workplace. The tasks in test are kept authentic to reflect the workplace demands. The present article starts by recounting the pilot stage of implementation where test design and development decisions were informed by the workplace needs and university English language curriculum and then goes on to explain the formidable technical challenges faced by test designers to keep a large-scale test like the EET authentic and relevant. In particular, this article explains the evolution of the EET from inception to its current form. In foreseeing future assessment demands and the logistics of running a far-reaching exit test, this article considers certain imperative issues and makes several recommendations to mitigate the risks.



**Fengping GUO**

- Lecturer
- Southeast University
- China



Got Bachelor's Degree in Nanjing Normal University and Master's Degree in Southeast University (SEU). Fengping GUO has taught mainly College English for non-English majors after graduation. Full preparations and flexible classroom tasks are the special features of Fengping GUO's teaching, especially within the teaching of IELTS courses, part of the Advanced College English Courses System in SEU.

**Co-presenter:**



**Maoxia YANG**

- Vice Director of the Testing Centre
- Southeast University
- China

Graduated from Nanjing Normal University in 1999 with a Master's Degree in English Language and Literature, Maoxia YANG has been working as a College English teacher in Southeast University for about 18 years. With great interest in teaching and testing, Yang Maoxia has collaborated with colleagues to design various types of College English tests, analysing results and researching into the effects of tests on EFL learning.

**Abstract: An Investigation of the Use of Apts for EFL Placement Decisions in Tertiary Education**

Since the release of College English Curriculum Requirements (2007) by the Ministry of Education in China, placement testing is universally adopted by Chinese colleges and universities to place freshmen into appropriate EFL courses. In some institutions, the score of College Entrance Examination is used to identify students for proper programs (Chun HAN, 2008). However, since College Entrance Examinations vary from province to province in both form and degree of difficulty, it is not easy to create student groups of homogeneous language ability with these scores. In other institutions, locally created, administrated and scored tests are used to distinguish students in terms of their language proficiency. In most of these placement tests, vocabulary range, grammatical range, listening, reading and writing ability of the students are assessed. However, evaluation of students' speaking ability has never been included (Li GUO, 2003; Rui WANG, 2006; Bona LI, 2011; Shaoyan ZOU, 2012; Long WANG, 2017).

This study reports the experiences of teachers and students, in the use of an internationally developed language test, Apts (consisting of reading, listening, writing as well as speaking modules), which was used for placing 120 freshmen into an EFL programme.

The report describes the university context in East China and the EFL programme in which the test was used. The results of qualitative

data, gathered from interviews with programme participants (2 teachers and 12 students), is presented and provides preliminary evidence supporting the use of a four- skill test as an initial screening tool.



**Dafu HUANG**

- Professor of Applied Linguistics and Dean of College of Humanities and Social Sciences
- Southern Taiwan University of Science and Technology
- Taiwan

Dr Huang is professor of applied linguistics and Dean of College of Humanities and Social Sciences of Southern Taiwan University of Science and Technology. He served as chairperson of Department of Applied English from 2003 to 2011, and has been director of Language Center since 2008. Receiving his PhD in applied linguistics from University of Texas at Austin in 2003, he specializes in language testing and evaluation, corpus linguistics, statistical modeling in applied linguistic research, and quantitative methods in EMI research. He focuses his recent research on English vocabulary learning, corpus linguistic approach to technology-enhanced language learning, corpus analyses of EMI and ESP, integration of corpus analysis tools and R statistics in applied linguistics

**Abstract: Developing a Global English Test as a Learning and Exit Measure for Taiwanese Polytechnic University Students**

Polytechnic university students in Taiwan have overall lower English entry proficiency and motivation as they come mostly from vocational high schools where English learning is not emphasized. Since polytechnic university students are trained to meet the industry needs in the international work contexts, equipping them with needed English communication skills should not be overemphasized. Still, with the rapid globalization of the English language and generally limited English input and learning hours available to the students, developing a practical, high-frequency words based curriculum and an exit test is crucial and feasible to enhance their English ability and motivation. English learning materials are thus created for formal class learning at Southern Taiwan University of Science and Technology which involves learning of the 5000 most frequently occurring words derived from BNC (British National Corpus)-COCA (Corpus of Contemporary American English) corpus and 1500 words of Nerrière's Globish. The text coverage of the learning materials by the 5000-word list reaches ca. 95% to ensure adequate comprehension by the students if they know the 5000 words. Moreover, based on the word list,

the Global English Test (GET) was developed as a criterion-referenced test to gauge students' familiarity with the high frequency target words as an alternative measure for those who fail to pass the English exit threshold set by STUST (i.e. English ability levels equivalent to CEF/ A2 or B1 required by different faculties). With two levels (fundamental and pre-advanced) and 95% coverage by the 5000-word list, the GET assesses students' ability to deal with real-life daily functions in authentic English use contexts. To ensure the concurrent validity and improve the content validity of the GET, cross-checking of test results of the GET is made with those of the GEPT (General English Proficiency Test), also a criterion-referenced test developed by the LTTC of Taiwan. The development of the GET is expected to bring about positive washback effects and shed insights into English teaching for vocation higher education in Taiwan.

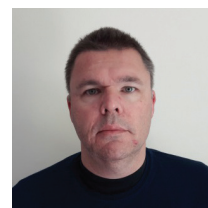


**Jamie Dunlea**

- Senior Researcher, Assessment Research Group
- British Council
- United Kingdom

Dr Jamie Dunlea is a Senior Researcher for the Language Assessment Research Group at the British Council, based in London. He has 25 years of experience working in EFL education, first as a teacher, then in test development and production and assessment research. He has worked on a number of large-scale national and international test development projects as well as advising institutions and ministries on assessment reform in education systems. His areas of interest and recent work include the application of validity theory to local testing contexts and linking examinations to frameworks and standards such as the CEFR, development contexts.

**Co-presenter:**



**Richard Spiby**

- Test Development Researcher, Receptive Skills
- British Council
- United Kingdom



Co-presenter:



### Nguyen Thi Ngoc Quynh

- Director of the Center for Language Testing and Assessment, University of Languages and International Studies
- Vietnam National University
- Vietnam

scale. In addition, a comprehensive proforma utilizing categories drawn from the socio-cognitive model and the growing body of CEFR alignment studies was developed to evaluate the constructs targeted by both tests. Questionnaire data from test takers also offers interesting insights into the attitudes of the university students regarding such aspects as differences in delivery mode for the productive skills.

Co-presenter:



### Nguyen Thi Quynh Yen

- Deputy Director of the Center for Language Testing and Assessment, University of Languages and International Studies
- Vietnam National University
- Vietnam

Nguyen Thi Quynh Yen is currently the deputy director of the Center for Language Testing and Assessment, The University of Languages and International Studies, Vietnam National University (ULIS-VNU). She is now a PhD candidate in English Language Assessment.

#### **Abstract: A Closer Look at Two Tests Aligned with the CEFR**

This paper reports on a comparability study of two EFL proficiency tests which use an international proficiency framework, the Common European Framework of Reference (CEFR) as an important source of feedback for test takers. The two tests are VSTEP, a pen-and-paper test in Vietnam targeting CEFR levels B1 to C1, and APTIS, an international computer-based test targeting CEFR levels A1-B2. The VSTEP is recognized by universities in Vietnam as certification of English proficiency for the purpose of meeting graduation requirements stipulated by the Ministry of Education. APTIS is used for a range of purposes in international settings, including by higher education institutions in EFL contexts. This study reports on the main phase of the project in which over 400 test takers at three universities in different regions of Vietnam took both tests, building on a smaller pilot phase at one university to trial the methodology. The socio-cognitive model for language test development and validation was used to design a multi-method approach. Statistical analysis of test scores includes factor analysis as well as a concurrent Rasch analysis to place test items from both tests on a common

## Breakout Session on Day 2



# Topic: Artificial Intelligence and Technology in Language Testing



## Hui LIN

- Co-Founder and Chief Scientist
- Shanghai Liulishuo Information & Technology Ltd.
- China

### Abstract: Large Scale Adaptive Language Testing via Artificial Intelligence

Artificial Intelligence has made a big leap forward recently thanks to the breakthroughs in deep learning technologies. A major benefit of the technology innovations in language testing is the ability of employ adaptivity and a variety of interactive tasks, e.g., speech interface, in tests. In this presentation, we show how these innovations have been used to develop an online placement test, serving over 10 million English learners.

With regard to adaptivity, we introduce a novel algorithm based Fisher information reduction to determine testing path and item bank usage. Moreover, an incremental approach to item bank construction is employed, and thanks to the large volume of active test takers, we are able continuously improve item bank quality and therefore test validity at an unprecedented speed.

Another important innovation is the ability to use speech recognition technology to score a student's ability to pronounce key phonemic patterns as well as their ability to chunk language in working memory as a measure of fluency. Equipped with deep learning technologies, the speech recognizer that we are using has proven to have an excellent ability to score pronunciation while also being robust to issues such as microphone and ambient noise levels.

Since the launch in July 2016, our placement test has served over 10 million English learners. Quantitatively, we show how the test validity evolves over time. We also show some interesting analysis on demographics of these 10 million learners as well as its correlations to language proficiency and etc.



## Xiaoyang HE

- Product Manager
- Foreign Language Teaching and Research Press
- China

Xiaoyang HE is the product manager in Language Assessment Department, Research and Development Center, Foreign language Teaching and Research Press. During 2009-2013, she studied in Xi'an International Studies University, with the major of English language and Culture. Then she got her master degree in Chinese University of Hongkong with a major of English Linguistics in 2014.

### Abstract: The Development of an Online Diagnostic Assessment System for Senior High Schools in China

In recent years, diagnostic tests have gained the attention of more and more researchers, which contributes to fruitful achievements in this area. As a kind of test facilitating teaching and learning, the application of diagnostic tests in actual context is particularly important. Udig is a web-based English diagnostic testing and learning system, which has started to be applied to teaching in practice. Based on the National Curriculum Standards for senior high schools in China, the system provides immediate feedback on the strengths and weaknesses of the subskills with reference to the learning targets and recommends exercises automatically to improve the weaknesses. Besides telling "what and where", the assessment also allows students to track their English language development as each time they take the language tests. At the same time, diagnostic information on the use of learning strategies related is also given in order to analyze "how to learn" problems. With the support, teachers and students can make adjustments on the plan as well as the process of their teaching and learning accordingly. This article aims to describe the contents of Udig and the way in which the system works. This is followed by an account of the construct, the approach to defining learning objectives and

the development process of the tests. The voice from the users is also touched upon to deepen our understanding of diagnosis in the practical context.

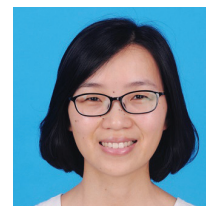


## Xiangdong GU

- Professor, PhD supervisor
- Chongqing University
- China

Xiangdong GU is a full time professor, PhD supervisor and director in Chongqing University, China. She also worked with Cambridge English Language Assessment, Cambridge University, UK as an academic consultant since 2012. Her main research interests span across language assessment, teacher development and foreign language educational data mining. She is leading a key research project on language assessment supported by the National Philosophy and Social Science Foundation of China [Grant No.14AYY010]. She has published dozens of journal articles and has given scores of invited keynote speeches, seminars and presentations in universities and at conferences nationally and internationally.

### Co-presenter:



## Yu ZHONG

- PhD candidate
- Chongqing University
- China

Yu ZHONG is a PhD candidate from Chongqing University, China. Her main research interests are language assessment and foreign language educational data mining. Over the last five

years, she has presented four and published five papers nationally and internationally, and played a key role in two research projects on language assessment supported by the National Social Science Foundation of China.

**Co-presenter:**



### Yao MENG

- Data Scientist
- Alibaba Corporation
- China

Yao MENG received her M.Sc degree in 2017 at College of Computer Science, Chongqing University, China. Now she is working as a data scientist at Alibaba Corporation, China. Her research interests include data mining and machine learning.

**Co-presenter:**



### Qing ZHOU

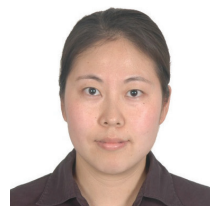
- Professor at College of Computer Science
- Chongqing University
- China

Qing ZHOU is a professor at College of Computer Science, Chongqing University, China. He received the B.Sc. degree of Engineering in 2002, the M.Sc. degree in 2005 and his doctor degree of Engineering in 2008, all at College of Computer Science, Chongqing University. His research interests include: data mining, machine learning, big data, and cloud computing and information security.

**Abstract: Analysing and Predicting English Majors' TEM-4 Performance in China**

Test for English Majors Band 4 (TEM-4) is a national English proficiency test for Chinese English majors at the end of the second university year. This paper focuses on analysing and predicting the TEM-4 performance of 77 English majors from a key university in China. With the permission of the school and cooperation of the students, a large amount of data was collected, including demographics and family status, learning-related achievements, learning motivations and learning journals kept by the students for a year. Three classification algorithms were used to predict students' TEM-4 performance. Results from the Naive Bayes

Classifier achieved the highest accuracy; in predicting those students achieving excellent, model accuracy was 90.9% and when predicting whether students could pass the exam, accuracy reached 98.7%. One contribution of this study was the analysis of rich data that might influence students' TEM-4 performance, and the selection of the most influential features that successfully predicted the test performance using data mining algorithms. Another aspect which had not been explored in previous TEM-4 research was that written learning journals were verified as improving the prediction accuracy.



### Jing CHEN

- Associate Professor in the School of Foreign Languages
- Sun Yat-sen University
- China

Dr Chen is an associate professor in the School of Foreign Languages at Sun Yat-sen University in China. Her research interests include systemic functional linguistics, language testing and second language writing.

**Co-presenter:**



### Pengpeng FENG

- Associate Professor in the School of Foreign Languages
- Sun Yat-sen University
- China

Dr Feng is an associate professor in the School of Foreign Languages at Sun Yat-sen University in China. Her research interests include computer-assisted language learning and second language speaking.

**Co-presenter:**



### Kai GUO

- MA student
- Shanghai Jiao Tong University
- China

Mr Guo is an MA student in the School of Foreign Languages at Shanghai Jiao Tong University in China. His research interests include language testing and computer-assisted language learning.

**Co-presenter:**



### Tan JIN

- Associate Professor in the School of Foreign Languages
- Sun Yat-sen University
- China

Dr Jin is a research associate professor in the School of Foreign Languages at Sun Yat-sen University in China. His research interests include language testing and data intelligence.

**Abstract: Re-fitting Text Complexity for Testing Reading: Introducing a Data-driven Approach**

While numerous corpus-based indices are available for assessing text complexity, item writers have mainly depended on their intuition and experience to level and adapt texts for testing reading. In this connection, this study reports on a data-driven approach employed to produce texts to be included in the reading component of the Sun Yat-sen University English Proficiency Test (SEPT). The data-driven approach features two stages, leveling and adapting. In the leveling stage, benchmark ranges of the text complexity were established. On the one hand, the National Matriculation English Test (NMET) and TOEFL iBT were investigated as two points of reference for the entry and exit levels of text complexity, respectively; on the other hand, the relationship between text complexity and student performance was examined based on 2017 SEPT test data for determining the distribution of text complexity adopted. In the adapting stage, a group of item writers were requested to conduct data-driven text adaptations through an online data-driven tool for English text adaptation, Eng-Editor. To facilitate the use of data-driven text adaptations, an online course was provided to the item writers, and semi-structured interviews were given to elicit their perceptions and feedback. Finally, a face-to-face workshop was held between testing experts and item writers to review and finalize the adapted texts. Overall, the data-driven approach encourages human-machine collaboration on the re-fitting of text complexity for testing purposes.



### Yong XU

- Associate Professor in English Language and Literature
- Shanghai International Studies University
- China

Associate professor in English language and literature. PhD. in applied linguistics from Shanghai International Studies University. Dr Xu's research interests: English language assessment, second language acquisition, English writing, cognitive and neurolinguistics.

item difficulty comprising these two sources. The grammar source involves two aspects (grammar form and grammar meaning) and six test focus types (phonological /graphological, lexical, morphosyntactic, cohesive, informational management, interactional), and the method source involves contextuality and reciprocity. Next, the article sets up levels for each attribute, fuzzy subsets of grammar MCQ difficulty, and degree of Membership Function, thus building a fuzzy rule to predict item difficulty with item features to be optimized with SVM. The article assigns attribute values to the 30 grammar MCQ items in two administrations of TEM4, and trained the SVM with responses of 452967 candidates. The effectiveness of the algorithm and implication for grammar assessment are also discussed.

#### Co-presenter:



### Shanshan SHANG

- Associate Professor in School of Business and Management
- Shanghai International Studies University
- China

Shanshan SHANG is an associate professor of information system and information management at school of business and management, Shanghai International Studies University, China. She received her Ph.D from the Tongji University. Shanshan's research focuses on knowledge management. For knowledge management, her research explores the factors influencing knowledge management, the relationship between the factors, the diffusion routes of knowledge through collaborative networks, how knowledge management helps enterprises to keep and enhance their key competence

#### **Abstract: Algorithm to Predict Item Difficulty of MCQ Grammar Assessment**

Developing an Algorithm to Predict Item Difficulty of Multiple Choice Item in Grammar Assessment

— the Case of Test for English Majors

This paper introduces an effort to development an algorithm that can automatically predict the difficulty of multiple-choice questions (MCQ) in English grammar assessment by drawing help from textual analysis, fuzzy mathematics and support vector machine (SVM). Starting with the perspective that language assessment is a "trait + method unit," the article attributes item difficulty to candidates' grammar knowledge and test method. Based on the grammar assessment theory of Purpura (2004), the article raises a framework for evaluating grammar





## Sponsors



### **IFLYTEK CO., LTD.**

IFLYTEK CO., LTD. (IFLYTEK), is a leading international company specializing in intelligent speech and artificial intelligence.



### **Shanghai Liulishuo Information Technology Co., Ltd.**

Liulishuo is China's leading "AI + Education" company in language learning. It was co-founded by Yi Wang, Hui Lin and Ben Hu in September 2012. Holding a PhD in computer science from Princeton University, Yi Wang is a distinguished expert of The Thousand Talents Program (established by the Government of China). Equipped with one of the best teams of AI scientists in the industry, Liulishuo has developed the world's most intelligent and effective AI English teacher that provides a personalized and adaptive learning path to every learner based on deep learning technologies. Liulishuo has closed a nearly \$100m series C funding and began to have scableable profit by Feb 2017. Its registered users have reached over 50 million.



### **Shenzhen SEASKYLAND Technologies CO., LTD.**

Shenzhen Seaskyland Technologies Co. Ltd. is a leading service provider of testing assessment in China and education evaluation services as well as a pioneering enterprise in the education informatization field in China, and is honored as the National Hi-tech Enterprise, Key Software Enterprise in the National Planning Layout, Key high-tech enterprise in the National Torch Program and Guangdong Integrity Model Enterprise.

**新东方** | 留学考试  
XDF.CN

### **New Oriental Education & Technology Group**

Since the establishment of New Oriental Education & Technology Group in 1993, the company's training services for overseas testing have helped millions of Chinese students study in universities abroad. New Oriental was the first Chinese brand committed to providing students with preparatory programs for foreign examinations in the US, UK, Canada and Europe, and is now the benchmark training institution for Chinese students preparing to study abroad.

 **高等教育出版社**  
HIGHER EDUCATION PRESS

### **Higher Education Press**

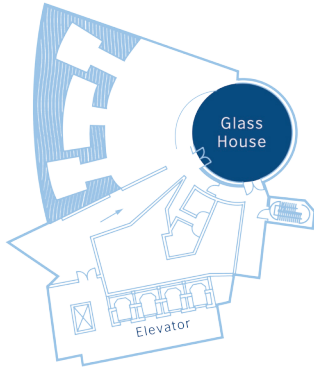
Founded in May 1954, Higher Education Press Limited Company (HEP), affiliated with the Ministry of Education, is one of the earliest institutions committed to educational publishing after the birth of P. R. China in 1949. After striving for six decades, HEP has developed into a major comprehensive publisher, with products in various forms and at different levels. Both for import and export, HEP has been striving to fill in the gap of domestic and foreign markets and meet the demand of global customers by collaborating with more than 200 partners throughout the world and selling products and services to more than 60 countries and regions in 24 languages. As a result, HEP ranks among China's top publishers in terms of copyrighting export volume for many years.

# Floor Plan



Ground Floor

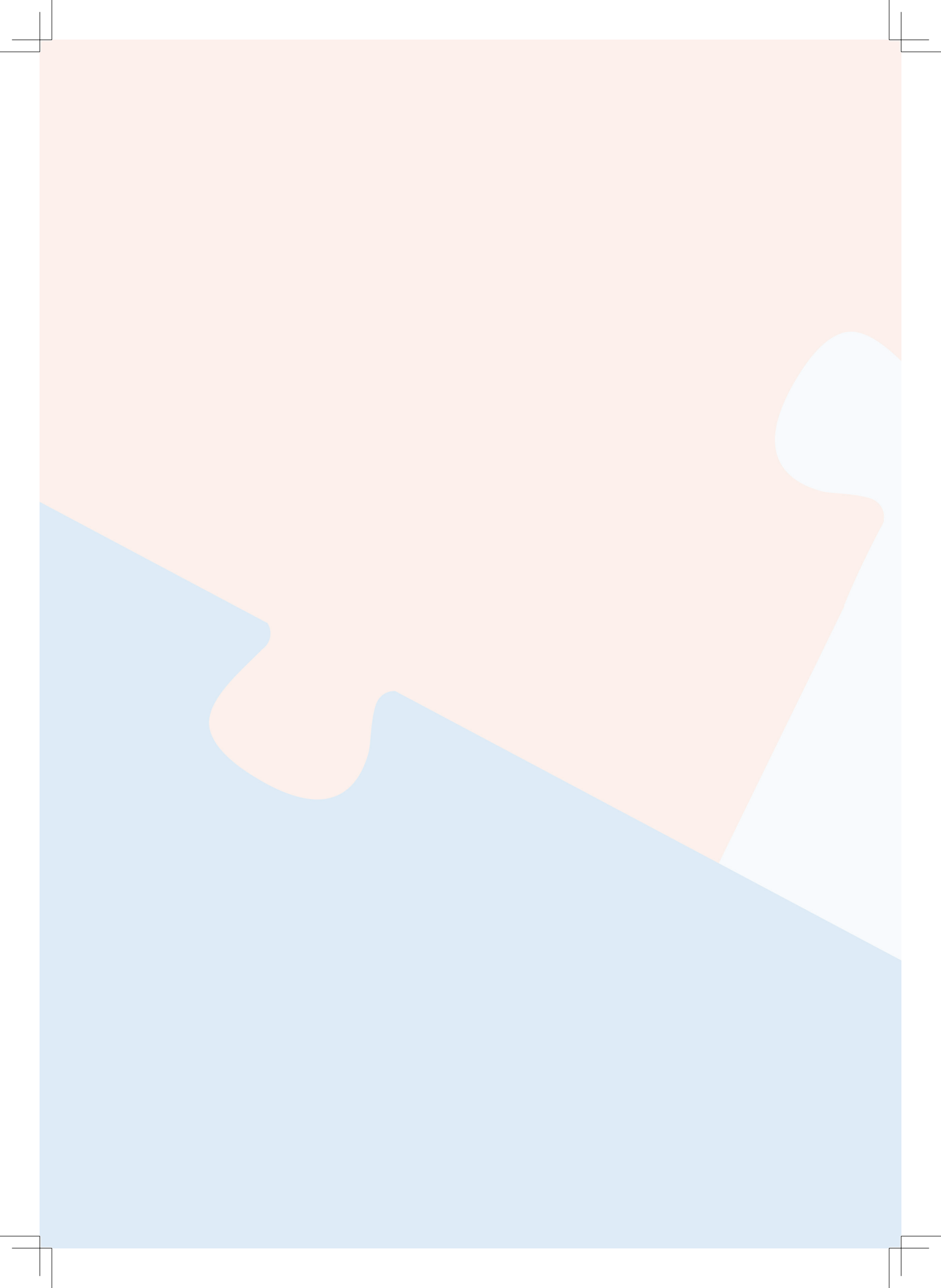
East Wing, 2F



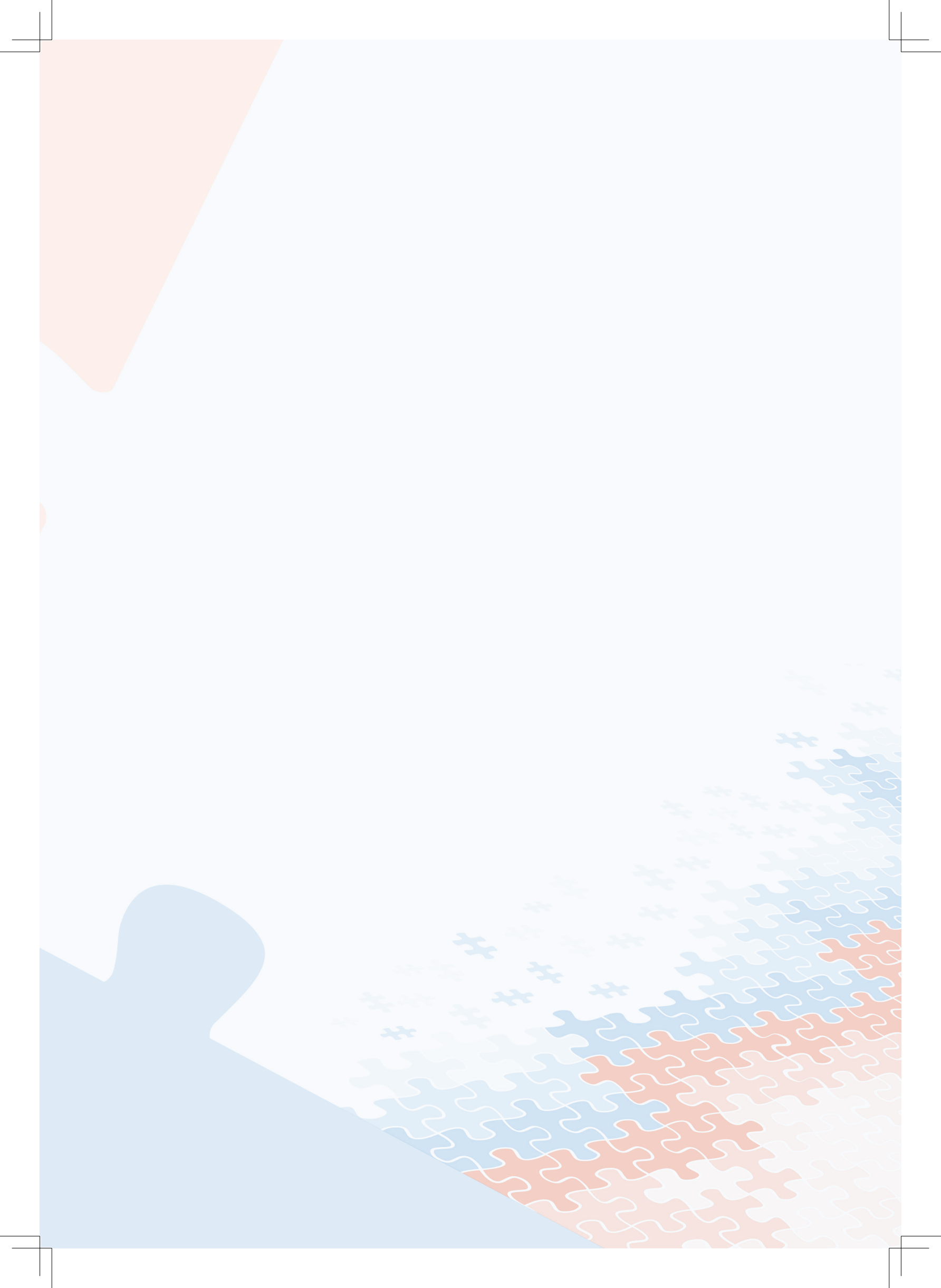
East Wing, Level M

### **Acknowledgement**

**We extend our deepest gratitude to  
Shanghai Jiao Tong University  
for their wonderful support and advice.**









英国文化教育协会



教育部考试中心  
NATIONAL EDUCATION EXAMINATIONS AUTHORITY