

19th EALTA Conference

# Sustainable Language Assessment

Improving Educational Opportunities at the Individual, Local, and  
Global Level

**13-18 June 2023, Helsinki, Finland**

## Pre-Conference

Workshops run parallel on three consecutive days

Tuesday, 13 June (14-17), Wednesday, 14 June (9-16), Thursday, 15 June, (9-12)

### Workshop 1

*P722*

EBB Rating Scale  
Development Bootcamp

*Beverly Baker & Heike  
Neumann*

### Workshop 2

*P674*

Examining the Dimensionality  
of the Language Tests Using  
Confirmatory Factor Analysis  
in R Statistical Software

*Gholam Hassan Khajavy*

### Workshop 3

*P673*

Using Sequential - Categorical  
Analysis to assess  
interactional competence

*David Wei Dai*

Thursday, 15 June, 2023

**12.00-14.00**

**Early Registration**

*Porthania, Aula*

**Parallel SIG Meetings**

**14.00-17.30**

**Assessment of Writing / Assessment for Academic Purposes**

*Sonja Zimmermann & Olivia Rütli-Joy*

*P673*

**CEFR**

*Neus Figueras*

*P111*

**Multilingual Language Assessment**

*Karin Vogt & Eleni Meletiadou*

*P674*

**Equality, Diversity and Inclusion (EDI)**

*Richard Spiby & Gemma Bellhouse*

*P723*

**16.00-18.00**

**Registration**

*Porthania, Aula*

**19.00-20.00**

**Welcome Reception**

**Old Court House (Aleksanterinkatu 20)**

**Host of the event is the City of Helsinki**

# Conference

Friday 16 June, 2023

08.00-09.00	<b>Registration</b> <i>Porthania, Aula</i>		
09.00-09.30	<b>Opening Ceremony</b> <i>PII</i> <i>Johanna Mäkelä, Dean of the Faculty of Educational Sciences, University of Helsinki</i> <i>Peter Lenz, EALTA President</i>		
09.30-10.30	<b>Keynote</b> <i>PII</i> <b>A language ecology lens for examining participatory action research in language assessment: Understanding inclusion, equity, and sustainability</b> <i>Jamie Schissel</i> <i>Chair: Peter Lenz</i>		
10.30-11.00	<b>Coffee</b> <i>Aula</i>		
11.00-12.30	<b>Parallel Paper Presentations</b>		
	<i>PIII</i> <i>Chair: Sonja Zimmermann</i>	<i>P674</i> <i>Chair: Barry O'Sullivan</i>	<i>P673</i> <i>Chair: Neus Figueras</i>
	11.00-11.25 <b>Applying a construct-driven approach to speaking task development in a digital environment</b> <i>Richard Spiby, Carolyn Westbrook, Johnathan Cruise, Mariano Felice &amp; Emma Bruce</i>	11.00-11.25 <b>Digital-First Language Assessment Systems as a Means of Achieving Accessible and Fair Language Assessment</b> <i>Yena Park, Ardeshir Geranpayeh &amp; Alina von Davier</i>	11.00-11.25 <b>Validating German Spelling Inventory for Adult Emergent Multiliterates: Evidence from the Rasch-Analysis</b> <i>Julia Edeleva &amp; Gina Do Manh</i>
	11.30-11.55 <b>The role of speech data in predicting fluency and proficiency assessments of L2 speech</b> <i>Heini Kallio &amp; Mikko Kuronen</i>	<del>11.30-11.55</del> <b>CANCELLED</b> <del><b>Developing a UDL-based accommodation protocol for digital exams</b></del> <del><i>Christine Troussart Van Bulck, Goedele Vandommele, Anne van Asseldonk, Lucia Luyten, Manon Willems, Lies Strobbe &amp; Jonas Wynants</i></del>	11.30-11.55 <b>A mixed-methods study into the role of the time variable in the construct of computer-administered C-Tests in three languages</b> <i>Anastasia Drackert, Anna Timukova &amp; Franziska Möller</i>

	<p>12.00-12.25</p> <p><b>What influences interactional competence? A structural equation modelling study of test-taker background variables</b></p> <p><i>David Wei Dai &amp; Ivy Chen</i></p>	<p>12.00-12.25</p> <p><b>The development and implementation of a policy for EDI (Equality, Diversity, and Inclusion) standards for the production of language tests</b></p> <p><i>Gemma Bellhouse</i></p>	<p>12.00-12.25</p>
<p><b>12.30-14.00</b></p>	<p><b>Lunch (not provided)</b></p>		

12.30-14.00	<b>Lunchtime Sessions for PhD students and young researchers</b>		
	<b>Open science practices for researchers: Dylan Burton.</b> <i>P673</i>	<b>How to get published: Claudia Harsch and Slobodanka Dimova.</b> <i>P111</i>	<b>How to write abstracts for conference calls: Benjamin Kremmel and Barry O’Sullivan</b> <i>P674</i>
14.00-15.30	<b>Parallel Paper Presentations</b>		
	<i>P111</i> <b>Chair: John Pill</b>	<i>P674</i> <b>Chair: Slobodanka Dimova</b>	<i>P673</i> <b>Chair: Richard Spiby</b>
	14.00-14.25 <b>Factors external to the rating criteria in Finnish high-stakes oral proficiency test</b> <i>Sari Ahola &amp; Mia Halonen</i>	14.00-14.25 <b>Developing assessment identity as a novice EFL teacher: A balance act between fairness and accountability</b> <i>Anna-Marie Csöreg</i>	14.00-14.25 <b>Beyond standard setting: Supporting score interpretation though mapping to proficiency levels</b> <i>Spiros Papageourgiou &amp; Venessa F. Manna</i>
	14.30-14.55 <b>Interpersonal affect and its relationship to language proficiency scores</b> <i>Dylan Burton</i>	14.30-14.55 <b>Affordances and limitations of incorporating video-vignettes as near-authentic stimuli in LSP oral performance tests</b> <i>Olivia Rütty-Joy</i>	14.30-14.55 <b>An innovative Method of Item Calibration: Use of the IRT Model in Standard Setting</b> <i>Merve Demiralp &amp; Yasmin Hatt</i>
	15.00-15.25 <b>Developing an automatic speaking proficiency assessment system for spontaneous L2 Finnish and Swedish on top of self-supervised speech models</b> Mikko Kurimo, Ragheb Al-Ghezi, Yaroslav Getman & Ekaterina Voskoboinik		15.00-15.25 <b>Considerations and challenges in setting up a retrospective DE&amp;I review of a high-stakes item bank</b> <i>Lottie Harrison &amp; Kasia Wojciechowska,</i>
15.30-16.00	<b>Coffee</b> <i>Aula</i>		

<b>16.00-17.30</b>	<b>Parallel Work-in-progress Presentations</b>		
	<i>P111</i> <i>Chair: Akihiro Matsuura</i>	<i>P674</i> <i>Chair: Toni Mäkipää</i>	<i>P673</i> <i>Chair: Gemma Bellhouse</i>
	16.00-16.25 <b>In search of a valid rating system for oral tests</b> <i>Gunn Inger Sture &amp; Sissel Randi Rong</i>	16.00-16.25 <b>Learning Oriented Assessment Approach in Pre-service Teacher Education Course</b> Merve Selcuk	16.00-16.25 <b>Exploring gender differences in the results of a language test</b> <i>Eli Moe, Hildegunn Helness Lahlum &amp; Norman Verhelst</i>
	16.30-16.55 <b>Speaking test sustainability through online administration</b> <i>Slobodanka Dimova</i>	16.30-16.55 <b>Teacher behaviour during speaking tests in beginners' classes</b> <i>Elisabeth Peyer, Gabriela Lüthi &amp; Nadia Ravazzini</i>	<del>16.30-16.55</del> <b>CANCELLED</b> <b>Assessment and inclusion in ELT training programs</b> <i>Norma Victoria Flores Martínez &amp; Tlaloc Domínguez-Navarro</i>
	17.00-17.25 <b>Assessing English as a Lingua Franca: Exploring the Implementation of Phonology Descriptors of the CEFR Using a New Intelligibility Measure</b> <i>Jingwen Wang &amp; Ying Zheng</i>	17.00-17.25 <b>Exploring novice item writing processes for classroom assessment</b> <i>Elisa Guggenbichler, Viktoria Ebner, Eva Konrad, Doris Moser-Frötscher, Simone Baumgartinger &amp; Kathrin Eberharter</i>	17.00-17.25 <b>Inclusive formative assessment practices (IFAP) in Higher Education: Promoting sustainable learning development</b> <i>Eleni Meletiadou</i>
<b>18.00</b>	<b>Social Event</b> <b>Walking tour starting from Senate Square. Free of charge.</b>		

# Conference

Saturday 17 June, 2023

9.00-10.30	Parallel Paper Presentations			
	<p><i>P111</i> Chair: Raili Hilden</p>	<p><i>P673</i> Chair: Dylan Burton</p>	<p><i>P674</i> Chair: Kaisa Hahl</p>	<p><i>P722</i> Chair: Eli Moe</p>
	<p>09.00-09.25</p> <p><b>How access to L1 speech affects L2 and L3 fluency assessment: Evidence from multilingual speakers</b></p> <p><i>Elina Lehtilä, Pauliina Peltonen &amp; Pekka Lintunen</i></p>	<p>09.00-09.25</p> <p><b>Using performance data as a lever for improvement</b></p> <p><i>Josep Maria Montcada Escubairó</i></p>	<p>09.00-09.25</p> <p><b>Remote proctoring in at-home language tests: Issues of Validity, Fairness, and Justice</b></p> <p><i>Benjamin Kremmel, Daniel R. Isbell &amp; Kim Jieun</i></p>	<p>09.00-09.25</p> <p><b>Jump-starting English in Chantilly: the impact of a ESOL learning and assessment programme for young learners in France</b></p> <p><i>Jonah Champaud &amp; Brigita Seguis</i></p>
	<p>09.30-09.55</p> <p><b>Exploring Repair Fluency Across Levels of Speaking Proficiency in the Aptis for Teens Speaking Test</b></p> <p><i>Akihiro Matsuura</i></p>	<p>09.30-9.55</p> <p><b>Language teachers' conceptions of the connection between sustainable development and assessment</b></p> <p><i>Minna Maijala, Päivi Laina &amp; Katja Mäntylä</i></p>	<p>09.30-9.55</p> <p><b><i>Moving towards minimum standards criteria for validating a high-stakes test in an at-home setting: preliminary findings</i></b></p> <p><i>Emma Bruce, Tony Clark &amp; Reza Tasviri</i></p>	<p>09.30-9.55</p> <p><b>Young language learners' beliefs about assessment in Swedish Compulsory school</b></p> <p><i>Ingela Finndahl</i></p>
		<p>10.00-10.25</p> <p><b><del>CANCELLED</del></b></p> <p><b><del>Local Tests, Local Contexts: The Italian language testing regime</del></b></p> <p><del><i>Sabrina Machetti, Lorenzo Rocca &amp; Giulia Paola Peri</i></del></p>	<p>10.00-10.25</p> <p><b>Marking what matters: The revision of a first-language speaking test for sign language interpreting students</b></p> <p><i>John Pill &amp; Tobias Haug</i></p>	<p>10.00-10.25</p> <p><b><i>Dealing with regional variety in standardized language exams</i></b></p> <p><i>Christine Troussart Van Bulck, Anne van Asseldonk, Sterre Turling, Manon Willems &amp; Goedele Vandommele</i></p>

**10.30-11.00**

**Coffee**  
*Aula*



11.00-12.30	Parallel Paper Presentations			
	<p><i>P673</i> Chair: Anna von Zansen</p>	<p><i>P111</i> Chair: Benjamin Kremmel</p>	<p><i>P722</i> Chair: Emma Bruce</p>	<p><i>P674</i> Chair: Voula Kanistra</p>
<p>11.00-11.25 <b>CANCELLED</b></p> <p><b>Simultaneous Diagnosis of L2 Listening Skills and Obstacles from Cognitive Diagnostic Approach</b></p> <p><i>Ya Wang &amp; Meng Ya-Ru</i></p>	<p>11.00 -11. 25</p> <p><b>Tracing pre-service language teachers' emerging language assessment literacy and item quality</b></p> <p><i>Doris Moser-Froetscher, Eva Konrad, Viktoria Ebner, Elisa Guggenbichler, Simone Baumgartinger &amp; Kathrin Eberharter</i></p>	<p>11.00 -11. 25</p> <p><b>Rethinking accessibility in the context of digital assessment: a review of evidence on the effectiveness of accessibility features</b></p> <p><i>Sylwia Macinska &amp; Carla Pastorino</i></p>	<p>11.00 -11. 25</p> <p><b>The effect of writing mode on the quality of young learners' writing tests</b></p> <p><i>Fauve De Backer &amp; Goedele Vandommele</i></p>	
<p>11.30-11.55</p> <p><b>Using interactive videos for assessing L2 listening: The effects of delivery modes and item formats</b></p> <p><i>Zhi Li, Ruslan Suvorov, Shanshan He &amp; Olga Kriukova</i></p>	<p>11.25-11.45</p> <p><b>Consulting pre- and in-service teachers: how their perceptions of difficulty inform sustainable course design promoting language assessment literacy</b></p> <p><i>Helen Heaney &amp; Armin Berger</i></p>	<p>11.25-11.45</p> <p><b>A neural transformer-based framework for complex text-based item generation</b></p> <p><i>Andrew Runge, Geoffrey T. LaFlair, and Yigal Attali</i></p>	<p>11.25-11.45</p> <p><b>Contribution of the CEFR-CV to developing a diagnostic rating scale for integrated writing</b></p> <p><i>Claudia Harsch, Valeriia Koval, Johannes Hartig &amp; Ximena Delgado</i></p>	
<p>12.00-12.25</p>	<p>11.50-12.25</p> <p><b>Enhancing teachers' and test takers' assessment literacy? Insights from test preparation for a digital high-stakes exam</b></p> <p><i>Sonja Zimmermann,</i></p>	<p>11.50-12.15</p> <p><b>Human judgement and technology in selecting texts for English for Academic Purposes Assessment</b></p> <p><i>Aylin Unaldi &amp; Bekir Ates</i></p>	<p>11.50-12.15</p> <p><b>Comparing traditional machine learning and deep learning approaches for automated essay scoring</b></p> <p><i>Erik Voss</i></p>	

		<i>Leska Schwarz, Anja Peters &amp; Günther Depner</i>		
<b>12.30-14.00</b>	<b>Lunch (not provided)</b>			
<b>14.00-15.30</b>	<b>AGM</b> <i>PII</i>			
<b>15.30-16.30</b>	<p><b>Coffee &amp; Poster Presentations</b> <i>Aula</i></p> <p><b>Secondary school students' and teachers' views on the usefulness of classroom speaking tests</b> <i>Gabriela Lüthi, Nadia Ravazzini &amp; Elisabeth Peyer</i></p> <p><b>Universal Screening of Multiple Language and Literacy Skills</b> <i>Miriam Semeniuk</i></p> <p><b>Strengthening Linguistic Interoperability among NATO and Partner Nations</b> <i>Merit Kompus &amp; Marju Laurits</i></p> <p><del><b>L2 writing assessment in a digital environment: lower-secondary school students' response processes when writing with and without full access to internet sources</b></del> <b>CANCELLED</b> <i>Dea Jespersen</i></p> <p><b>Meeting challenges in specifying and operationalizing the construct of teacher language competence: a balancing act</b> <i>Doris Moser-Froetscher, Lukas Bleichenbacher, Mirjam Egli Cuenat, Thomas Messerli &amp; Wilfrid Kuster</i></p> <p><b>Alternative Assessment in Language Learning (Case Study)</b> <i>Barbara Lesničar</i></p> <p><b>What would I do differently next time? Beyond Self-assessment in Presentational Writing in the L2 Classroom</b> <i>Borja Ruiz de Arbulo Alonso</i></p>			
<b>16.30-17.30</b>	<b>Parallel Work-in-Progress Presentations</b>			
	<i>PIII</i> <i>Chair: David Wei Dai</i>	<i>P674</i> <i>Chair: Karin Vogt</i>	<i>P722</i> <i>Chair: Eleni Meletiadou</i>	<i>P673</i> <i>Chair: Marita Härmälä</i>
	16.30-16.55 <b>Towards automated speaking</b>	16.30-16.55 <b>Teachers' perspectives on assessment in</b>	16.30-16.55 <b>Raters' cognitive load in assisted and</b>	16.30-16.55 <b>What factors support consistent</b>

	<p><b>assessment: L2 Finnish learner's perceptions of a computer-based speaking test</b></p> <p><i>Anna von Zansen &amp; Raili Hilden</i></p>	<p><b>foreign language education in Finland</b></p> <p><i>Toni Mäkipää Kaisa Hahl, Laura Lahti, Milla Luodonpää-Manni</i></p>	<p><b>unassisted rating of source-based writing</b></p> <p><i>Valeriia Koval</i></p>	<p><b>decision making in online standard setting contexts?</b></p> <p><i>Jane Lloyd</i></p>
	<p>17.00-17.25</p> <p><b>Research into the effect of test-taker personality on performance in the university test of English language (UTEL) in Sri Lanka</b></p> <p><i>Anton Pushparajah Tharamathan</i></p>	<p>17.00-17.25 <b>CANCELLED</b></p> <p><b>Different Stakeholder Groups, Different Language Assessment Literacy Requirements</b></p> <p><i>Ebtesam Abdalhaleem</i></p>	<p>17.00-17.25</p> <p><b>AI-based solutions in exam preparation classes: a silver bullet or conceptual mismatch?</b></p> <p><i>Anna Soltyska</i></p>	<p>17.00-17.25</p> <p><b>The added value of feedback for teachers in a digital assessment tool</b></p> <p><i>Jozefien Loman, Kris Van den Branden, Jolien Wouters, Goedele Vandommele &amp; Mariet Schiepers</i></p>
<p><b>19.00-22.30</b></p>	<p><b>Conference Dinner</b></p> <p><i>Hilton Helsinki Kalastajatorppa (Kalastajatorpantie 1)</i></p>			

# Conference

Sunday 18 June, 2023

9.30-11:00

**Symposium**

*PII*

*Chair: Martina Hulesova*

**Multilingualism, policy, assessment practice, policy and implementation**

*Barry O'Sullivan, Elana Shohamy, Anba Gani, Mina Patel, Lorenzo Rocca & Cecilie Hammes Carlsen*

11.00-11.30

**Coffee**

*Aula*

11.30-12.30

**Keynote**

*PII*

**Fairness in language assessment – Why do the purpose and context of assessment matter?**

*Ari Huhta*

*Chair: Claudia Harsch*

12.30-13.00

**Closing Ceremony**

*PII*

**Trip to Suomenlinna (self-funded)**