

**aea** '25  
EUROPE



---

05-08 November  
The Hague, Netherlands

---

# Contents

---

Theme .....	5
Message from the AEA-Europe President .....	7
Message from the Scientific Programme Committee (SPC) .....	8
Sponsors .....	12
Committees .....	14
Reviewers .....	15
Keynote Speakers .....	16
Workshops .....	20
Special Interest Groups .....	44
Programme .....	50
Social Events .....	75
About The Hague, Netherlands.....	77
Venue .....	79
Sponsors .....	80

# Theme

## Designing tomorrow's assessment landscape for positive impact on learning

Assessment and learning are fundamentally intertwined. It is widely acknowledged that the design and implementation of educational assessments directly and indirectly impact how teaching and learning are approached in the classroom. In high-stakes assessments like pre-university qualifications, final grades significantly impact students' future academic and career paths. These assessments also have major implications for teachers, schools and other stakeholders, and feed into broader issues of public trust in educational systems. Pressures and tensions associated with high-stakes assessments therefore inevitably create a backwash (or washback) effect on classroom-based activities, and risk practices of 'teaching to the test'. Assessment for Learning (AfL) influences very directly and intentionally the interactions between the teacher (assessor) and learner in the classroom. Other low-stakes assessments can also impact classroom practices. This is seen in international large-scale assessments – for example, with 'PISA shock' reverberating through national education policymaking and curriculum-based decisions in many countries. Thus, whether assessments are designed for summative or formative purposes, whether they are intended to reflect content in a particular curriculum, or to measure skills considered to be generalizable across different groups of students – assessment and learning are inexorably interconnected. Yet, despite the symbiotic relationship of assessment and learning, they are so often dealt with separately – in both theory and practice (Baird et al, 2017). Yet, if assessment and learning are not consciously integrated from the point of design, intended learning outcomes can be undermined. Thus, the theme for the 26th annual AEA Europe conference in 2025 focuses on: Designing tomorrow's assessment landscape for positive impact on learning. Several focal points within this overarching theme demonstrate the myriad of connections between and avenues for integration of assessment and learning.

The first focal point of the conference theme is: **Integrating assessment and curriculum design**. Assessment is the means for students to demonstrate their learning in relation to certain curricula. However, an artificial disconnect between assessment and learning can be exacerbated by conflicting priorities from curriculum and assessment specialists, as well as the lack of established, underlying theory to unite them (Newton, 2024). Despite the best intentions in assessment design, practices of 'teaching to the test' can risk superficial learning experiences, such as a disproportionate focus on rote memorization over critical thinking skills (Hamer & van Rossum, 2017). It can also result in increased anxiety, decreased motivation, and negatively impact students' wellbeing. The onus is therefore on assessment specialists to strive for a positive backwash effect, one example being consciously designing authentic assessments that provide opportunities to demonstrate learning outcomes that are valued. This can include encouraging the development of higher order thinking skills or supporting inquiry-based learning. This strand may include areas such as approaches to uniting theories of learning and assessment, designing curriculum-based assessments, backwash effect, and assessing authentic skills with real-world application. With the recent explosion of generative AI, this latter area in particular also connects with emerging technologies, which is the focus of the next strand.

The second focal point of the conference theme is: **AI and authentic assessment**. Here, the 2025 conference theme also builds on the AEA-Europe 2024 theme: 'Advances in Educational Assessment Practices: Considering the use of Technology, Artificial Intelligence, and Process Data for Assessment in the 21st Century'. In the rapidly evolving world of digital technologies, challenges and opportunities related to AI and large language models (LLMs) are a reality faced by teaching and learning settings around the world. As authentic assessments are intended to reflect real-life learning, the impact of AI in the classroom cannot be ignored in assessment design. Concerns have naturally been raised about authenticity of student work, and the risks to academic honesty and malpractice issues posed by easily available AI tools, so new approaches are needed to determine the veracity of assessed work. However, emerging technologies also offer new opportunities, such as more adaptive, personalised assessment approaches and the

# Theme

inclusion of interactive elements and simulations requiring more integration with the learning itself (Bennett, 2015). New and innovative assessment designs may also allow students to engage critically with AI generated materials, such as by encouraging them to fact-check sources, investigate possible AI hallucinations, and question biases that may be perpetuated or reinforced in AI-generated content. Assessment task design can play an important role in these aims, such as by encouraging critical thinking skills, and the ethical use of AI. Carefully designed summative assessments, that promote critical thinking skills, awareness of inherent biases in and limitations of AI-generated content, will equip learners for the future where AI and AI generated content will be embedded in daily life. Similarly, there is a need for research on and development of guidance and frameworks for educators on how to design formative tasks for AI use and embed the ethical AI practices in the classroom.

However, it is also important to recognize the limitations of AI tools in educational contexts, such as inconsistencies in the quality and availability of resources in LLMs across different languages. This brings us to the third and final focal point: **Inclusive assessments for all learners**. Educational assessments offer social value in that they can contribute to the improvement of student learning, whilst also contributing to a more equitable society (Sireci, 2021). However, to do so, they must be underpinned by core values, such as recognizing that there are no inherent differences in the capacity to learn across different groups, such as cultural, linguistic or racial (ibid). The final strand therefore focuses on designing assessments that are inclusive and valid for all groups of learners. This may include exploring validity in cross-lingual and cross-cultural assessment, such as how exams are developed for use in different languages, and how scores are evaluated and compared across different language versions. Research into accommodations that can support student learning in the classroom as well as in assessments would be relevant here, as would universal design approaches that aim to reduce barriers for all learners. This may also include exploring how systematic biases against subgroups of students (e.g., race, gender or neurodiverse) can be avoided in assessments design, thereby supporting and creating a positive impact on the learning experience and outcome for all students.

## References

- Baird, J. A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: Fields apart? *Assessment in Education: Principles, policy & practice*, 24(3), 317-350.
- Bennett, R. E. (2015). The changing nature of educational assessment. *Review of Research in Education*, 39(1), 370-407.
- Hamer, R., & Van Rossum, E. J. (2017). Students' conceptions of understanding and its assessment. In *Innovative practices for higher education assessment and measurement* (pp. 140-161). IGI Global Scientific Publishing.
- Newton, Paul (2024). (In search of) Educational Certification Theory. The 24th annual meeting of the Association for Educational Assessment-Europe (AEA-Europe). 6-9 November, 2024. Paphos, Cyprus.
- Sireci, S. G. (2021). NCME presidential address 2020: Valuing educational measurement. *Educational Measurement: Issues and Practice*, 40(1), 7-16.

# Welcome Note

## Message from the President

Dear Colleagues

Van harte welkom op de AEA Europe Conferentie van 2025 en in de prachtige stad Den Haag.  
A warm welcome to the AEA Europe Conference 2025 and to the beautiful city of The Hague.

AEA Europe was established on 3 April 2000. For over a quarter of a century, the Association has provided a platform for the sharing and discussion of ideas by individuals and organisations involved in educational assessment, serving an ever-expanding constituency in Europe and further afield. This interchange of ideas supports high-quality research and practice in assessment. Thank you for registering for our 26th Annual Conference. I do not underestimate the effort made by individual and corporate members and other delegates in attending in The Hague and I commend your generosity in sharing your ideas and your valuable time with us.

This year's conference theme is "Designing tomorrow's assessment landscape for positive impact on learning". It recognises that assessment and learning are closely intertwined. In a world characterised at present by much instability, social upheaval, and conflict, we as an assessment community can leverage the tools and practices at our disposal for societal benefit. The work of delegates attending the conference impacts directly or indirectly on learners and on the activities of professionals who support learners, in educational/workplace settings or in the broader policy arena. Our theme encourages developers, users and consumers of assessments to adopt a holistic view of assessment as integrated with, rather than separate to, learning. Debates about the different purposes of assessment abound, but underpinning all purposes is the realisation that assessment can and does influence teaching, learning, and curriculum. As a cross-national community of assessment professionals we need to champion an overarching expectation that assessment is a positive variable in education and training. This year's theme provides ample opportunity for delegates to interrogate the relationship between assessment and learning and the implications for a range of stakeholders and for society as a whole.

I hope that you feel enriched by the conference programme and by the opportunity to renew friendships. I extend a special welcome to delegates joining us for the first time - I hope you find us to be a vibrant, warm and welcoming community. We look forward to getting to know you. We thank our sponsors for showcasing their work and services at the conference and I encourage delegates to meet with our sponsors/exhibitors during the event.

As I conclude my first year as President of our Association, I have been reflecting on what brings us together as a community. Whereas the conference is our flagship event annually, AEA Europe offers a range of activities and benefits for members throughout the year.

- Special Interest Groups (SIGs) meet regularly and offer topical, informative online webinars that provide professional development opportunities for participants located anywhere in the world. You are invited to attend brief presentations by SIG committee members during our first conference coffee break on Thursday 6 November (10:15 – 10:45).
- Our website hosts blog discussions posted by members, offering opportunities to share emerging ideas and generate discussion in an informal space.
- AEA Europe's professional accreditation programme allows members to have their assessment contribution recognised by their peers. Accreditation is offered at 3 levels - Associate, Practitioner and Fellowship - so there is a level appropriate for everybody's stage in career. 13 members have achieved accreditation this year and will receive their certificates during the conference.

# Welcome Note

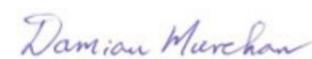
- A number of prestigious awards are presented annually, including the Kathleen Tattersall New Assessment Researcher Award and the AEA Europe Poster Award.

These and other benefits of membership provide opportunity to develop your interest in assessment within a supportive, friendly community of researchers, practitioners, policymakers, postgraduate students and others with interest in assessment. So, in addition to attending our conference, I invite you to renew your membership, or if you are not yet a member, consider joining. Check details on our website <https://aea-europe.net/> Members can also avail of discounted rates to register for the annual conference.

Many people worked hard to plan the 2025 conference. I thank especially our colleagues in the International Baccalaureate (IB) for hosting us in The Hague. My sincere appreciation goes to Rebecca Hamer, Antony Furlong, Vanessa Scherman, Shehzad Jeeva and their other colleagues in the IB for giving us the opportunity to meet in The Hague. I extend thanks also to the Scientific Programme Committee, the Conference Organising Committee, our conference partners EasyConferences, keynote speakers, workshop presenters and all presenters and session chairs. Thanks also to the many individuals who reviewed conference proposals, accreditation applications and award nominations.

Finally, I am humbled by the dedication of so many AEA Europe members to the work of the Association on a continuing basis. In this I include my colleagues on the Council and members of the Professional Development Committee, Communications Committee, Audit Committee, Secretariat (AQA) and SIG Steering Groups. They all give very generously of their time and expertise throughout the year, for which I am extremely grateful.

I wish you a stimulating, enjoyable conference.



**Damian Murchan**  
President AEA Europe

## Message from the Scientific Programme Committee (SPC)

Dear Conference Delegates

The preparatory work of the Scientific Programme Committee (SPC) for the 2025 annual conference began in earnest in early December 2024. Coordinating construction of the programme can be a complex business and involves close collaboration between the SPC and the Council; the Conference Organising Committee; the Local Organising Committee; the Publications Committee; the Professional Development Committee; and the Special Interest Groups. A truly collective endeavour!

This year's conference theme takes as its focus Designing tomorrow's assessment landscape for positive impact on learning. The 2025 conference programme embraces a full range of varied submission formats including keynote addresses, open paper presentations, discussion groups, symposia, workshops, and poster presentations on topics relating to this theme.

The 2025 conference submission data reflects both Europe-wide and international interest with nearly 250 high-quality proposals. The number of participating countries making submissions this year has increased to 36 (from 28 countries in 2024).

The overriding responsibility of the SPC is to ensure the preparation and production of a quality, balanced and timely programme for the November annual conference. On occasions, however, such a responsibility can be challenging. For example, the number of proposals submitted for the 2025 annual conference have again exceeded our expectations and, although very encouraging, has created a raft of related issues. In 2024, the SPC received a total of 231 submissions (across all formats). The total for this year has been 242.

Increased submissions in recent years have impacted the reviewing process. Proposals are considered through a blind peer review process and - before this year - each proposal has been reviewed by three independent reviewers. The shift away from 'three' reviewers to 'two' reviewers this year has eased the management of the reviewing process, resulting in fewer independent observations.

	2022	2023	2024	2025
Number of reviews	459	591	641	444
Number of reviewers	46	60	74	61

Whilst the SPC is indebted to all those volunteers for their willingness to review and for their hard work, recruitment has continued to be an issue. (The SPC welcomes further requests from members to join the reviewing team for next year.) At the request of Council at a Meeting in Cyprus (April, 2024), the SPC was instructed to switch from three to two reviews per submission for future conferences (2025 and beyond). Such a move has eased pressure on the annual recruitment of reviewers. Following extensive deliberation, the SPC agreed on a '2-Reviewer' model to be implemented for this year's conference.

The revised evaluation criteria for judging the quality of submissions - introduced in 2022 and which were intended to be used to evaluate both conceptual/theoretical and empirical proposals, are now fully embedded in the evaluative process. The new criteria appear to be working well.

# Welcome Note

Scores	Evaluation Criteria
0 - reject	Unacceptable scholarly/scientific quality
1 - poor	The content, style, examples, and/or description of previous work, conceptual ideas or results need further development or represent a work in progress.
2 - fair	Borderline submission of uneven quality – in other words, some aspects are well-developed, and others not so well.
3 - good	A good submission that would benefit from some changes to strengthen its conceptual/theoretical, technical or presentation qualities.
4 - very good	A very good quality submission that would benefit from some minor changes in preparing the presentation.
5 - excellent	This submission represents excellent conceptual/theoretical and/or technical work.

The evaluation criteria for pre-conference workshops were also revised by the SPC and used for the first time this year.

Criterion	Description
Relevance	<ul style="list-style-type: none"> <li>Offers a relevant topic, including a fresh perspective on an existing topic, important to assessment</li> <li>Displays scientific rigour and up-to-date insights</li> </ul>
Appeal to delegates	<ul style="list-style-type: none"> <li>Has appeal for delegates</li> <li>Can expect a 'reasonable' number of registrations</li> </ul>
Quality of submission	<ul style="list-style-type: none"> <li>Underlying rationale/argument is sound</li> <li>Proposal is clear and well-written</li> </ul>
Adherence to submission instructions	<ul style="list-style-type: none"> <li>Submission adheres to instructions described in the 'Call and instructions for submission'</li> <li>submission consists of a short abstract (200-word max.) and fully and accurately completed 'Appendix A'</li> </ul>
Balance (between presentation and participant interaction activities)	<ul style="list-style-type: none"> <li>Participants have opportunities to engage in ample discussion and relevant activities throughout the workshop.</li> </ul>

The 2025 conference submission data reflects both Europe-wide and international interest with over 240 high-quality proposals (a number greater than previous totals) from 36 countries (eight more than in 2024).

The SPC continues to face new and interesting challenges in constructing the conference programme especially given the increasingly larger number of high-quality presentation proposals received and accepted each year (an anticipated trend going forward). Given that the conference structure and duration of the conference has remained the same over the years, internal features of the programme have been reconfigured to accommodate potentially higher numbers of presentations. Main changes to the internal structure of the programme for 2025 include:

- In 2024, the poster session was divided in two across two successive days in parallel with (for the first time) other presentation formats. In addition, two 'presenter' sessions were built into the conference programme (including during the coffee break immediately after each poster session). This was in direct response to the feedback from previous conferences i.e. concerns expressed by conference attendees that poster presenters were not always available to speak to at their poster. However, in 2024 feedback from the conference survey strongly indicated that conference attendees did not welcome these changes. Accordingly, the 2025 programme has returned to the one poster session format of 75-minute duration on the first day. The coffee break presenter sessions have been retained, however. Additionally, the duration of each poster presentation has been reduced from 90 seconds to 60 seconds to accommodate the increased number of submissions accepted. Thus far, there have been no reports from poster presenters that the reduced time is problematic.
- As a consequence of changes made this year to the poster session, Discussion Groups are now no longer located in parallel with the first of the two poster sessions. The Discussion Groups sessions have returned to their original format.
- By agreement with Council, the Ignite session has been removed from the conference programme this year. As this session has traditionally been in parallel with the symposia, the programme remains unaffected by the loss.
- All other changes made to the 2024 conference programme (e.g. keynotes on each day, KTNRA winner keynote on the final day before the 'Closing Ceremony incl. Poster Award & Accreditation Awards') were encouragingly embraced by conference attendees and have, therefore, been retained.

Finally, given that some of the existing SPC will no longer participate in future SPC activities (e.g. some members of the SPC representing conference hosts), we would welcome interest from other AEA-Europe members who are keen to join the SPC.

In the meantime, we hope that you and your colleagues both enjoy, and engage with, the conference programme to the fullest extent.



Stuart Shaw (Chair of the SPC)

# Sponsors

## Host Institution

## Sponsors



## Coordinator



ORGANIZATION	SPONSORSHIP TYPE
City of The Hague	Supporter
IEA (International Association for the Evaluation of Educational Achievement)	Welcome Reception
AQA	Kathleen Tattersall New Assessment Researcher Award & Exhibitor
GA Partnership	Networking Coffee Break
Renaissance Global	Exhibitor
Pearson	PhD Fellows Event
Inspera AS	Exhibitor
Cambridge University Press & Assessment	Sponsoring the future of assessment research
Duolingo	Sponsoring the future of assessment research
CITO	AEA Europe Poster Award
International Baccalaureate	Host Institution
Learnlab	Exhibitor

# Committees

## Council Membership

Damian Murchan, President AEA-Europe (Professor in Education at the School of Education, Trinity College Dublin)  
 Cor Sluijter, Vice President AEA-Europe (Independent Educational Measurement Specialist)  
 Sebastian de Klerk, Treasurer (Director of Cito International, Netherlands)  
 Alex Scharaschkin, Executive Secretary AEA-Europe (Director of Research and Regulation at AQA)  
 Linda Wye, AEA-Europe Secretariat (AQA)  
 Deborah Chetcuti, Non-Executive Member (Professor, Faculty of Education, University of Malta)  
 Elena Papanastasiou, Non-Executive Member (Dean of the School of Education at the University of Nicosia),  
 Lesley Wiseman-Orr, Non-Executive Member (Research Fellow at the School of Education University of Glasgow)  
 Stuart Shaw, Non-Executive Member (Honorary Professor of University College London in the Institute of Education – Curriculum, Pedagogy & Assessment)

## Conference Organising Committee/Local Organising Committee

Damian Murchan (Chair), Trinity College Dublin (AEA Europe President)  
 Cor Sluijter, Vice- President AEA-Europe  
 Stuart Shaw, Chair of SPC (Institute of Education, University College London), UK  
 Elena Papanastasiou, University of Nicosia  
 Linda Wye, AEA-Europe Secretariat (AQA)  
 Rebecca Hamer, International Baccalaureate, The Netherlands  
 Antony Furlong, International Baccalaureate, The Netherlands  
 Vanessa Scherman, International Baccalaureate, The Netherlands  
 Petros Stratis/EasyConferences, Cyprus

## Scientific Programme Committee

Anthony Furlong (IB, Netherlands)  
 Angela Verschoor (CITO, Netherlands)  
 Dario Pirotta (University of Malta, Malta)  
 Doreen Said Pace (Ministry for Education, Sports, Youth, Research and Innovation, Malta)  
 Michalis Michaelides (University of Cyprus, Cyprus)  
 Rebecca Hamer (IB, Netherlands)  
 Chair: Stuart Shaw (Institute of Education, University College London, UK)

## Communications Committee

Deborah Chetcuti (Chair) (University of Malta)  
 Amina Afif (Pedagogue, Luxembourg)  
 Judith Gracey (AQA, UK)  
 Jeanne Marie Ryan (AQA, UK)

## Professional Development Committee

Elena Papanastasiou (Chair) (University of Nicosia, Cyprus)  
 Bas Hemker (CITO, Netherlands)  
 Shakeh Manassian (PSI Services, UK)

## Audit Committee

Graham Hudson (GA Partnerships, UK)  
 Fazilat Siddiq (University of South-Eastern, Norway)  
 Guri Nortvedt (University of Oslo)

# Reviewers

Alex Scharaschkin	Mariana Manzo
Andrew Boyle	Marte Sove Syverud
Andrew Watts	Mary Richardson
Angela Verschoor	Matt Glanville
Anna Lind Pantzare	Michael A. Buhagiar
Antony Furlong	Michalis Michaelides
Catarina Correia	Militsa Ivanova
Cate Sparkes	Mkululi Wami
Christina Wikstrom	Nathalie Younes
Christoph Schneider	Newman Burdett
Claudine Muscat	Nico Dieteren
Cor Sluijter	Parveen Sandhu
Dario Pirotta	Paul Newton
Dina Tsagari	Rachel Cunningham
Doreen Said Pace	Raphaël Pasquini
Eleanor Andressen	Rebecca Chivers
Elena Papanastasiou	Rebecca Engledew
Fabrizia Flynn	Rebecca Hamer
Francois Zammit	Rolf Vegar Olsen
Fraser Lawrence	Rose Clesham
Grace Grima	Ryan O'Connor
Harald Eriksen	Sam Passeport
Helen Griffiths	Sebastiaan de Klerk
Iasonas Lamprianou	Stuart Shaw
Isabel Nisbet	Vanessa Scherman
Jana Strakova	
Kay Evans Johns	
Laura Gardner	
Lenka Firtova	
Lesley Wiseman-Orr	
Lise Vikan Sandvik	
Louise Badham	
Marcus Hjarne	
Maria Galahe Ramos	

# Keynote Speakers



*Prof. Jo-Anne Baird*  
University of Oxford

**Title:** Raising educational standards

**Chair:** Stuart Shaw  
Friday, 7th November, 11.00 - 11:45 GMT

**Abstract:** A perennial challenge for the field is how assessments can support learning and thereby raise educational standards. Part of the answer to this must involve clear messaging about what the educational standards are, to provide goals for teachers and learners.

In this address, I will consider how standards are thought about differently in three assessment paradigms: construct-, curriculum- and outcomes-based approaches. I have previously characterised these paradigms as prototypical ways of thinking about assessments, since many assessment systems are hybrid versions, borrowing across paradigms. In each case, the philosophy of assessment differs in terms of the attribute of interest, the definition of standards and expectations of outcomes. This creates different relations between what kind of thing is valuable learning, the ways in which we think they ought to be assessed and which quality criteria we prioritise. Different paradigms promise information of different kinds about standards for teachers and learners.

In research conducted with over 900 stakeholders in Scotland during the pandemic, we investigated how stakeholders think of national assessments and their standards. One of our main findings was that they think of the same assessments in relation to different paradigms. I will outline how this raised significant issues for the management and communication of standards by exam boards, that go beyond misunderstandings or lack of assessment literacy. In a smaller-scale qualitative study in Wales, we investigated how industry-insiders and teachers thought about standards and their communication. A striking finding from this study was that not only were teachers unaware of the standard setting processes, but they did not believe that they needed to know. This raises questions about exactly what teachers and learners need to know about standards and whether our industry-insider perspective has been miscommunicating.

**Bio:** Professor Jo-Anne Baird is Professor of Educational Assessment and Director of the Oxford University Centre for Educational Assessment. Her research addresses systemic aspects of educational assessment, and she works extensively with governments and assessment organisations to support the development and evaluation of national and international assessment systems.

Professor Baird has played leading roles in major reviews of qualifications in Wales, Scotland, and England. She holds a PhD in Educational Assessment and has also served as President of the Association for Educational Assessment-Europe, as well as Director (Dean) of the Department of Education at the University of Oxford.

She currently serves as Co-Editor of the Oxford Review of Education, one of the field's premier journals, and is a Standing Advisor to the UK Parliament's Education Select Committee. A former member of the influential Assessment Reform Group, Professor Baird has helped shape global policy on formative assessment. She is a Fellow of the Academy of Social Sciences and holds an Honorary Doctorate from the University of Bergen.

Her recent research explores assessment standards, the resilience of qualification systems, international tests, policy dynamics, and the paradigms that frame educational assessment.



*Dr. Matt Glanville*  
International Baccalaureate

**Title:** All the World Is a Stage – Enter the Androids: Navigating Current and Future Challenges of International Baccalaureate Assessment

**Chair:** Damian Murchan  
Thursday, 6th November, 09:30 – 10:15 GMT

**Abstract:** The International Baccalaureate (IB) has an unusual role in aiming to offer an educational philosophy and approach for the whole world. In assessment terms, this presents the IB with a number of challenges – some of which will be familiar to national assessment systems and some which are unique. After providing a brief overview of the IB and its qualification I will reflect on these unusual and interesting matters.

However, the assessment landscape is not standing still and developments in technology and society are impacting on many of our foundations of education. In the second half of my presentation, I will talk about how the IB is thinking about these changes (AI being key amongst them) and speculate on what the known unknowns may be for future assessment.

**Bio:** Dr Matt Glanville is the Director of Assessment at the International Baccalaureate (IB), where he is responsible for the assessment models across the Middle Years Programme, Career-related Programme, and Diploma Programme. These models serve more than 250,000 students annually, in over 150 countries and across more than 250 subjects.

Dr Glanville holds a doctorate and began his professional life as a mathematics teacher in the United Kingdom. He later led research teams for regional and vocational skills agencies and held key positions at the Qualifications and Curriculum Development Agency (QCDA) and Ofqual, the regulator of qualifications in England.

Since joining the IB over nine years ago, Dr Glanville has contributed to assessment innovation and strategic reviews, including the reform of 16–19 education pathways. His interests span broader educational issues, and he is recognised for promoting high-quality, equitable, and forward-looking assessment practices within global education systems.

# Keynote Speakers



*Dr. Militsa Ivanova*

*Winner of the Kathleen Tattersall New Researcher Award*

**Title:** Who Tries and When in the Digital Age: Measuring and Modeling Test-Taking Effort through Process Data in Large-Scale Assessments

**Chair:** Elena Papanastasiou, AEA Europe Council Member  
Saturday, 8th November, 11:45 – 12:30

**Abstract:** Achievement tests aim for valid estimation of proficiency, but inadequate test-taking effort can introduce construct-irrelevant variance and threaten score validity, particularly in low-stakes contexts. Since 2010s, international programs like the Programme for International Student Assessment (PISA) have tried to measure effort, mainly through self-reports. With the digitalization of large-scale assessments, the large amount of process data now allows for tracing examinees' test-taking behavior, avoiding the biases of self-reports. Process data variables, such as response time, have proven valuable in estimating examinee effort on multiple-choice items, but research on constructed-response items – often linked to lower effort – remains limited. The literature on how thresholds can be determined to distinguish effortless from effortful responses also remains inconclusive. While individual characteristics have been widely studied as predictors of effort, family- and school-level factors have not. Cross-national differences in effort and its relationship to performance are evident, but predictions of effort across countries are yet to be explored in depth. This presentation will discuss indicators, thresholds, and predictors of test-taking effort using process data.

**Bio:** Militsa G. Ivanova earned her Ph.D. in Clinical Psychology from the University of Cyprus in June 2024. She also holds a BS in Social and Behavioral Sciences with a major in Psychology from European University Cyprus and a MA in School Psychology from the University of Cyprus. She is a licensed School Psychologist in Cyprus. During her graduate studies, she taught research methodology and contributed to several funded projects focused on student motivation, test-taking engagement, and responders' behavior in large-scale assessments such as the Trends in International Mathematics and Science Study (TIMSS) and the Program for International Student Assessment (PISA). In 2019, she completed an ERASMUS+ funded internship at the Umeå University in Sweden, where she expanded her research skills and statistical expertise.

Her research interests include motivation, effort, and engagement in assessments and surveys, as well as individual, contextual, and item-level factors influencing test-taking behavior. She has experience working with large-scale international databases and process data, and her work was presented at national and international academic conferences and published in peer-reviewed journals. In November 2024, she was awarded a competitive postdoctoral fellowship from the Cyprus Research and Innovation Foundation (Restart program, 2025–2027) to investigate innovative multiple-choice item formats and their impact on student engagement and performance.



*Prof. Jennifer Randall*

*University of Michigan*

**Title:** Assessment as a Tool for Liberation: Come Dream with Me

**Chair:** Cor Sluijter, AEA Europe Vice President  
Saturday, 8th November, 09:00 – 09:45

**Abstract:** This talk reimagines educational assessment as a transformative force for justice and liberation, rather than a mechanism of oppression and marginalization. Drawing on critical pedagogies of care (Noddings) and discomfort (Boler), I illuminate how assessment systems can elevate learners' critical consciousness, respond to the expressed needs of marginalized communities, and affirm diverse ways of knowing. Through personal narrative (i.e. storytelling) and scholarship, I describe the historical and ongoing harms perpetuated by conventional assessment practices that center whiteness; and advance a liberatory framework for assessment that disrupts these practices. I propose a set of principles, framed within a shared responsibility for justice, that center the voices and lived experiences of rights-holders and actively seek to disrupt structures of inequity. Ultimately, I invite my colleagues to dream with me about assessment systems that not only measure academic achievement, but actively foster growth, social justice, inclusion, and liberation for all learners.

**Bio:** Professor Jennifer Randall is the Dunn Family Chair of Psychometrics and Test Development in the School of Education at the University of Michigan, and the founder of the Center for Measurement Justice. Her work pioneers the development of justice-oriented educational assessment, challenging traditional paradigms that often marginalize underserved student populations.

Professor Randall holds a Bachelor's degree (1996) and a Master's degree (1999) from Duke University, and earned her PhD from Emory University in 2007. She began her career as a public high school social studies teacher, working with racially and ethnically minoritized students. These early experiences catalysed her critical examination of assessment practices and their impact on equity in education.

Her scholarship seeks to disrupt white supremacist logics embedded in traditional assessment by advancing practices that are explicitly and unapologetically antiracist. Professor Randall's work centres the sociocultural identities of learners—not as an afterthought, but as a foundational principle in the design and implementation of assessments. She continues to collaborate with minoritized communities and allies to reimagine assessment systems that are equitable, inclusive, and justice-driven.



### Workshop 1

**Presenters:** Stuart Shaw and Ezekiel Sweiry

#### **Identifying and analysing evidence to determine whether tasks elicit the intended constructs: bridging the gap between modern validity theory and innovative validation practice**

Establishing that assessment tasks elicit performances that reflect the intended constructs is a fundamental component of assessment validation. However, literature on this question is dominated by theoretical perspectives, with little practical guidance for practitioners. In addition, recent developments in the assessment landscape, including the emergence of technological advances (e.g. process data) and novel forms of assessment (including 21st-century skills such as collaboration and self-reflection), mean that established guidance may not adequately address contemporary requirements.

Through a combination of presentations, discussion and group activities, this workshop explores the challenges in identifying and analysing validation evidence to determine whether assessment tasks elicit performances that reflect the intended constructs. Participants will explore a range of validation evidence sources, from traditional statistical and qualitative approaches to innovative methods like process data and eye-tracking. Attendees will work through real-world scenarios, critically examining the alignment between task design and cognitive processes, and considering the implications of emerging technologies on validation practices.

Participants will leave with a deeper understanding of how to strengthen their validation arguments, navigate trade-offs in assessment design, and leverage both classic and emerging techniques to enhance validity in diverse educational contexts.

**WORKSHOP TITLE:** Identifying and analysing evidence to determine whether tasks elicit the intended constructs: bridging the gap between modern validity theory and innovative validation practice.

#### **Presenters' Bios** (500 words max per presenter)

**Stuart Shaw** is an educational assessment consultant, researcher, and author, and is Honorary Professor of University College London in the Institute of Education – Curriculum, Pedagogy & Assessment. Before he became an independent consultant, he worked for international awarding organisations for over 20 years where he was particularly interested in demonstrating how educational, psychological, and vocational tests seek to meet the demands of validity, reliability, and fairness. He has a wide range of publications in English second language assessment and educational/psychological research journals (around 150 publications). He is Chair of the Board of Trustees of the Chartered Institute of Educational Assessors (CIEA) and a Fellow of the CIEA. Stuart is a Fellow of the Association for Educational Assessment in Europe (AEA-Europe), an elected member of the Council of AEA-Europe, and is Chair of its Scientific Programme Committee.

**Ezekiel Sweiry** is an Associate Director at Ofqual, a non-governmental department that regulates qualifications, examinations and assessments in England. He has over 25 years' experience in assessment, with the majority of those years spent in paper-based and digital test development and item validity research. He has worked for the three largest UK awarding bodies, where he has been involved in the development of a range of high-stakes assessments and responsible for the training of examiners. His particular research interests include the factors that affect the difficulty and accessibility of test items, the item and mark scheme features that affect marking reliability, and the comparability of paper-based and digital assessments.

#### **Why AEA members should attend this workshop:**

Despite the primacy of validity as a theoretical concept and the increasing number of conceptual frameworks designed to guide validation effort, there is relatively little in the way of practical guidance for would-be validators: good validation studies still prove surprisingly challenging to conceptualise, let alone to implement (Lissitz, 2009). Accordingly, calls for a more pragmatic approach to validation abound (Kane, 2013; 2006).

One key validation research question, often given scant attention due to the challenge it presents, seeks to determine whether assessment tasks elicit the intended test constructs. Corresponding evidence garnered in support of such a claim of validity is often limited to marker reports or basic consideration of item level statistics. For example, candidate response process validation evidence[1] is relatively scarce (Hubley & Zumbo, 2011; Zumbo & Chan, 2014) – though this situation is changing with the advent of increasingly more sophisticated sources of assessment evidence (Zumbo, 2017). This challenge has arguably resulted in a lack of practical guidance, both in terms of the nature of the post-assessment evidence that can be collected, and the ways in which it might be analysed.

As well as this lack of practical guidance, the question of determining whether assessment tasks elicit the intended test constructs is particularly pertinent for two further reasons. First, recent technological advances in assessment have made available a number of more novel evidence sources, such as candidate process data and eye tracking. These sources, derived from richly situated learning environments, have the potential to offer a more fully developed articulation of “assessment” and enhance the capacity to make more valid and accurate decisions about student learning and pedagogy.

Second, the question merits increased attention due to the emergence of more novel forms of assessment. 21st century skills (e.g. collaborative problem-solving, creativity, critical thinking, technology literacy and decision-making), “hard-to-measure” complex, interactive performances, classroom-based formative assessment, and computer-based testing is making the process of validation a more multi-faceted, complex endeavour.

This workshop takes as its focus a very practical orientation, exploring challenges in identifying and analysing validation evidence that seeks to address the validity question: Do the assessment tasks elicit performances that reflect the intended constructs? In doing so, the workshop attempts to bring various different validation techniques to life.

#### **Who this workshop is for:**

The responsibility for assessment providers to demonstrate robust and thorough validity evidence is a long-established expectation (Messick, 1992, p.89), as are warnings about the “potentially serious consequences” (Kane, 2009, p.61) of shirking such responsibilities. Even assessment providers and users of such assessments that have limited resources will still have a responsibility to demonstrate the quality and validity of their assessments. This workshop is intended to make the complexities of validation theory and practice less challenging and more readily operational. Accordingly, the workshop is relevant to many key actors in educational assessment. Test developers who specify the content and format of tasks; examination officials frequently working within statutory agencies or approved assessment providers/awarding bodies; policy makers – those charged with formulating, leading and funding assessment; and to researchers and postgraduate students with a wide variety of assessment interests will find this workshop of great relevance.

# Workshops

## Workshop 1

### Overview of workshop (max. 600 words):

The workshop divides into four parts, each comprising sessions that afford group activity, reflection, and discussion.

**Part 1** offers a brief introductory overview of assessment validation. Important questions are raised which will punctuate the workshop throughout, including:

- How is it possible to judge whether a qualification has necessary and sufficient validity? How much validity is sufficient validity?
- What is the relationship between validity and validation?
- What makes for a sound validity/validation argument?
- What is validation evidence and what do we mean by analysing validation evidence?

**Part 2** offers an overview of sources of validation evidence, categorised into four broad types: Statistical, Qualitative, Documentary, and Innovative. Example approaches from each category will be presented, with a focus on how each evidence source can be analysed and interpreted to determine whether tasks elicit performances that reflect the intended constructs.

In **Part 3**, participants will be presented with several different scenarios based on four sources of innovative validation evidence:

- ‘Hard-to-measure’ skills (such as ‘collaboration’ and ‘self-reflection’), which refers to the personal, social, and emotional attributes, skills and dispositions that enhance cognitive abilities and academic performance.
- Impact (in English second language assessment), which refers to the consequential aspects of validity. Consequential validity focusses on three main areas: differential validity, washback and effects on society. In other words, the effects and consequences within the educational system and within society more widely.
- Process data, which refers to any data automatically collected about test takers’ response processes, such as response times and the actions performed by the test-taker (e.g. keystrokes, mouse clicks, selection, and navigation).
- Eye tracking data, which refers to the information gathered from tracking the test taker’s gaze as they interact with assessment material, and includes metrics such as fixation points (where the eyes stop to focus) and saccades (quick eye movements between fixations).

For each of the four scenarios, attendees will have the opportunity to analyse and interpret authentic assessment data. Working in groups, they will determine what can be concluded about the cognitive processes that are elicited, and the extent to which those processes align with the intentions of the assessment. Attendees will also have the opportunity to reflect on the potential and limitations of each type of evidence in relation to eliciting evidence of construct validity, as well as the barriers (e.g. costs, resources) to utilising each type of evidence.

**Part 4** explores issues relating to:

- How empirical evidence and logical analysis are used to test the claims in a validation argument in order to evaluate its overall strength,
- How to identify the kinds of evidence and analysis that can be used either to falsify or to support a measurement argument claim,
- How we know that a qualification is measuring what it is supposed to be measuring, and
- How traditional (classic) and innovative techniques can be combined.

The discussion will broaden to encompass future directions of validation methodology including the use of digitised technologies such as Artificial Intelligence.

The workshop will conclude with discussion activities relating to fundamental validation concerns such as:

- why it is very hard to judge (a) what kind of validation evidence is necessary and (b) how much is sufficient;
- the inclusion of consequences of testing as a legitimate source of validation evidence or analysis;
- ways to optimise validity;
- inevitable assessment design trade-offs and compromises and their impact on validity;
- ‘validation-by-design’ versus ‘validation-of-design’ and, the relationship between validity and reliability, fairness and practicality.

### Preparation for the workshop:

Useful (though not essential) preparatory reading:

Cook, D. A., & Hatala, R. (2016). Validation of educational assessments: a primer for simulation and beyond. *Advances in Simulation*, 1(31), DOI 10.1186/s41077-016-0033-y.

Crisp, V., & Shaw, S. (2012): Applying methods to evaluate construct validity in the context of A level assessment, *Educational Studies*, 38:2, 209-222 <http://dx.doi.org/10.1080/03055698.2011.598670>.

All texts will be provided for participants in advance of the workshop.

### Proposed Schedule

Time	Session	Presenter
09.00	ARRIVAL & WELCOME	Stuart Shaw & Ezekiel Sweiry
9.30 – 10.00	<b>PART 1: INTRODUCTION</b>	
	What is validation?	
	How is it possible to judge whether a qualification has necessary and sufficient validity?	Stuart Shaw
	To what degree is evidence and analysis consistent with the overarching validation (measurement) claim?	
10.00 -10.45	<b>PART 2: SOURCES OF VALIDATION EVIDENCE</b>	
	Overview of sources of evidence and analysis for validation research:	
	<b>Statistical</b>	
	• Types of correlational analysis	
	• Distractor analysis	
	• Differential Item Functioning	
	• Rasch modelling	
	<b>Qualitative</b>	
	• Subject expert analysis and judgement	
	• Analysis of candidate responses	
	• Think aloud protocols	
	<b>Documentary</b>	
	• Literature reviews	Stuart Shaw & Ezekiel Sweiry
	• Syllabus and paper development documentation	
	• Assessment evaluation reports	
	<b>Innovative:</b>	
	• Cognitive laboratory experiments (e.g. cognitive interviewing, think-aloud protocols and vignettes, cognitive demand, behavioural observations, interviews)	
	• Physiological (e.g. eye tracking)	
	• Psycho-sensory	
	• Process data (e.g. computer-generated log files)	
	• Impact (with a focus on Critical Language Assessments)	



# Workshops

## Workshop 1

### 10.45 – 11.00 TEA/COFFEE BREAK

#### 11.00 – 11.45 Activity and reflection (1)

Participants will consider in groups a number of realistic scenarios, with a view to reflecting on the following prompts:

- Which sources of evidence would you use to stake a claim of validity with respect to whether the tasks elicit performances that reflect the intended constructs?
- Why would you choose such these sources of evidence?
- What other sources of evidence might you use?

What constraints might impact on your choice?

Stuart Shaw & Ezekiel Sweiry

### 12.00 – 13.00 LUNCH

#### 13.00 – 14.05 Part 3: ExEMPLIFYING SOURCES OF VALIDATION EVIDENCE: EXAMPLES

Participants, working with authentic data, will be presented with four scenarios based on the following innovative validation approaches:

- Candidate ‘reflection’ and/or ‘collaboration’ scripts
- Impact instrumentation
- Eye tracking
- Process data

#### 14.05 – 14.30 Activity and reflection (2)

- What empirical evidence and logical analyses is required to stake a claim of validity for a qualification?
- What are the relative merits and demerits of the practical examples explored?
- Why and how would you employ these types of evidential sources in your own contexts?

Stuart Shaw & Ezekiel Sweiry

### 14.30 – 14.45 TEA/COFFEE BREAK

#### 14.45 – 16.00 Part 4: Discussion

Participants will be asked to reflect in groups on the workshop and in response to the semi-structured prompts, such as:

- How can validation techniques be combined?
- Looking to the future and AI

#### 16.00 – 16.30 Activity and reflection (3)

Issues for discussion will include:

- What have we learned?
- Who gets to judge sufficient validity, and through what due process?
- How little validation evidence can we rely upon? Should we rely upon? Do we rely upon?

Ezekiel Sweiry

Stuart Shaw

### References

Hubley, A. M., & Zumbo, B. D. (2011). *Validity and the consequences of test interpretation and use*. *Social Indicators Research*, 103(2), 219–230. <https://doi.org/10.1007/s11205-011-9843-4>.

Kane, M.T. (2013). *Validating the interpretations and uses of test scores*. *Journal of Educational Measurement*, 50 (1), 1-73.

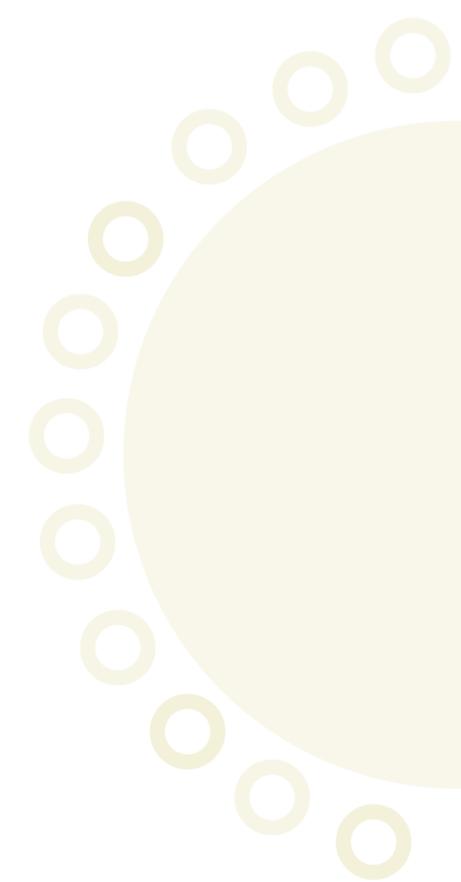
Kane, M.T. (2009) ‘Validating the interpretations and uses of test scores.’ In R.W. Lissitz (ed.), *The Concept of Validity*. Charlotte, NC: Information Age, 39-64.

Kane, M. T. (2006). *Validation*. In R. B. Brennan (Ed.), *Educational Measurement (4th ed., pp. 17-64)*. Westport, CT: American Council on Education/Praeger.

Lissitz, R. W. (ed.). (2009). *The Concept of Validity: Revisions, New Directions, and Applications*. Charlotte, NC: Information Age Publishing.

Messick, S. 1992. *Validity of test interpretation and use*. In *Encyclopedia of Educational Research (6th ed.)*, ed. M.C. Alkin. New York: Macmillan.

Zumbo, B. D., & Chan, E. K. H. (Eds.). (2014). *Validity and Validation in Social, Behavioral, and Health Sciences*. Springer International Publishing/Springer Nature. <https://doi.org/10.1007/978-3-319-07794-9>.





### Workshop 2

**Presenters:** Sam Passeport, Andrew Watts, Nathalie Younès, Constanze Höpfner & Marianne Talbot

#### Exploring feedback dialogues for a transformative feedback culture

This interactive pre-conference workshop invites participants to critically engage with formative assessment or Assessment for Learning (AfL) by focusing on the articulation and shared understanding of feedback. Research highlights feedback as a powerful intervention (Hattie, 2009; Hattie & Timperley, 2007), yet its impact varies. Some approaches view feedback as information transfer, while others see it as a process. Building on the latter, this session examines feedback dialogues through a socio-material lens, exploring how relationships, power dynamics, tools, technologies, and institutional structures shape feedback encounters.

Participants will engage with key research on feedback literacies, gaining insights into how feedback cultures are enacted and experienced by students and assessors. Through discussions and activities, they will reflect on their feedback experiences and explore strategies for fostering meaningful, dialogue-based interactions.

Designed for educational researchers and practitioners—including teachers, lecturers, and curriculum coordinators in secondary and higher education—this session requires no prior expertise, just curiosity and an interest in student agency and dialogic feedback.

By the end, participants will have a deeper understanding of formative feedback as a relational, situated practice and leave with practical strategies to cultivate sustainable, dialogue-rich assessment cultures.

**APPENDIX A:** Template pre-conference workshop

**WORKSHOP TITLE:** Exploring feedback dialogues for a transformative feedback culture

**Presenters:**

Sam Passeport (she/her)  
 Andrew Watts (he/him)  
 Nathalie Younès (she/her)  
 Constanze Höpfner (she/her)  
 Marianne Talbot (she/her)

**Presenters' Bios (500 words max per presenter):**

**Sam Passeport**

Sam works as an Instructional Designer at Tilburg University in the Netherlands. She is also a Professional Doctoral candidate with the University of Dundee in Scotland. She investigates the (missed) feedback encounter between undergraduate students' and their teachers from a critical theory point of view and using interpretative phenomenological analysis. With over 13 years of experience in international education (PK-12 and higher education), Sam promotes a relational and critical approach to coaching and dialogic feedback, creating space for connections and true dialogues between students and teachers. Sam is also the Chair of the Assessment Cultures SIG at AEA-Europe since December 2024.

**Andrew Watts**

Dr. Andrew Watts began his career as a teacher of English in secondary schools in the UK. After eleven years he moved to Singapore where he taught in a Junior College for over four years. He then worked for five years in the Ministry of Education in Singapore, focusing on curriculum development and in-service teacher development. In 1990 he returned to England and worked with Cambridge Assessment from 1992 to 2008. From 2004 he set up and ran the Cambridge Assessment Network, which provides professional development opportunities for assessment professionals internationally. Since leaving Cambridge Assessment he has worked on numerous assessment projects on a freelance basis. In 2019 he completed work for a PhD on the history of examinations in England, based at the University of Cambridge Faculty of Education. He has

made various contributions to AEA-Europe conferences including running pre-conference workshops. He was one of the initiators of the Assessment Cultures Special Interest Group within the Association and continues to serve on the Steering Group for that SIG.

**Nathalie Younès**

Nathalie Younès is a professor of education sciences at the University of Clermont-Auvergne (UCA) in France. Her research focuses on assessment for learning and ecological assessment in higher education. She directs the ACTe laboratory's theme: "Design and Evaluation of Tools and Systems" and a research programme on higher education. She designs and delivers the teacher training programme for teachers and pedagogical advisors at UCA. She has led several collaborative research projects with primary and secondary school teachers and higher education pedagogical advisors to develop assessment for and as learning practices and identify the conditions for their effective development. Between 2016 and 2021, she was president of ADMEE-Europe, a French-speaking network of researchers, evaluators, trainers and teachers who question the methodology of evaluation in education systems.

**Constanze Höpfner (she/her)**

Constanze works as an instructional designer and junior assessment specialist at the Faculty of Humanities and Digital Sciences of Tilburg University. She collaborates with a wide range of stakeholders such as lecturers, education policy staff, and the university's Network for Educational Development and Innovation. As part of a Dutch national initiative ('Smarter Academic Year'), she works towards stimulating assessment practices that reduce pressure on students and lecturers, promote critical thinking, and better integrate assessment and curriculum design.

**Marianne Talbot (she/her)**

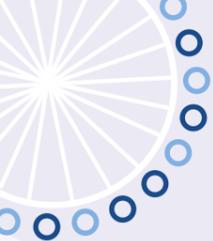
Marianne is a post-transfer PhD researcher at the University of Leeds School of Education. Her area of interest is the impact of professional development in educational assessment on assessors, including qualified teachers, university lecturers, and those assessing in professional workplace environments. Her doctoral research focuses on the personal assessment practice, feedback landscape, and unique assessment ecologies of Chartered Educational Assessors who are also secondary school teachers in England. She is researching their influence as educational assessment experts on the practice of others around them, their individual perceptions of assessment expertise, their confidence to deploy that expertise, and their assessment agency and identity. Her experience is in qualifications, curriculum, and assessment development and evaluation, project management including impact and risk assessments, independent evaluation of outreach and widening participation activities, and course leadership for the Chartered Institute of Educational Assessors, based at the University of Hertfordshire. More information is available on LinkedIn and in her researcher profile at the University of Leeds.

**Why AEA-E members / conference delegates should attend this workshop:**

AEA members/conference delegates should join this workshop to: (1) explore assessment feedback cultures, (2) learn from the latest empirical and theoretical research insights regarding formative assessment, Assessment for Learning, and feedback, and (3) co-create practical strategies for implementation in their own contexts.

Explore Assessment Feedback Cultures: Feedback or Formative Assessment is more than information being transferred from teachers to students. Nowadays, formative feedback cultures aim to center dialogic processes, for instance, through novel assessment models like Programmatic Assessment (van der Vleuten et al., 2012) and using a variety of tools (rubrics, coversheets, AI) and processes (self-assessment, peer learning, coaching).

Research Insights on Feedback: Much research has been carried out in the past few decades, shifting research on feedback from a focus on teachers and their actions, and then students' self-regulation and feedback reciprocity, to finally arrive at the importance of a partnership and shared responsibility between students and teachers through feedback literacy/ies.



# Workshops

## Workshop 2

Co-Create and Take Away Practical Strategies: Participants will contribute insights, practices, and reflections from their own contexts and engage in collaborative exercises and solution design. Through interactive activities and a co-creation session (experimenting with feedback tools, prototyping and sharing ideas for implementation on a digital wall), attendees will leave with tangible strategies and resources to implement in their institutions.

Who this Workshop is for:

This workshop is designed for educational researchers and practitioners, including teachers, lecturers, and programme or curriculum coordinators in secondary schools and higher education (vocational and university settings). It is ideal for those interested in exploring dialogic, relational, and socio-material approaches to feedback to enhance assessment practices in their institutions.

### Overview of workshop (500-600 words):

This PCW will examine the critical role of feedback dialogues in shaping assessment cultures. While feedback is often framed as a simple transmission of information between teachers and students, a learning-centered approach emphasizes its emotional, relational, and situated nature (Winstone & Carless, 2019). Feedback is not merely a cognitive process but first and foremost a socially situated practice influenced by human interactions, institutional structures, and material (Gravett, 2022), which participants will dive into.

### Process

The workshop consists of three sessions following the Design Thinking process, culminating with a co-creation session that will employ bricolage (Kincheloe & Berry, 2004) as a reflective (Schön, 1983) and hands-on inquiry approach to help participants translate insights into actionable strategies. This approach has been recognized as an effective professional learning practice (Campbell, 2018).

### Alignment with the Conference Theme

Formative assessment is regarded as one of the most effective interventions for enhancing student learning (Black & Wiliam, 1998; Hattie, 2009), however, research highlights several barriers in feedback engagement, as reported by Winstone et al. (2017), such as students' lacking the strategic skills for feedback uptake as well as course design problem.

Recent research emphasizes the need to design feedback for uptake (Winstone & Carless, 2019) and to foster a shared responsibility for the feedback process (Nash & Winstone, 2017; Winstone, Pitt & Nash, 2020). This workshop will address these complexities, promoting feedback dialogues that enhance inclusive assessment practices (Tai et al., 2022). In doing so, it touches upon design elements, as well as themes relating to equity and inclusion.

### Workshop Structure

#### Session 1: Feedback Cultures & Dialogues (Empathize & Define)

The workshop begins with an exploration of feedback cultures through listening to an audio material capturing students' voice, enabling participants to empathize with students' experiences. They will examine feedback cultures through different contextual layers (Bronfenbrenner, 1979) and the question of psycho-socio-technical mediations to be considered in a culture of emancipatory evaluation (Younès, 2020, 2025). By reflecting on their own experiences as feedback recipients and providers, participants will explore feedback dialogues, recognizing that student agency and engagement are shaped by both social interactions and material conditions. Using research on feedback literacies as a socio-material practice (Gravett, 2022) and mini-case studies, they will critically analyze missed feedback opportunities and the role of relational pedagogies (Biesta & Tedder, 2007).

#### Session 2: Bricolage Session (1/2) – Ideate

In this co-creative session, participants will reflect on their own contexts, examining how socio-material factors (e.g. relationships, emotions, identity, power, norms, time, space, tools, and structures) shape feedback encounters. Engaging in dialogic feedback processes such as peer coaching, they will collaboratively explore strategies for enhancing feedback interactions and fostering relational trust.

#### Session 3: Bricolage Session (2/2) – Prototype & Test

Participants will engage in a group activity to examine tools and co-create actionable plans for transforming feedback cultures. Through structured reflection and collaborative exercises, they will adapt feedback practices to their institutional contexts. For example, administrators who lead programs will be able to create intentional structures to support teachers and lecturers in more intentional processes for feedback engagement and uptake, and teachers/lecturers will be able to use (modifiable) tools and strategies in the classroom, to promote students' engagement with feedback. The session will conclude with the creation of tangible takeaways, including a collective digital wall (e.g. Padlet), fostering a sense of community and commitment to implementing feedback dialogues.

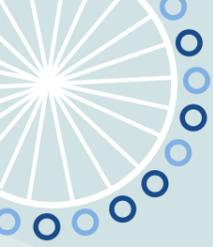
By the end of the workshop, participants will have developed a deeper understanding of feedback as a relational, situated practice. They will leave with practical strategies to cultivate sustainable, dialogue-rich feedback cultures in their assessment contexts.

#### Preparation for the workshop:

Time	Session	Presenter
9.30-12.00 (Inc. a 15 min break 10:30-11:00)	Block I <b>Feedback Cultures &amp; Dialogues [Empathise and Define]</b>	Sam Passeport, Marianne Talbot, Nathalie Younès
12.00 - 13.00	L U N C H	
13.00-14.30	Block II <b>'Bricolage' Session (1/2) [Ideate &amp; Prototype]</b>	Sam Passeport, Andrew Watts
14.30 - 14.45	TEA / COFFEE BREAK	
14.45-16.30	Block III <b>'Bricolage' Session (2/2) [Prototype and Test]</b>	Constanze Höpfner, Sam Passeport

#### References

- Black, P. and Wiliam, D. (1998) 'Inside the black box: Raising standards through classroom assessment,' *Phi Deltan Kappan*, 80(2), pp. 139-148.
- Bronfenbrenner, U. (1979) *The ecology of human development*. Harvard: Harvard University Press.
- Campbell, L. (2018) 'Pedagogical bricolage and teacher agency: Towards a culture of creative professionalism,' *Educational Philosophy and Theory*, 51(1), pp. 31-40.
- Gravett, K. (2022) 'Feedback literacies as sociomaterial practice,' *Critical Studies in Education*, 63(2), pp. 261-274.
- Hattie, J. (2009) *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*.: Routledge.
- Hattie, J. and Timperley, H. (2007) 'The Power of Feedback,' *Review of Educational Research*, 77(1), pp. 81-112.
- Kincheloe, J. L. and Berry, K. S. (2004) *Rigour and complexity in educational research: conceptualising the bricolage*. Maidenhead, Berkshire: Open University Press.
- Schön, D. (1983) *The Reflective Practitioner: How Professionals Think in Action*. London: Temple Smith.
- van der Vleuten, C. P., Schuwirth, L. W., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K. and van Tartwijk, J. (2012) 'A model for programmatic assessment fit for purpose,' *Med Teach*, 34(3), pp. 205-14.
- Winstone, N. and Carless, D. (2019) *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. Routledge.
- Younès, N. (2025). *Vers une évaluation émancipatrice dans un écosystème formatif*. Dans J.F. Marcel et Gremion, C. (Dir.), *Évancipation dans l'institution. Oser le rapprochement entre émancipation et évaluation* (p. 169-183). Editions Cepaduès.
- Younès, N. (2020). *Vers une évaluation écologique dans l'enseignement supérieur : dispositifs et transformations en jeu*. [Note de synthèse pour l'habilitation à diriger des recherches]. Nancy, Université de Lorraine. France. tel-03032417v1.



# Workshops

## Workshop 3

### Workshop 3

**Presenters:** Marieke van Onna, Bas Hemker, Cor Sluijter

#### Your best friend the psychometrician: The preventive role of psychometrics in test development

This workshop will help you to get a further insight in the advantages of timely involvement of a psychometrician when setting up a new testing program. It is useful for non-psychometricians to find out on how many more issues they can call on their friendly neighbourhood psychometrician. For psychometricians, the workshop may help to increase their added value.

During the workshop, we will use a scheme of all activities involved in test development. We'll discuss several general psychometric topics, and relate these to the decisions you will have to make for these activities. We'll show in which way a psychometrician might contribute to each activity. In each block, we'll give guidelines and illustrate best practices. We'll invite you to share your experiences with the topics and ask us for advice.

No R, no formulas, still all psychometrics.

**APPENDIX A:** Template pre-conference workshop

**WORKSHOP TITLE:** Your best friend the psychometrician: The preventive role of psychometrics in test development

**Presenters:**

Marieke van Onna, Bas Hemker, Cor Sluijter

Presenters' Bios (500 words max per presenter):



**Marieke, Bas and Cor**

- Studied psychology and wrote a Ph.D. thesis on a psychometric topic
- Have at least 20 years of experience at Cito, where they work or worked at the psychometric department CitoLab
- Are fellows at AEA-Europe
- Have given numerous lectures, classes, and workshops in the Netherlands, as well in many other countries on five continents.
- Are involved in wide variety of educational measurement projects, that require a combination of statistics, psychometrical skills and knowledge of educational practice.
- Love working with motivated colleagues on improving educational assessments.

**Marieke**

Wrote her thesis on Bayesian estimation of latent class models and nonparametric IRT

- Worked a few years as an assistant professor for statistics and methodology in psychology before joining Cito
- Coordinates all psychometric analyses that are necessary for the national exams in Dutch secondary education
- Apart from that, works on some smaller testing programs

**Bas**

- Has been the head of the psychometric department of Cito
- Is involved in CitoLab International, Cito's international department
- Works at Cito, specialized as an educational measurement researcher, with quality of school exams as one of his projects
- Is a member of the Dutch Committee on Test Matters (COTAN)
- Was a member for many years of the Professional Development Committee of AEA Europe

**Cor**

- Also a former head of psychometric department of Cito
- Works as an independent consultant on test quality and test use
- Is a lecturer in educational measurement at Fontys University of Applied Sciences
- Is external member of several Exam Boards of various universities in the Netherlands
- Is the Vice-President of AEA-Europe since November 2024

**Why AEA-E members / conference delegates should attend this workshop:**

As Ronald Fisher said: To consult the statistician after an experiment is finished is often merely to ask him to conduct a post mortem examination. He can perhaps say what the experiment died of.

Read 'psychometrician' for 'statistician', and 'educational test' for 'experiment'.  
Psychometricians like to work in prevention!

This workshop will help you to get a further insight in the advantages of timely involvement of a psychometrician when setting up a new testing program. It is useful for non-psychometricians to find out on how many more issues they can call on their friendly neighbourhood psychometrician. For psychometricians, the workshop may help to increase their added value.

No R, no formulas, still all psychometrics.

**Who this Workshop is for:**

Anybody involved in setting up (new) exams, tests and testing programs: managers, project leaders, psychometric leads, test developers, practitioners in the field of educational testing. No mathematical pre-knowledge necessary. (Max. 25 participants.)

**Overview of workshop (500-600 words):**

Psychometricians are often asked to 'run the analyses'. However, they can only present you sensible tables as output, if they know which meaning will be attached to the results. Who are the stakeholders, and what is the type of message, you want to convey to each of these stakeholders? Therefore, psychometricians will ask you 'what is the purpose of your test?', shortly followed by 'what is the format of your test report, for each stakeholder?', as the full meaning of the test will only be in the use of the reports by the stakeholders.

In order to run their analyses, psychometricians need data, attributes, classifications, higher-level information and, most importantly, information on decisions. All decisions you make during the development of a testing program, have an influence on the quality of the data and the options for reporting. E.g. you can't report on subdomains if your test matrix does not include a balanced representation of them: your psychometrician will start complaining about too little score points and large measurement errors.

# Workshops

## Workshop 3

If you have a preventative discussion with a psychometrician when setting up a testing program, they'll point out the implications of decisions, and allow you to make an informed decision. They can also help with improving the quality of the data, because they know: "Garbage in, garbage out".



During the workshop, we will use the figure above as a scheme of all activities involved in test development. We'll discuss several general psychometric topics, and relate these to the decisions you will have to make for these activities. We'll show in which way a psychometrician might contribute to each activity.

After an introduction to the test development scheme, we'll discuss the topics below. In each block, we'll give guidelines and illustrate best practices. We'll invite you to share your experiences with the topics and ask us for advice.

- Validity and the desired coherence between report purpose, test matrix and tasks, by means of the evidence-centered design model and constructive alignment.
- Different kinds of standard setting and the implied reporting options. Standards may be relative or absolute, and standard setting procedures may be test-, item-based, population- or person-based.
- Threats to validity in the case of missing values or differential item functioning. Several reasons why candidates do not answer questions are discussed. Subgroups of candidates may differ in general ability level. Differential item functioning, however, refers to deviating scores on a single question.
- Reliability, and the effect of rater agreement. We'll distinguish between global and local reliability, and the relation with the purpose of the test. In the case of questions that cannot be marked automatically, rater agreement might be an issue. A psychometrician can advise on the amount of raters that are needed, or detect items that need better marking schemes.
- When (not) to use adaptive testing. We'll discuss the pros and cons of adaptive testing. We'll include the difference between full adaptive testing and multi-stage testing. In addition we'll discuss automated test assembly (ATA), which may also be useful in non-adaptive settings.

We'll finish with a list of what annoys psychometricians most (e.g. lacking item version administration).

Preparation for the workshop:

### Please think about the following questions beforehand:

- What assessment program are you working on, or planning to work on?
- Which problems do you experience in your program?
- Which part of the test cycle is malfunctioning, or missing, in your program?
- What do you think is/might be the contribution of a psychometrician for your program?

### Tentative Schedule

Time	Session	Presenter
9.30-12.00 (Inc. a 15 min break 10:30-11:00)	Block I: Introduction to test cycle Validity (ECD, Constructive alignment) Reporting (standard setting)	Marieke, Bas, Cor
12.00 - 13.00	Lunch	
13.00 - 14.30	Block II: Threats to validity (DIF, missing data) Reliability and rater agreement	Marieke, Bas, Cor
14.30-14.45	Tea/coffee break	
14.45-16.30	Block III: When (not) to use adaptive testing Automated test assembly	Marieke, Bas, Cor



### Workshop 4

**Presenters:** Celine van der Lienden & Laurinde Koster

#### From awareness to action: embedding inclusive assessment in teacher development programs in higher education

In today's diverse learning environments, inclusive assessment is essential to ensure fair and valid learning outcomes. Moreover, assessment should support higher education students in their learning process, enabling them to demonstrate their knowledge and skills. Inclusive assessment practices align with broader educational values such as contributing to a more accessible and equitable society.

To achieve such inclusive assessment practices, it is crucial that educators are supported in developing the necessary awareness, knowledge, and skills. Teacher development programs play a key role in this process. In the Netherlands, structured development trajectories are used to build assessment literacy in among university teaching staff, with inclusive assessment increasingly integrated as a central theme.

This workshop focuses on how inclusive assessment can be meaningfully integrated into teacher development programs in higher education. Led by an educational measurement expert and an inclusive education specialist from Risbo (Erasmus University Rotterdam), we take an evidence-based approach by translating research into practical strategies for program and course design, assessment construction, and grading. Through interactive activities and peer dialogue, participants will explore how inclusive assessment can be implemented in their own context, with concrete tools that contribute to fairer, more valid assessment practices and improved student learning outcomes.

#### APPENDIX A: Template pre-conference workshop

**Workshop Title:** From awareness to action: embedding inclusive assessment in teacher development programs in higher education Presenters: Celine van der Lienden, Msc. (Risbo, Erasmus University Rotterdam) Laurinde Koster, MA. (Risbo, Erasmus University Rotterdam)

#### Presenters' Bios (500 words max per presenter):

**Celine van der Lienden:** Celine van der Lienden is an education advisor at Risbo, Erasmus University Rotterdam, where she specialises in assessment innovation, curriculum development and teacher training in higher and secondary education. Her focus is on feedback design, programmatic assessment, and the integration of emerging technologies such as AI in assessment practices. A key part of her work involves developing and facilitating professional development programs for academic teachers, with a particular focus on strengthening assessment literacy. With a strong focus on the learning function of assessment, she helps institutions rethink traditional approaches and adopt more holistic and student-centred strategies. Her role often includes guiding lecturers through complex design questions, translating educational theory into practical, context-sensitive solutions. Celine holds a Master's in educational science from Radboud University and Educational Measurement from Fontys University of Applied Sciences. She holds a bachelor's degree in Pedagogical Sciences from Erasmus University Rotterdam, which forms the foundation of her educational thinking. Combining theoretical expertise with practical experience, Celine is committed to strengthening assessment as a tool for student learning. She brings a thoughtful, analytical perspective to complex educational challenges and works closely with educators to help shape assessment practices that are meaningful, future-oriented, and pedagogically sound.

**Laurinde Koster:** Laurinde Koster (MA) is an educational consultant and trainer with experience in higher education, committed to fostering inclusive and innovative teaching practices. Since 2017, she has worked at Risbo, Erasmus University Rotterdam, where she advises faculties, lecturers, and external partners on course design, blended and hybrid learning, and educational innovation. As a trainer, she develops and delivers professional development programs, including University Teaching Qualification (UTQ) and Basic Examiner Qualification (BEQ) trajectories, PhD training programs, and teacher workshops. Her work is rooted in a strong belief in inclusive education, with a focus on supporting educators in creating learning environments that embrace diversity, accessibility and inclusion. She has guided teaching teams in developing inclusive curricula and assessment practices that recognize and value diversity as assets in the classroom. With an international academic background—including a Master's in Multilingualism (University of Groningen), a Master's in English and Education (University of Copenhagen), and a bachelor's in education with TESOL—Laurinde combines theoretical insight with hands-on experience. She brings a practical and empathetic approach to her work, characterized by creativity and a deep commitment to educational equity and sense of belonging.

#### Why AEA-E members / conference delegates should attend this workshop:

By connecting theory to practice, this session provides concrete tools and frameworks that participants can apply in their own institutions. Participants will engage with peers across contexts, share challenges and solutions, and leave with actionable insights and strategies to enhance fairness and quality in assessment design and practice. The workshop will have a focus on professionalising educators through structured development trajectories.

#### Who this Workshop is for:

All interested parties: researchers, practitioners, higher education teachers, policy makers, educational advisors.

#### Overview of workshop (500-600 words):

In an increasingly diverse higher education landscape, inclusive assessment has become a key element for promoting fairness, improving validity, and supporting all students in demonstrating their knowledge and skills. Inclusive assessment aims to ensure that students, regardless of background or ability, can demonstrate their learning fairly. However, "historically, assessment has struggled to meet the needs of student diversity in higher education" (McArthur, 2016; Nieminen, 2022). Moving beyond accommodations, inclusive assessment advocates for proactive design. As emphasized, "assessment should be designed to be inclusive for all students in the first place – not retrospectively" (Hanafin et al., 2007; Nieminen, 2022). This approach aligns with the principles discussed in Assessment for Inclusion in Higher Education (Ajjawi et al., 2022), which highlight the need for equity-driven, responsive practices in assessment design. Therefore, achieving inclusive assessment practices requires more than policy—it requires an investment in teacher professionalization to enable them to design inclusive assessment practices. Educators need structured opportunities to build their assessment literacy, develop inclusive mindsets, and acquire practical strategies to apply in their daily practice. This workshop explores how inclusive assessment can be integrated as a central theme in teacher development programs in higher education. Drawing on experience with professionalization trajectories in the Netherlands, we discuss how such programs can raise awareness and create meaningful impact on learning and assessment quality. The workshop is led by an educational measurement expert and an inclusive education specialist from Risbo (Erasmus University Rotterdam).

#### Structure of the workshop

The workshop starts with a short introduction to the key principles of inclusive assessment and its importance in higher education. Drawing on current research and international frameworks such as Universal Design for Learning (UDL), we reflect on how inclusive assessment contributes not only to fairness and validity, but also to learning outcomes. Participants will be invited to reflect on their own context and share current practices and challenges. In the second part, the focus is on the design of teacher development programs at Risbo. Using examples from our work in the Dutch higher education context.



Participants will analyze examples of professionalization interventions and frameworks, and discuss what elements could be adapted or applied in their own institutional or national context. The final part of the workshop is hands-on. Participants will engage with practical tools and resources designed to support inclusive assessment in course design, exam construction, and grading. Through peer dialogue and group discussion, they will reflect on how these tools can be integrated into professional development programs and how to create impact within their own institutions.

### Outcomes and Takeaways

By the end of the workshop, participants will:

- Understand the value of inclusive assessment for fairness, validity, and student learning
- Be familiar with strategies for embedding inclusive assessment into teacher development programs
- Gain practical tools to support educators in developing inclusive assessment practices
- Exchange ideas and approaches with peers working in various international contexts

This workshop is relevant for researchers, practitioners, higher education teachers, assessment specialists, policy makers, and education advisors involved in the professionalization of educators and the improvement of assessment in higher education.

### References

Ajjawi, R., Tai, J., Boud, D., & Jorre de St Jorre, T. (Eds.). (2022). *Assessment for inclusion in higher education: Promoting equity and social justice in assessment*. Routledge. <https://doi.org/10.4324/9781003293101>

Hanafin, J., Shevlin, M., Kenny, M., & McNeela, E. (2007). *Including young people with disabilities: Assessment challenges in higher education*. *Higher Education*, 54(4), 435–448. <https://doi.org/10.1007/s10734-006-9005-9>

SpringerLink+1ResearchGate+1 McArthur, J. (2016). *Assessment for social justice: The role of assessment in achieving social justice*. *Assessment & Evaluation in Higher Education*, 41(7), 967–981. <https://doi.org/10.1080/02602938.2015.1053429>

journal.alдинhe.ac.uk+2East Texas A&M University, ETAMU+2ResearchGate+2 Nieminen, J. H. (2022). *Assessment for inclusion: Rethinking inclusive assessment in higher education*. *Teaching in Higher Education*, 27(8), 1038–1052. <https://doi.org/10.1080/13562517.2021.2021395>

Taylor & Francis Online+4East Texas A&M University, ETAMU+4journal.alдинhe.ac.uk+4 Preparation for the workshop: No preparation required. Tentative Schedule Time Session Block I Presenter 9.30- 12.00 (Inc. a 15 min break 10:30-11:00) Celine van der Lienden and Laurinde Koster 12.00- 13.00 Lunch 13.00- 14.30 Block II Celine van der Lienden and Laurinde Koster 14.30- 14.45 Tea/coffee break 14.45- 16.30 Block III Celine van der Lienden and Laurinde Koster.

## Workshop 5

**Presenters:** Stuart Gallagher and Georgie Billings

### Establishing valid qualification equivalency with qualitative judgement

Stuart Gallagher and Georgie Billings Where statistical equating methods are not available, through lack of common items or common candidates, but equivalency between two qualifications is required, it can be difficult to provide robust evidence.

**Session 1** of the workshop focuses on a novel standard setting methodology that allowed for IGCSE scores to be translated into Mississippi end-of-course performance levels and integrated into the state accountability system. The method draws on aspects of both Body of Work (BoW) and Bookmarking methods to create an operationally feasible process.

**Section 2** explores why there is a need for qualification equivalency, how a qualification can be broken down into content, demand and awarding standards, and some possible methodologies for establishing standards equivalency when the data for psychometric equating is not available.

The practical activity will involve delegates becoming comfortable with using the CRAS framework to evaluate the demand of questions, and how to set up, run and evaluate a comparative study using the No More Marking platform. We will discuss the benefits and limitations of this methodology for establishing demand equivalency, and what follow up work can usefully be done with the results.

APPENDIX A: Template pre-conference workshop

**Workshop Title:** Establishing valid qualification equivalency with qualitative judgement

**Presenters:** Stuart Gallagher and Georgie Billings

Presenters' Bios (500 words max per presenter):

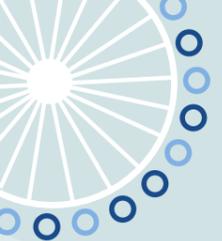
#### Stuart Gallagher

Stuart Gallagher is a Principal Assessment Advisor in the International Education group at Cambridge University Press & Assessment. His responsibilities include advising on assessment issues for the US market, evidencing and developing the technical quality of large-scale, high-stakes international assessments, and innovating practice around reporting, analysing and articulating candidate performance and awarding standards.

He has wide experience of UK assessment, having worked in Research and Technical Standards and as a Chair of Examiners for OCR. This includes assessment design and development, general qualification reform, awarding standards in vocational and technical qualifications, and major process change for key operational activities.

Before moving into educational assessment, Stuart taught French and Spanish, having gained a Postgraduate Certificate in Education and Qualified Teacher Status at the Faculty of Education in Cambridge. As a secondary school MFL practitioner and head of department, his key interests were innovating and improving curriculum design and delivery, developing use of data for reporting and analysis, and introducing immersion-style teaching of modern languages.

Stuart holds a BA in French and Linguistics from the University of York and an MEd in Researching Practice from the University of Cambridge.



# Workshops

## Workshop 5

### Georgie Billings

Georgie Billings is Head of Assessment Quality within the Assessment Reform Group at Cambridge University Press & Assessment. The team enacts assessment reform projects for ministries, school groups and NGOs globally. Her responsibilities include heading up the quality assurance and research programme for accredited qualifications, advising on standard setting across a vast range of tests, including Africa, Asia and the Middle East, and heading up reform projects in contexts such as Cox’s Bazar refugee camps.

Georgie has a wealth of prior experience, including teaching, working for the educational UK charity The Prince’s Trust and working for the Department for Education.

Georgie holds a BAHons in English, PGCSE (Cantab) in Secondary English, MEd(Cantab) in Researching Practice, and holds RITTech accreditation as a data analyst. She has completed post-graduate studies in Creative Writing and Forensic Chemistry, and is currently finishing her MSc in International Development at University of Edinburgh.

### Why AEA-E members / conference delegates should attend this workshop:

Where statistical equating methods are not available, through lack of common items or common candidates, but equivalency between two qualifications is required, it can be difficult to provide robust evidence. This workshop explores some methodologies for doing this, based on real experiences in the field, and will provide practical tools that can be applied to other qualifications.

### Who this Workshop is for:

Anyone interested in establishing comparability or equivalency in the demand or awarding standard of qualifications through non-psychometric methodologies, particularly across different international contexts.

### Overview of workshop (500-600 words):

Part 1: In 2022, the US Department of Education approved Cambridge IGCSE assessments in English Language, Mathematics and Biology to be administered in place of the corresponding high school end-of-course tests within the Mississippi state testing program, known as the MAAP. Cambridge IGCSE and the MAAP represent the different educational testing traditions in the UK and the US: Cambridge IGCSE assessments include a significant number of constructed response and multi-part items and have a strong focus on evaluating candidate-generated evidence, while MAAP end-of-course tests are grounded in psychometric tradition and make considerable use of field-tested banks of selected-response items.

This section of the workshop focuses on a novel standard setting methodology that allows for IGCSE scores in these subjects to be translated into Mississippi end-of-course performance levels and integrated into the state accountability system. The method draws on aspects of both Body of Work (BoW) – a holistic standard-setting method often used when an assessment contains many open-ended or constructed response items[1] – and Bookmark methods to create an operationally feasible process.

Delegates will learn about how this blended approach was designed and operationalised, including the statistical methodology used to develop ordered item booklets for constructed response items, and consider the challenges at each stage. They will have the opportunity to compare and contrast the BoW and Bookmark aspects of the process side-by-side, and explore the contributions made by these activities to the overall validity of the outcome.

Part 2: Cambridge International accredits a variety of qualifications globally which are owned and administered by other awarding organisations, resulting in co-certification and a statement of equivalency with a Cambridge qualification.

The session will explore why there is a need for qualification equivalency, how a qualification can be broken down into content, demand and awarding standards, and some possible methodologies for establishing standards equivalency when the data for psychometric equating is not available.

The practical activity will involve delegates becoming comfortable with using the CRAS framework[2] to evaluate the demand of questions, and how to set up, run and evaluate a comparative study using the No More Marking platform. We will discuss the benefits and limitations of this methodology for establishing demand equivalency, and what follow up work can usefully be done with the results.

### Preparation for the workshop:

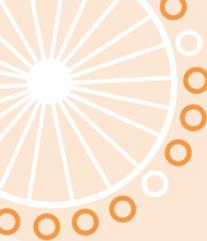
While no explicit knowledge or prior reading is required, delegates will need to have a laptop with the ability to connect to WiFi at the venue. They may also wish to bring two question papers which they are interested in comparing the demand of.

### Tentative Schedule

Time	Session	Presenter
9.30-12.00 (Inc. a 15 min break 10:30-11:00)	Block I  Blending Bookmark and Body of Work standard setting methods: a practical solution to linking grading scales across testing traditions. This will include discussion of the problem statement and the development of the methodology.  Delegates are then invited to undertake a critique of the method, exploring the impact with real data.	Stuart Gallagher
12.00-13.00 L u n c h		
13.00-14.30	Block II  Discussion of the need for comparability Exploring the concepts of content, demand and awarding standards Development of qualitative methods to establish equivalency	Georgie Billings
14.30-14.45 T e a / c o f f e e b r e a k		
14.45-16.30	Block III  Introducing the CRAS framework Practical task setting up a demand standard comparative task in No More Marking and analysing the results.	Georgie Billings

[1] Wyse, A. E., Bunch, M.B., Deville, C., and Viger, S. G. (2014). A Body of Work Standard-Setting Method With Construct Maps. Educational and Psychological Measurement, 74(2) pp.236–262

[2] Johnson, M. and Mehta, S. (2011). Evaluating the CRAS Framework: Development and recommendations. Research Matters(12) pp. 27-32



### Workshop 6

**Presenter:** Iasonas Lamprianou

#### Network Analysis for the investigation of Rater Effects (using R)

This workshop introduces the application of Network Analysis (NA) to rater-mediated assessments. NA analyzes rating datasets by considering pairwise comparisons between raters.

Participants will learn how to detect and interpret key rater behaviors, including severity/leniency, inconsistency (misfit), halo effects, bias, drift (changes over time), and the formation of rater sub-communities. A key feature of the workshop is the comparison of NA results with those from traditional approaches such as the Rasch model.

NA is a flexible method that can handle nominal, dichotomous, ordinal, and numeric data. Unlike traditional models that rely on strong assumptions (e.g., local independence or unidimensionality), NA operates with minimal requirements, making it especially suitable for complex or non-standard rating contexts. Visualizations further enhance interpretability.

The workshop emphasizes hands-on experience using open-source R code and real datasets from published studies. A brief theoretical overview will also be provided. Participants are encouraged to bring their laptops and follow along.

#### Reading list:

- Lamprianou (2018) in *Educational and Psychological Measurement*, 78(3), 430–459
- Lamprianou (2023) in *Sociological Methods and Research*, 55(1), 525–553
- Lamprianou (2025, in press) in *Research Methods in Applied Linguistics*
- Lamprianou et al. (2023) in *Assessing Writing*, 56, 100713

#### APPENDIX A: Template pre-conference workshop

**Workshop Title:** Network Analysis for the investigation of rater effects

**Presenters:** Iasonas Lamprianou, University of Cyprus

#### Presenters' Bios (500 words max per presenter):

Iasonas Lamprianou is an Associate Professor of Quantitative Methods at the Department of Social and Political Sciences, University of Cyprus. His methodological interests include rater and coding effects, Rasch models, and appropriateness measurement (person-fit).

His recent books include 'A Step-by-Step Guide to Applying the Rasch Model Using R: A Manual for the Social Sciences' (2024, 2nd Edition) and 'Network Analysis for Rating Datasets in R: A Multi-Disciplinary Perspective' (2025, in press), both by Routledge.

Iasonas has published in journals such as *Educational and Psychological Measurement*, *Sociological Methods and Research*, *International Journal of Social Research Methodology*, *Journal of Educational Measurement*, *Research Methods in Applied Linguistics*, *Language Testing*, *Assessing Writing*, *Assessment in Education*, *International Journal of Testing*, and others.

#### Why AEA-E members / conference delegates should attend this workshop:

Participants will develop data analysis skills using open-source, cutting-edge tools that can be used either as stand-alone approaches or in complement to existing methods (e.g., Rasch models).

#### Who this Workshop is for:

Researchers, academics, students, and practitioners who want to enhance their empirical analysis toolbox, particularly in rater designs and rater effects within the broader field of social sciences.

#### Overview of workshop (500-600 words):

This workshop introduces the application of Network Analysis (NA) to rater-mediated assessments. NA analyzes rating datasets by considering pairwise comparisons between raters.

Participants will learn how to detect and interpret key rater behaviors, including severity/leniency, inconsistency (misfit), halo effects, bias, drift (changes over time), and the formation of rater sub-communities. A key feature of the workshop is the comparison of NA results with those from traditional approaches such as Rasch measurement.

NA is a flexible method that can handle nominal, dichotomous, ordinal, and numeric data. Unlike traditional models that rely on strong assumptions (e.g., local independence or unidimensionality), NA operates with minimal requirements, making it especially suitable for complex or non-standard rating contexts. Visualizations further enhance interpretability.

The workshop emphasizes hands-on experience using open-source R code and real datasets from published studies. A brief theoretical overview will also be provided. Participants are encouraged to bring their laptops and follow along.

#### Reading List:

- Lamprianou, I. (2018). *Investigation of rater effects using Social Network Analysis and Exponential Random Graph Models*. *Educational and Psychological Measurement*, 78(3), 430-459.
- Lamprianou, I. (2023). *Measuring and visualizing coders' reliability: New approaches and guidelines from experimental data*. *Sociological Methods and Research*, 55 (1), 525-553.
- Lamprianou (2025, in press). *Network Analysis for the investigation of rater effects: a comparison of ChatGPT vs human raters*. *Research Methods in Applied Linguistics*.
- Lamprianou, I., Tsagari, D., & Kyriakou, N. (2023). *Experienced but detached from reality: Theorizing and operationalizing the relationship between experience and rater effects*. *Assessing Writing*, 56, 100713.

*Preparation for the workshop:*

*Laptop with R, RStudio. More information will be provided.*

# Workshops

## Tentative Schedule

Time	Session	Presenter
9.30 - 12.00 (Inc. a 15 min break 10:30-11:00)	Block I <ul style="list-style-type: none"><li>Theory of Rater Effects</li><li>Introduction to Network Analysis</li><li>Severity/Leniency</li></ul>	Iasonas Lamprianou
12.00 - 13.00 Lunch		
13.00 - 14.30	Block II <ul style="list-style-type: none"><li>Consistency / Inconsistency</li><li>Halo Effects</li></ul>	Iasonas Lamprianou
14.30 - 14.45 Tea/coffee break		
14.45 - 16.30	Block III <ul style="list-style-type: none"><li>Differential Functioning (drift)</li><li>Detection of sub-communities</li></ul>	Iasonas Lamprianou

## Special Interest Groups (SIGs)

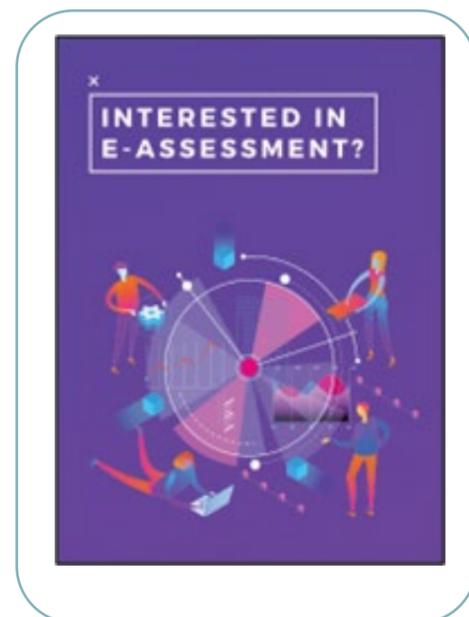
# Special Interest Groups (SIGs)

## e-Assessment SIG

Since its introduction at the Prague conference in 2017, the eAssessment SIG has offered a range of activities for sharing specialist knowledge and professional skills relevant to a wide range of e-assessment topics. However, the worldwide shift of learning and assessment into the digital realm has – to a certain extent – made e-assessment specialists of us all. E-assessment is no longer a specialism of the future; in its broader interpretation, it has become more a ‘common’ than a ‘special’ Interest, and the e-Assessment SIG recognises and accommodates that important development. Please join us at our events to contribute to discussions about how the world of e-Assessment is changing and how e-Assessment is changing the world.

During the conference, please come and find out more at the ‘Meet the SIGs’ event. We would be very happy to welcome you to our SIG. Joining is free for AEA-E members, with a registration link on our webpage at

<https://aea-europe.net/eassessment-sig/>



### eAssessment SIG Steering Committee members:

- **Graeme Clark**, (Chair), Scottish Qualifications Authority
- Dee Arp, NEBOSH
- Rebecca Conway, NCFE
- Irene Custodio, Pearson
- Alison Matthews, University of Oxford
- Stuart Shaw, University College London
- Ben Stafford, AQA

## Assessment Cultures SIG

The Assessment Cultures SIG offers an opportunity for members to examine their own cultures and values, and to learn about and from the assessment cultures of other countries. The SIG offers a collaborative ‘community of practice’ in which ‘to share what they know, to learn from one another regarding some aspects of their work and to provide a social context for that work’ (Lave and Wenger, 1998).

The SIG’s identity is created by the contributions of its members, whose individual perspectives – for example, historical, educational, psychological, sociological or methodological – are variously brought to bear on the work of assessment. Members come together to study the assessment cultures within which they work, through discussion, debate, and the exchange of information and ideas.

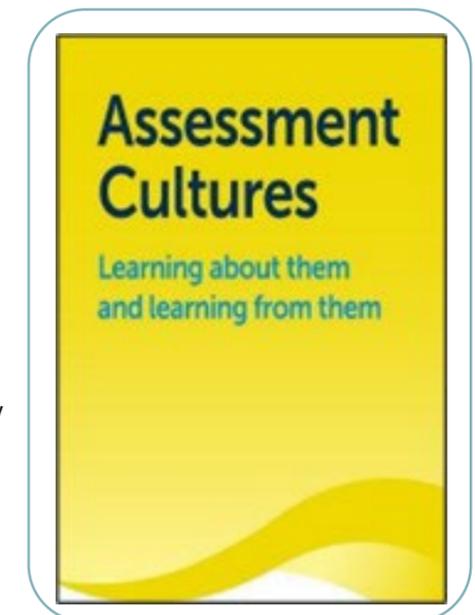
Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815355>

During the conference, please come and find out more at the ‘Meet the SIGs’ event. We would be very happy to welcome you to our SIG. Joining is free for AEA-E members, with a registration link on our webpage at

<https://aea-europe.net/assessment-cultures-sig/>

### Assessment Cultures SIG Steering Committee members:

- Isabel Nisbet, Independent, London, UK
- Raphaël Pasquini, University of Teacher Education Vaud, Switzerland
- **Sam Passeport**, (Chair), Vrije Universiteit Amsterdam, Netherlands
- Lise Vikan Sandvik, Norwegian University of Science and Technology, Norway
- Christoph Schneider, Trier University, Germany
- Stuart Shaw, University College London, UK
- Marte Sørve Syverud, Norwegian Defence University College, Norway
- Andrew Watts, Wolfson College, Cambridge, UK
- Nathalie Younès, Université Clermont-Auvergne, France



# Special Interest Groups (SIGs)

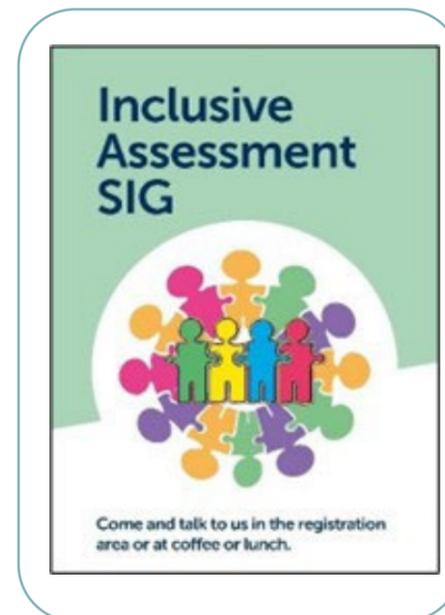
## Inclusive Assessment SIG

The internationalisation of Education worldwide has broadened student diversity in educational institutions in recent years. Inclusive assessment or Assessment for Social Justice promotes developmental, timely feedback which supports students' understanding of the assessment requirements and allows them to reflect on how they can improve their academic performance. It also ensures fairness and inclusivity for all students. Assessment is a powerful driver of learning which may engage our diverse student cohorts in processes that support their development, academic success, and employability, while connecting with their own linguistic identities, learning experiences, and cultural capital. Our SIG offers a platform to present and discuss research, debate current events, and discuss issues concerning fairness, inclusion, and advocacy, so that educators' work includes the voices of all students irrespective of background.

During the conference, please come and find out more at the 'Meet the SIGs' event. We would be very happy to welcome you to our SIG. Joining is free for AEA-E members, with a registration link on our webpage at <https://aea-europe.net/inclusive-assessment-sig/>

### Inclusive Assessment SIG Steering Committee members:

- **Katy Finch**, (Chair), AQA, UK
- **Irini Papakammenou**, University of Staffordshire, UK
- **Elodie Persem**, Ministry of National Education and Youth, France
- **Thomais Rousoulioti**, Aristotle University, Greece



## Postgraduate and Early Career Researcher Network SIG

Being a postgraduate student or early career researcher can be a lonely and challenging journey. Our SIG was established to bring together AEA-E members facing such challenges, providing them with opportunities to meet, interact with other students and researchers, be part of a supportive community, and build a professional network. In our SIG, we can meet, connect, discuss, and share our work and research within a supportive community. As a collaborative research community, the Postgraduate Student and Early Researcher Network provides members with opportunities to come together to get to know each other, to inspire and to share research, experiences, and practices within the field of assessment in education. We welcome all postgraduate students and early career researchers to join our SIG and be a part of our network.

Whether you are already a member or you are just curious about our SIG, we want to get to know you and your research. During the conference, please come and find out more at the 'Meet the SIGs' event. We would be very happy to welcome you to our SIG. Joining is free for AEA-E members, with a registration link on our webpage at <https://aea-europe.net/postgraduate-early-career-network/>

### Postgraduate Student and Early Career Researcher Network SIG Steering Committee members:

- **Estelia Borquez Sanchez**, (Chair), University of Glasgow
- **Gabriel Cipriano**, CIES-ISCTE, Instituto Universitário de Lisboa
- **Julie Leonardsen**, Department of Teacher Education in Vocational Subjects, Norwegian University of Science and Technology
- **Shakeh Manassian**, University of Glasgow/PSI Services
- **Pernille Moe**, Department of Pedagogy, Western Norwegian University of Applied Sciences
- **Dan-Anders Normann**, (Proxy Chair for 2025 Conference), Department of Teacher Education in Vocational Subjects, Norwegian University of Science and Technology.



# Special Interest Groups (SIGs)

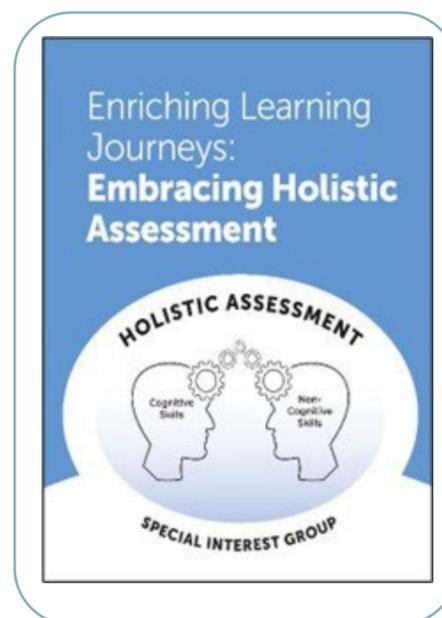
## Holistic Assessment SIG

To succeed in the modern society, including in the workplace, learners need to complete their education with more than the academic knowledge, skills and understanding assessed in traditional school subjects. They need a broad range of 21st-century competencies coupled with other personal attributes such as solid ethical values, positive learning dispositions, and high levels of personal wellbeing. This is the concept of holistic education, which has a deep-rooted history in several European pedagogical philosophies. Both cognitive skills (such as problem-solving, critical thinking, and creativity) and so-called 'non-cognitive' skills (such as communication, collaboration, and perseverance) are important. Assessing personal competencies and attributes accurately, and understanding how they interact to affect educational outcomes and life chances, are significant challenges in the assessment field. Despite its importance, however, holistic assessment is still an under-researched area. This SIG aims to bring members together who are interested in theories and practices for assessing cognitive and non-cognitive skills and attributes, and provide recommendations for effective assessment tools.

During the conference, please come and find out more at the 'Meet the SIGs' event. We would be very happy to welcome you to our SIG. Joining is free for AEA-E members, with a registration link on our webpage at <https://aea-europe.net/holistic-assessment-sig/>

### Holistic Assessment SIG Steering Committee members:

- **Gulbakhyt Sultanova** (Chair), Center for Pedagogical Measurements, NIS, Kazakhstan
- Carolyn Hutchinson, University of Glasgow, UK
- Stuart Shaw, Institute of Education, University College London, UK
- Irenka Suto, Cambridge University Press and Assessment, UK
- Tim Oates, Cambridge University Press and Assessment, UK



## Work-Related Assessment SIG

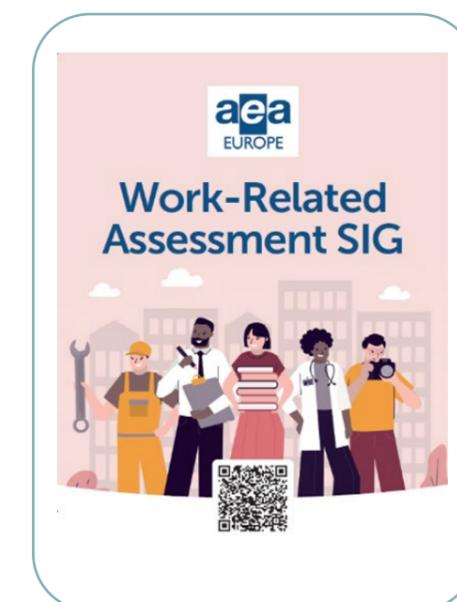
Launch of this new SIG is at the conference on Thursday, 6th November, 10.15 -10:45 GMT

The new Work-Related Assessment SIG provides a forum for the exchange of ideas, knowledge and information about work-related assessment. Work-related assessment encompasses a broad range of assessment settings, purposes and formats, reflecting the diverse nature of education and training for work across European nations. Terms such as vocational and technical, applied, work-based and competence-based are often contested, as are the nature of the learning programmes and skills and knowledge to be assessed, and how to ensure cross-border equivalence. Assessment may take many forms, even within one system, creating an overall picture that can sometimes be viewed as fragmentary. This SIG will provide a focal point for AEA-E members who share an interest in all aspects of work-related assessment. It will aim to draw together and develop specialist skills, knowledge and professional competences to support best practice, and will act as an expert resource to AEA-E and its members on knowledge, practice and the development of policies and strategies related to work-related assessment.

During the conference, please come and find out more at the 'Meet the SIGs' event. We would be very happy to welcome you to our SIG. Joining is free for AEA-E members, with a registration link on our webpage at <https://aea-europe.net/work-related-assessment-sig/>

### Work-Related Assessment SIG Steering Committee members:

- **Lena Gray** (Chair), Department of Education, University of Oxford
- Eleanor Andressen, Trinity College, United Kingdom
- Rebecca Conway, NCFE, United Kingdom
- Jo Fowler
- Jagdeep Kaur, Pearson, United Kingdom
- Nicola Mellor, Chartered Insurance Institute, United Kingdom
- Paul Newton, Ofqual
- Isabel Nisbet, Faculty of Education, University of Cambridge
- Stuart Shaw, Institute of Education, University College London



SIG conference flyers are available on the conference website:

<https://aea-europe.net/sigs-networks>



## Conference Programme 2025

### Wednesday, 05 November 2025

9:00 - 16:30	<b>Pre-Conference Workshops</b>
9:00	<p><b>Workshop 1</b>            Room: Rotterdam 2            Identifying and analysing evidence to determine whether tasks elicit the intended constructs: bridging the gap between modern validity theory and innovative validation practice  <i>Stuart Shaw, Ezekiel Sweiry</i></p>
9:00	<p><b>Workshop 2</b>            Room: Eindhoven 2            Exploring feedback dialogues for a transformative feedback culture  <i>Sam Passeport, Andrew Watts, Marianne Talbot, Nathalie Younès, Constanze Höpfner</i></p>
9:00	<p><b>Workshop 3</b>            Room: Eindhoven 3            Your best friend the psychometrician: The preventive role of psychometrics in test development  <i>Marieke Van Onna, Bas Hemker, Cor Sluijter</i></p>
9:00	<p><b>Workshop 4</b>            Room: Utrecht            From awareness to action: embedding inclusive assessment in teacher development programs in higher education  <i>Celine van der Lienden, Laurinde Koster</i></p>
9:00	<p><b>Workshop 5</b>            Room: Eindhoven 1            Establishing valid qualification equivalency with qualitative judgement  <i>Georgie Billings, Stuart Gallagher</i></p>
9:00	<p><b>Workshop 6</b>            Room: Rotterdam 3            Network Analysis for the investigation of Rater Effects (using R)  <i>Iasonas Lamprianou</i></p>
19:00 - 20:30	<b>Welcome Reception for all participants - NH Hotel (Venue)</b>

Thursday, 06 November 2025	
8:00	Registration
9:00 - 9:30	<p>Welcome Addresses Room: Holland Hall</p> <p>Prof. Damian Murchan (President AEA Europe) Dr. Shehzad Jeeva, Chief Education Officer, IB</p>
9:30 - 10:15	<p><b>Keynote Speech</b> Chair: Damian Murchan, AEA Europe President Room: Holland Hall</p> <p><b>All the World Is a Stage – Enter the Androids: Navigating Current and Future Challenges of International Baccalaureate Assessment</b> Matt Glanville</p>
10:15 - 10:45	<p><b>Coffee Break</b></p> <p>Coffee is served outside Holland Hall, in Amsterdam Lounge (Ground Floor) &amp; Outside meeting rooms (Second Floor)</p> <p>During Coffee Break 'Meet the SIGs' event at Holland Hall</p>

Thursday, 06 November 2025	
10:45 - 12:00	<p><b>Poster Presentations</b> Chair: Cor Sluijter Room: Holland Hall</p>
1	<p>Aligning National Mathematics Curriculum and West African Senior School Certificate Examination Questions: Experiences from Ghana and Nigeria <i>Alhaji Pateh Bah, Dr Adunola Osoba, Mr Kwaku Dankwa</i></p>
2	<p>Participant ecologies: how the unique ecology of every individual affects their engagement with and responses to my research questions about the impact of qualifying as a Chartered Educational Assessor <i>Marianne Talbot</i></p>
3	<p>Developing students' argumentation skills through situational tasks in preparation for external assessment <i>Lazzat Kabakbayeva, Bakhytzhhan Abylkhatov</i></p>
4	<p>Facilitating Learner Reasoning: Digital Annotation Tools for High-Stakes Assessment <i>Abdullah Khan</i></p>
5	<p>Students with disability as research advisors: Exploring fairness and justice in university assessment <i>Karina Dyliaeva</i></p>
6	<p>Designing Inclusive Assessments: Strategies for Equity and Accessibility in Education <i>Dina Tsagari</i></p>
7	<p>Washback effects on teaching practice: the Greek Language University Entrance Examination <i>Stavroula Tsiplakou, Dina Tsagari</i></p>
8	<p>Grading Effort in Physical Education. A Norwegian Case Study <i>Runa Korshavn, Egil Galaaen Gjølme, Lise Vikan Sandvik</i></p>
9	<p>Cambridge Early Adopter Programme: building a pathway to digital exams <i>Stephen Kemmery</i></p>
10	<p>Predictive Validity of Digital Mock Assessments <i>Carmen Vidal Rodeiro, Tim Gill</i></p>
11	<p>Assessing consistency in exam standards across sessions <i>Mkululi Wami, Antony Furlong</i></p>
12	<p>Utility of expert judgement for setting grade boundaries in England <i>Alejandra Miranda, Tom Benton</i></p>
13	<p>Sociodemographic characteristics, prior attainment and university choices in the UK: an intersectional approach <i>Diana Tonin</i></p>
14	<p>The effectiveness of group discussion, assessment criteria, peer assessment in improving the geography essay writing skills of 11 grade students <i>Madina Yeskeldi, Miras Kurbangaliyev</i></p>

Thursday, 06 November 2025	
15	Trends in between-school and between-class variation in student achievement in the Nordic countries based on IEA assessments from 2011 to 2023 <i>Sami Määttä, Sirkku Kupiainen, Jacob Christensen, Anders Christensen</i>
16	In the eye of the digital storm: A blind spot for fairness in screen-based learning? <i>Christopher Vincent</i>
17	From pen to keyboard: Is fairness affected by mode of input? <i>Faye Walker, Syeda Alina Husain, Christopher Vincent, Victoria Armstrong</i>
18	Simulating Expert Behavior with LLMs for Question Difficulty Estimation: Angoff-like procedure versus Pairwise Comparison <i>Diana Kolesnikova, Kirill Fedyanin, Matthieu J. S. Brinkhuis, Maria Bolsinova</i>
19	Stakeholders' Perceptions on Assessment Systems in Secondary Education <i>Joost Godschalk</i>
20	Efficiently and effectively increasing the role of formative assessment in large-scale programs <i>Astrid Kramer, Ya Ping Hsiao, Miranda Stienstra</i>
21	Lost in transcription: Why accuracy matters in AI marking <i>Katie Gilmartin, Lucy Howarth</i>
22	Authentic Learning Journeys through Collaboration <i>Miranda Stienstra, Astrid Kramer, Ya Ping Hsiao</i>
23	The Impact of Accessibility Features in GCSE Science Assessments <i>Emily De Groot</i>
24	Antecedents of AI's Perceived Usefulness in Grading: Insights from Norwegian High School Teachers <i>Harald Eriksen, Eyvind Elstad</i>
25	Using longitudinal datasets to explore labour market outcomes for learners taking Vocational and Technical Qualifications in England <i>Blake Ashworth, Hayley Dalton, Ria Bhatta</i>
26	A Theoretical Framework for Digital-First Assessment: Validity, Reliability, and Accessibility <i>Liyuan Liu, Irene Custodio</i>
27	The reliability of examiners' judgements of the likelihood of scripts being produced by background speakers of Japanese <i>Peter Hallsworth</i>
28	"I'll get a dog if I get an A" – Rewards as a Parental Strategy to Support Students' Academic Success <i>Dan-Anders Normann, Henning Fjørtoft, Lise Vikan Sandvik</i>

Thursday, 06 November 2025	
29	Peer Feedback? Maybe. Instructor Feedback? Definitely: Exploring Student Perceptions of Feedback in Higher Education <i>Nasser Hasan, Fajer Shamsaldeen</i>
30	Test Equating for Maintaining Standards Using Comparative Judgement <i>Kevin Mason</i>
31	Chatbots in Language Education: Challenges and Possibilities <i>Øystein Gilje</i>
32	Dealing with disruption: the evolving nature of teacher assessment practices in technology-driven formative assessment <i>Keith Menary</i>
33	Reference-Based Selection for Comparative Judgments with Natural Language Processing: Improving Efficiency of Large-Scale Writing Assessments <i>Michiel De Vrindt, Marije Lesterhuis, Anaïs Tack, Wim Van Den Noortgate, Renske Bouwer</i>
34	Development of CVD-accessible assessments as standard <i>Carly Hill-Banks, Gareth Roberts, Sally Melhuish</i>
35	Using AI and a clear set of rules to consistently and objectively improve the reading accessibility of assessments <i>Gareth Gillard, Jonathan Paul, Sally Melhuish</i>
36	AI-Powered Language Assessment: Can Linguistic and Cognitive Features Predict Item Difficulty and Discrimination? <i>Lina Loxley, Mark Elliott, Nadir Zanini</i>
37	From Selection to Success: Measuring Academic Competencies in High- and Low-Stakes Contexts <i>Milagrosa Sánchez-Martín, Juan F. Luesia, Juan Francisco Plaza</i>
38	Supporting the Identification of Learning Needs: Strengthening Science Teachers' Diagnostic Assessment Competence <i>Katrin Vaino, Triin Rosin, Helen Kann</i>
39	Exploring Students' Comprehension of Graphic Texts in the National Assessment of Reading Literacy <i>Tove Hårsaker</i>
40	Real Life – How Much Real Life? <i>Guri A. Nortvedt, Karianne Berg Bratting, Else Devold, Henrik Hung Haram, Oksana Kovpanets, Andreas Pettersen, Eren Sübül</i>

Thursday, 06 November 2025							
12:00 - 13:30	<b>Open Paper Session I</b>						
	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION
	Assessment Cultures I  Chair: Liyuan Liu Room: Amsterdam 4	Assessment of Practical Skills I  Chair: Mary Richardson Room: Amsterdam 3	Psychometrics and Test Development I  Chair: Sebastiaan de Klerk Room: Amsterdam 2	e-Assessment I  Chair: Graham Hudson Room: Rotterdam 1	National Tests and Examinations I  Chair: Dario Pirotta Room: Amsterdam 1	Artificial Intelligence and Assessment I  Chair: Rebecca Hamer Room: Holland Hall	Holistic Assessment  Chair: Carolyn Hitchinson Room: Eindhoven 1
12:00	How well are they? Assessing wellbeing <i>Isabel Nisbet, Stuart Shaw</i>	Designing a Test to Strengthen the Functional Literacy of Students at Nazarbayev Intellectual Schools <i>Diana Sartauova, Fariza Khamza</i>	How does different linkage plans affect the equating transformation? <i>Marie Wiberg</i>	Are schools ready for digital exams? – Creating a framework to evaluate digital readiness <i>Sanjay Mistry, Stephen Kemmerly</i>	Beyond Grammar – A Holistic Approach to KS2 Writing Assessment <i>Rebecca Clarkson</i>	Principles of AI use in high stakes marking <i>Joanna Williamson</i>	Debunking a dichotomy: An analysis of future skills in knowledge-rich qualifications <i>Irenka Suto, Sarah Nelson, Judith Roberts, Aman Sidhu, Lesley Spence</i>
12:30	Private judgments, public stakes: How can reliable final assessment be achieved in a culture of autonomy? <i>Lise Vikan Sandvik, Marita Byberg Johansen, Anne Holten Kvistad, Bodil Svendsen</i>	The validity of virtual labs for assessing science practical skills <i>Emma Walland, Alison Rodrigues</i>	A Peculiarity in Educational Testing Practices <i>Mark White</i>	Assessment mode effects and their relationship with item characteristics <i>Carmen Vidal Rodeiro, Carmen H.J. Lim</i>	Using Aberrant Responses to Identify Backwash in National Mathematics Tests <i>Marta Mikite, Girts Burgmanis</i>	Responsible Use of AI in Research and Assessment: Emphasizing Reliability through Repeated Evaluations <i>Elena Papanastasiou, Agni Stylianou Georgiou, Georgios Lampropoulos</i>	Understanding the construct of ‘future skills’: some critical reflections emerging from a comparison of recent and older future skills frameworks <i>Filio Constantinou, Alejandra Miranda, Silvia Pirola</i>
13:00	Critical consciousness and well-being: Questioning and transforming teachers’ relationship with assessment <i>Catarina F. Correia</i>	How are oracy skills assessed in England and internationally? <i>Latoya Clarke, Stephen Holmes</i>	Effects of using IRT scoring on admission outcomes <i>Per-Erik Lyrén, Inga Laukaityte, Mattias Wickberg, Torbjörn Lindquist, Håkan Pettersson</i>	HyPES project - assessment of learning in hybrid learning environments in higher education: what methods and tools? <i>Noben Natasha, Nathalie Younès, Julie Pironom, Réginald Burton, Luc Massou</i>			Creativity at the Crossroads of STEM and Non-STEM Subjects: Pathways to Holistic Education <i>Gulbakhyt Sultanova, Aikorkem Zhapparova, Yerbol Almenov</i>
13:30 - 14:30	<b>Lunch</b> Room: Hotel restaurant - “The Verne” - First floor and Rotterdam 2 and 3 Halls - Second floor						

# Programme

Thursday, 06 November 2025

14:30 - 15:30

## Discussion Groups

### Discussion Group 1

Room: Amsterdam 3  
How do we reconfigure assessment to meet the challenges of cultures driven by technology?  
*Eleanor Andressen, Graeme Clark, Stuart Shaw*

### Discussion Group 2

Room: Eindhoven 1  
The 'Assessment Cultures' SIG – who are we and what do we want to be talking about?  
*Christoph Schneider, Sam Passeport, Marte Søve Syverud, Andrew Watts, Nathalie Younès*

### Discussion Group 3

Room: Rotterdam 1  
AI-supported translation tool for linguistically inclusive classrooms  
*Tony Burner*

### Discussion Group 4

Room: Eindhoven 2  
Should we have confidence in high-stakes assessments?  
*Isabel Nisbet, Mary Richardson, Stuart Shaw, Lesley Wiseman-Orr*

### Discussion Group 5

Room: Amsterdam 4  
Integrating assessment and curriculum design: Promoting authentic connections between learning and assessment  
*Jo Hayman, Vicki Westwell, Jude Scanlon*

### Discussion Group 6

Room: Amsterdam 2  
Beyond Alignment: Assessment at the Intersection of Constructiveness, Accountability, and Tacit-Explicit Knowledge Dynamics  
*Ya Ping Hsiao, Astrid Kramer, Miranda Stienstra*

### Discussion Group 7

Room: Holland Hall  
Artificial Intelligence for Holistic Learning: Challenges and Possibilities in Assessing Collaboration and Empathy  
*Gulbakhyt Sultanova, Irenka Suto, Carolyn Hutchinson*

### Discussion Group 8

Room: Eindhoven 3  
Universal Design for Assessment: What it is, what it is not and the tensions it must balance  
*Kala Parasuram, Mariana Manzo*

### Discussion Group 9

Room: Utrecht  
From grades to growth: Reframing the dialogue on learning, assessment and mastery  
*Kathinka Blichfeldt, Egil Weider Hartberg, Ingrid Jacobsen, Kaja Haaland, Trude Slemmen Wille*

### Discussion Group 10

Room: Amsterdam 1  
Why should assessment of vocational or professional education and training matter to this conference?  
*Lena Gray, Rebeca Conway, Nicola Mellor, Horacy Debowski, Victoria Morgan*

aea '25  
EUROPE



Conference Programme 2025



**Thursday, 06 November 2025**

15:30 - 16:30							Open Paper Session II								
ORAL SESSION		ORAL SESSION		ORAL SESSION		ORAL SESSION		ORAL SESSION		ORAL SESSION		ORAL SESSION			
International Assessments I  Chair: Judith Gracey Room: Rotterdam 1		Assessment Cultures II  Chair: Eleanor Andressen Room: Amsterdam 3		Artificial Intelligence and Assessment II  Chair: Therese Hopfenbeck Room: Holland Hall		Comparative Judgement I  Chair: Antony Furlong Room: Eindhoven 1		e-Assessment II  Chair: Graeme Clark Room: Amsterdam 4		Perspectives of End-users and the General Public on Assessment I  Chair: Gulbakhyt Sultanova Room: Amsterdam 2		Psychometrics and Test Development II  Chair: Mareka Van Onna Room: Amsterdam 1			
15:30		Standards in the wild: Setting and maintaining standards in a high-stakes global context <i>Lauren Miller, Sarah Hughes</i>		Determinants and consequences of behavioral grades in schools <i>Ragnar Alne, Andreas Fidjeland, Rune B. Reiling</i>		A Comparative Analysis of AI-Assisted and Human Interpretations of the Bloom's Taxonomy in the Classification of Test Items <i>Oluyemi Adetoro, Dauda Adeniyi Lawal</i>		Simplified pairs comparative judgement method for equating via expert judgement of script quality – further evidence of accuracy and bias <i>Milja Curcin, Ming Wei Lee</i>		Unpacking Student Performance: The influence of Process data on International English Language Exam Results <i>Liyuan Liu, Blake Ashworth, Hayley Dalton</i>		Curriculum for Wales: Insights from the Inaugural Stakeholder Surveys <i>Gillian Parkinson, Ben Tylden-Smith, Hefin Thomas, Joanne Davis, Laura Verdasco Menendez, Claire Sinnema</i>		Assessing Beyond the Score: The Role of Differential Item Functioning in Valid and Fair English Language Testing <i>Joanne Topham</i>	
16:00				Joining Forces: Teachers and AI in Early Reading Risk Detection <i>Per Henning Uppstad, Bente Walgermo, NjAL Foldnes, Kjersti Bjønnes, Annette Aarflot</i>				Student approaches to annotation in onscreen assessments: Exploring annotation techniques and their impact on test-taker strategie <i>Zoe Mair, Irene Custodio</i>		What Shapes Realisation of the New Curriculum for Wales? Understanding the Perceptions of Teachers, Learners and Parents/ Carers Using a New Theory of Change <i>Laura Verdasco Menendez, Ben Tylden-Smith, Claire Sinnema, Hefin Thomas, Gillian Parkinson, Joanne Davis</i>					
16:30 - 16:45							<b>Coffee Break</b> Coffee is served outside Holland Hall, in Amsterdam Lounge (Ground Floor) & Outside meeting rooms (Second Floor)								

	<b>Thursday, 06 November 2025</b>						
16:45 - 18:15	<b>Open Paper Session III</b>						
	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>
	<b>Fairness &amp; Social Justice I</b>	<b>Artificial Intelligence and Assessment III</b>	<b>Formative Assessment I</b>	<b>e-Assessment III</b>	<b>National Tests and Examinations II</b>	<b>Assessment of Practical Skills II</b>	<b>Assessment Cultures III</b>
	Chair: Ben Stafford Room: Eindhoven 1	Chair: Alex Scharaschkin Room: Holland Hall	Chair: Stuart Shaw Room: Amsterdam 4	Chair: Irene Custodio Room: Amsterdam 1	Chair: Lesley Wiseman-Orr Room: Rotterdam 1	Chair: Andrew Boyle Room: Amsterdam 3	Chair: Sam Passeport Room: Amsterdam 2
16:45	<b>Predictive Validity of Unified National Testing in Kazakhstan</b> <i>Elena Kardanova, Lyazzat Shinetova, Bolatbek Abdrassilov, Alina Ivanova, David Orlov</i>	<b>Evaluating AI assisted auto-marking systems using a range of item types</b> <i>Mark Frazer, Sanjay Mistry</i>	<b>Navigating the Path to Effective Learning: Progressions in Assessment and Feedback Literacy</b> <i>Vanessa Scherman, Fabienne van der Kleij, Rayanne Shakra</i>	<b>What examiners know: to what extent is detailed content knowledge necessary for reliable marking in broad, open-ended assessments?</b> <i>Maria Galache Ramos, Mkululi Wami</i>	<b>"It didn't work well for me, but it works overall": Exploring the relationship between validity and trust in a university entrance examination</b> <i>Pok Jing (Jane) Ho</i>	<b>Foundations for Success: Reimagining Accessible Assessment Across the Years</b> <i>Marijne Medhurst, Julie McMillan, Greta Rollo, Kathryn Richardson</i>	<b>Effectiveness of using open-ended items in high-stake examinations</b> <i>Narmin Abdurahmanova, Shukufa Mirzayeva</i>
17:15	<b>Gender Gap in Test Scores: Do Different Guessing Approaches and Test Creators' Gender Have an Impact?</b> <i>Lenka Firtova</i>	<b>Artificial Intelligence-Driven Test Item Generation: Enhancing Quality and Efficiency in Assessment Design</b> <i>Richard Ekpo</i>	<b>Learning Oriented Assessment at an institutional level: Key design and validity prerequisites for positive impact</b> <i>Angeliki Salamoura</i>	<b>Shifting Landscapes for Screening of Reading Difficulties in the Early Grades</b> <i>Bente Walgermo</i>	<b>Developing qualifications to support the Curriculum for Wales and to promote positive backwash</b> <i>Delyth Jones</i>	<b>Measuring Literacy Skills for Young Learners: A School Grade-Based and Pluriliteracies-Informed Approach</b> <i>Laura Katkeviča</i>	<b>The transition to the new SEC model. Did it succeed in providing a learner-centred model or did a onesize fits all model remain in place?</b> <i>Dario Pirotta, Malcolm Micallef, Francois Zammit, Lorraine Vassallo, Analise Grixti, John Muscat</i>
17:45	<b>Investigating the Validity and Fairness of the Mandarin Leaving Certificate Examination in Ireland</b> <i>Xiao Huang, Damian Murchan</i>	<b>Investigating the Validity of ChatGPT in the Assessment of Open-Ended University Tasks: A Comparative Study with Human Grading</b> <i>Giuseppe Carmelo Pillera</i>	<b>School leaders' and teachers' attributions in making sense of norm-referenced school performance data</b> <i>Glen Molenberghs, Roos Van Gasse, Jan Vanhoof, Sven De Maeyer, Evelyn Goffin</i>	<b>Typing Speed as a Proxy for Digital Literacy in Onscreen Assessments</b> <i>Sumita Ishaque, Ellen Barrow, Ana Ulicheva</i>			<b>Assessment of Transversal Competences in Vocational Education and Training. Theoretical and Practical Considerations</b> <i>Dr Lena Gray, Dr Horacy Debowski</i>
18:15 - 19:00	<b>General Assembly Room: Holland Hall</b>						
18:15 - 19:00	<b>PhD Students Room: Bar in the "Verne" restaurant in the first floor</b>						
19:30 - 21:00	<b>Event for AEA Europe Fellows, Practitioners &amp; PhD students Location: VOCO</b>						

	Friday, 07 November 2025						
9:00 - 10:30	Open Paper Session IV						
	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION
	Assessment Cultures IV Chair: Andrew Watts Room: Rotterdam 1	Assessment of Practical Skills III Chair: Iasonas Lamprianou Room: Amsterdam 3	Psychometrics and Test Development III Chair: Michalis Michaelides Room: Amsterdam 2	Inclusive Assessment I Chair: Katy Finch Room: Amsterdam 4	Formative Assessment II Chair: Isabel Nisbet Room: Amsterdam 1	Artificial Intelligence and Assessment IV Chair: Elena Papanastasiou Room: Holland Hall	Fairness and Social Justice II Chair: Tim Stratton Room: Eindhoven 1
9:00	External moderation of school based assessment: A review of systems' experiences with implementation and factors associated with selecting approaches <i>Damian Murchan, Stuart Shaw, Evgenia Likhovtseva</i>	Borderline Regression Analysis: A Data-Driven Framework for Fairer Educational Decisions <i>Dr Thomas Kropmans</i>	Exploring the impact of optimising the consistency of examiner scores used in model training on automarker performance <i>Trevor Breakspear, Hannah Bouteba, Mark Brenchley, Abhirup Chakravarty, Ian Lewin, Theerdha Sajimon, Yan Huang</i>	Redesigning assessments for inclusion: A developmentally-aligned, criterion-related approach for students aged 11-16 <i>Shwetangna Chakrabarty, Shalini M Easwardas, Lenny Dutton</i>	Making Computational Thinking visible as framework for inquiry and assessment <i>Tim Knight, Ross Robertson</i>	Towards Scalable Personalised Feedback: Technology-Supported Insights into Student Feedback Use and Self-Regulated Learning <i>Fabienne van der Kleij, Anastasiya Lipnevich, Therese Hopfenbeck</i>	Towards the automated optimisation of examination time-tables <i>Antony Furlong, Dick Den Hertog, Julien Grand Clement</i>
9:30	Implementing an Assessment Framework in Cox's Bazar refugee camps: two years on <i>Georgie Billings, Anita Oomeer</i>	Implications of different types of test items on applicants' results in school-leaving exams <i>Shukufa Mirzayeva, Narmin Abdurahmanova</i>	Targeting the top: Evaluating differentiation among high-performing candidates in items designed to stretch and challenge <i>Ana Ulicheva, Sarah Hughes</i>	The Way You Teach Matters! How Different Teaching Practices to Develop Reading Comprehension Skills Impact Italian Students' Reading Abilities <i>Silvia Mazzuca</i>	The effects of the launch of the Learning Outcomes Framework on teaching, learning and assessment in the French as Foreign Language Classroom. The case of Malta <i>Ruth Bonello</i>	Enhancing automated essay scoring with artificial intelligence (ai): the role of natural language processing (nlp)-based assessment tools in improving Scoring systems <i>Olaitan Oyeniya, Rosemary Ojo-Odide</i>	Strengthening Assessment and Learning through Culturally and Linguistically Responsive Design <i>Louise Badham, Camelia Constantinescu</i>
10:00	Bridging perception and performance in teacher appraisal: a multi-faceted study from Kazakhstan <i>Nuryam Shora, Baurzhan Yessingeldinov, Aidana Shilibekova, Akmaral Zhumykbayeva</i>	The challenge of designing an inclusive, high-stakes, digital baseline assessment for 4-year-olds in England <i>Miranda Hornett, Claire Whitby</i>		Enhancing inclusion and cross-cultural validity in PISA. Towards an improved index of economic, social and cultural status <i>Francesco Avvisati, Said Ettejjari, Alejandra Osses Vargas</i>		The future of non-examined assessment (NEA) – malpractice detection and the challenges and opportunities of generative AI <i>Tony Leech, Raimonda Laugalyte</i>	
10:30 - 11:00	Coffee Break - Coffee is served outside Holland Hall, in Amsterdam Lounge (Ground Floor) & Outside meeting rooms (Second Floor)						
11:00 - 11:45	<p>Keynote Speech</p> <p>Chair: Stuart Shaw, AEA Europe Council Member Room: Holland Hall</p> <p>Raising educational standards</p> <p>Jo-Anne Baird, University of Oxford</p>						

	Friday, 07 November 2025						
11:45 - 13:15	Open Paper Session V						
	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION
	e-Assessment IV Chair: Dee Arp Room: Amsterdam 3	National Tests and Examinations III Chair: Catarina Correia Room: Amsterdam 2	Assessment Cultures V Chair: Lise Vikan Sandvik Room: Rotterdam 1	Artificial Intelligence and Assessment V Chair: Vanessa Scherman Room: Holland Hall	Comparative Judgement II Chair: Amina Afif Room: Eindhoven 1	Test Development Chair: Raphaël Pasquini Room: Amsterdam 4	Other Chair: Dan-Anders Normann Room: Amsterdam 1
11:45	Face Validity in Focus: Engaging Test-Takers in the Development of High-Stakes English Language Tests <i>Ellen Barrow</i>	Evidence-Based Prompting: Searching for the Holy Grail <i>Ondrej Hrubes</i>	Dialogic Assessment Conversations as Dynamic and Learning-Oriented Assessment in English Language Education <i>Elin Løvnæseth Hauer</i>	AI in Assessment: So much more than ChatGPT <i>Nathan Thompson, Tammy Trierweiler</i>	Graded Response Modeling of Forced-Choice Measures <i>Dimiter Dimitrov, Dimitar Atanasov</i>	The impact of response format on performance in a maths test for gifted and talented primary children <i>Ellie Darlington</i>	Redesigning assessment objectives for authentic and aligned learning and assessment: moving away from Bloom's taxonomy <i>Sabina Sorrentino</i>
12:15	Automating Effective Feedback for Comparative Judgment Assessments of Essays <i>Michiel De Vrindt, Tine van Daal, Marije Lesterhuis, Davy Mortier, Wim Van Den Noortgate, Anaïs Tack, Renske Bouwer</i>	Beyond oracy – 30 years of using Oral assessments to assess candidates' knowledge in high stakes exams at 16 and 18. <i>Alun Evans, Joanna Maziarz</i>	Are political beliefs associated with assessment policies? <i>Michelle Meadows, Jo-Anne Baird</i>	Abilities aren't numbers: Alfred Binet and the road not taken in educational assessment <i>Alex Scharaschkin</i>	Investigating the validity of comparative judgement: What influences judges' decisions when assessing STEM subjects? <i>Merlin Walter</i>	Key principles for determining the optimal item format <i>Ezekiel Sweiry</i>	Decolonising Assessment in South African Education: Towards an Ubuntu-Centered Approach <i>Anil Kanjee</i>
12:45		Scottish educators' attitudes towards AI use in the classroom: A mixed methods consultation <i>Maggie Mroczkowski, Jake Green, Jamie Lawson, Rachel Whitford</i>	Reframing Dialogues of Learning, Assessment and Mastery: Insights from Norwegian Lower Secondary Schools <i>Egil Weider Hartberg, Kathinka Blichfeldt, Ingrid Jacobsen, Kaja Haaland, Trude Slemmen Wille</i>	Evaluating AI-Generated Feedback: Insights on usefulness, accuracy, relevance and actionability <i>Lucy Howarth, Katie Gilmartin</i>	Beyond Reliability: Challenges of Using Adaptive Comparative Judgment in High-Stakes Legal Assessment <i>Kjetil Egelandstad, Jan-Ove Færstad</i>	Achievement Levels in Irish Literacy among Post-Primary Students Educated in Irish-Medium Schools <i>Muireann Bergin, Helena Ní Rócháin, Conall Ó Duibhir, Brenda Donohue</i>	An adaptive national assessment: technical and policy challenges and insights <i>Andrew Boyle, Chris Hope</i>
13:15 - 14:15	<p style="text-align: center;"><b>L u n c h</b></p> <p style="text-align: center;">Hotel restaurant - "The Verne" first floor and Rotterdam 2 and 3 Halls - second floor</p>						

Friday, 07 November 2025							
<b>Open Paper Session VI</b>							
	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION	ORAL SESSION
	Artificial Intelligence and Assessment VI Chair: Bas Hemker Room: Holland Hall	Assessment Cultures VI Chair: Christoph Schneider Room: Amsterdam 4	Formative Assessment III Chair: Ezekiel Sweiry Room: Rotterdam 1	National Tests and Examinations IV Chair: Irenka Suto Room: Eindhoven 1	Fairness and Social Justice III Chair: Guri Nortvedt Room: Amsterdam 2	Summative Assessment Chair: Grace Grima Room: Amsterdam 3	Higher Education and Assessment Chair: Dina Tsagari Room: Amsterdam 1
14:15 - 15:45	Improving OSCE Feedback Using ChatGPT-4.0 Dr Thomas Kropmans, Kamilla Mahrlamova, Akram Alshafi	Time, the one-handed clock, and educational certification theory Richard Harry	Exploring the Formative Potential of Standardized Reading Tests: Enhancing students' competence in critical reading through assessment results Cecilie Weyergang, Tove Stjern Frønes	Exploring the use of similar items for monitoring grading standards Charlotte Draper	From Deficit to Asset: Reframing Validity Through Linguistic Diversity in Classroom-Based Assessment Pooja Nagpal	Enhancing science teachers' assessment literacy through authentic modelling of multiple assessment methods Berit S. Haug, Sonja M. Mork	Understanding the Access Arrangements System in Wales Amy Harrison
14:15	AI-Driven Predictions of Mathematics and Science GCSE Exam Results using Mock Papers Sebastian Nastuta	The importance of political and cultural embeddedness in the enactment of Assessment for Learning: Lessons from the "Network of schools for Assessment Innovation (Red sin Notas)" in Chile Maria Teresa Florez Petour	Formative Assessment with Impact: What can we learn from different contexts Brooke Wyatt	Raising educational standards in the Kingdom of Bhutan: a case study of an assessment framework James Frith, Georgie Billings	Minimising construct-irrelevant barriers for neurodivergent learners in assessment: Exploring neuroinclusivity by design Rebecca Chivers, Louise Badham	Going gradeless in upper secondary school. A multiple case study of a natural experiment Henning Fjørtoft, Lise Vikan Sandvik, Iveta Kohanová, Sverre Tveit, Dan-Anders Normann, Arne Amdal, Stine Aarønes Angvik, Andrea Seljeseth, Kari Smith	To Test or Not to Test. Evolving Landscape of Higher Education Admissions in the Czech Republic and Slovakia Andrej Kutarna
14:45	Incorporating structural elements in Automated Question-Difficulty Estimation models can improve model accuracy and efficiency Gabriel Ortega, Abelino Jiménez, Séverin Lions, Pablo Dartnell			The Predictive Validity of High-Stakes Testing: A Meta-Analysis of Long-Term Academic and Career Outcomes Rosemond Wilson, Boateng Owusu-Ansah	High Stakes and High Accountability: Public Post-Test Item Review in Czech University Admission Process Barbora Lunakova		
15:15	<b>Coffee Break</b> Coffee is served outside Holland Hall, in Amsterdam Lounge (Ground Floor) & Outside meeting rooms (Second Floor)						
15:45 - 16:15	<b>Coffee Break</b> Coffee is served outside Holland Hall, in Amsterdam Lounge (Ground Floor) & Outside meeting rooms (Second Floor)						

	<b>Friday, 07 November 2025</b>				
16:15 - 17:45	<b>Symposium Session</b>				
	<b>SYMPOSIUM</b>	<b>SYMPOSIUM</b>		<b>SYMPOSIUM</b>	<b>SYMPOSIUM</b>
	Symposium Session 1 Integrating assessment and curriculum design: harnessing analysis of TIMSS data to positively impact the enacted curriculum  Chair: Mary Richardson Discussant: Jennie Golding Room: Rotterdam 1	Symposium Session 2 Assessing higher order thinking skills in an Assessment for Learning perspective: Challenges and prospects facing teachers in Switzerland  Chair: Raphaël Pasquini Discussant: Tine Prøitz Room: Amsterdam 1		Symposium Session 3 Capturing Diverse Student Voices in Assessment  Chair: Katy Finch Discussant: Bryan Maddox Room: Amsterdam 3	Symposium Session 4 Rating systems for dynamic assessment in adaptive learning environments  Chair: Maria Bolsinova Discussant: Matthieu Brinkhuis Room: Amsterdam 4
16:15	Understanding TIMSS data: an insider critique of the frameworks, assessments and questionnaires <i>Grace Grima</i>	Under which conditions implementing criterion-based assessment? <i>Raphaël Pasquini</i>		Study on the education and well-being of disabled pupils in France <i>Elodie Persem</i>	Tracking progress in subdomains: When are multiple ratings better than one? <i>Angelina Kuchina</i>
16:45	TIMSS 2023 evidencing mathematics and science education as widely gendered endeavours: the issues and some pedagogic (and broader) responses <i>Jennie Golding</i>	Does a new curriculum require new assessment tools? <i>Robbert Smit</i>		Trialling a New Numeracy Product: Capturing students' experiences of maths learning in Further Education and Alternative Provision settings <i>Sheradan Miller, Faye Walker</i>	Extending the Elo rating system for between-item multidimensionality <i>Hanke Vermeiren</i>
17:15	We don't belong here: views on school belonging absenteeism and achievement <i>Mary Richardson</i>	How can teachers implement a competency-based assessment embedded in a backward planning approach? <i>Miriam Salvisberg</i>		Multilingual students' performance in high school history classes: teaching and assessment needs <i>Thomas Rousoulioti</i>	Elo vs. Urnings: A comparison of the measurement properties of different rating-based algorithms in adaptive learning systems <i>Bence Gergely</i>
20:00 - 22:00	<b>Conference Dinner</b> Location: The Kurzaal of the Amrath Kurhaus hotel in Scheveningen				

	<b>Saturday, 08 November 2025</b>						
9:00 - 9:45	<p align="center"><b>Keynote Speech</b>                  Chair: Cor Sluijter, AEA Europe Vice President    Room: Holland Hall</p> <p align="center">Assessment as a Tool for Liberation: Come Dream with Me                  Jennifer Randall</p>						
9:45 - 11:15	<b>Open Paper Session VII</b>						
	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>	<b>ORAL SESSION</b>
	<b>Assessment Cultures VII</b>	<b>Artificial Intelligence and Assessment VII</b>	<b>International Assessments II</b>	<b>e-Assessment V</b>	<b>Technical, Vocational and Applied Assessments</b>	<b>Formative Assessment IV</b>	<b>Inclusive Assessment II</b>
	Chair: Antony Furlong Room: Eindhoven 1	Chair: Rose Clesham Room: Holland Hall	Chair: Alison Matthews Room: Amsterdam 2	Chair: Cor Sluijter Room: Amsterdam 1	Chair: Lena Grey Room: Rotterdam 1	Chair: Doreen Said Pace Room: Amsterdam 3	Chair: Henning Fjørtoft Room: Amsterdam 4
9:45	The role of Irelands National Assessments in policy and curriculum development <i>Joanne Kiniry, Sharon Nelis</i>	Authoring with AI using a human in the loop approach – is this the future of exam authoring? <i>Kay Evans Johns, Fraser Lawrence</i>	Do students respond more inconsistently to low-stakes conditions? An experimental study manipulating the stakes of the assessment <i>Michalis Michaelides, Evi Konstantinidou</i>	Exploring the relationship between students’ use of ICT and performance in PISA Mathematics, Reading and Science digital assessments <i>Irene Custodio, Sebastian Nastuta, Liyuan Liu</i>	Investigating the relationship between problem-solving elicitation, language demands and difficulty <i>Diana Tonin, Ezekiel Sweiry</i>	Sustainable Assessment-for-Learning (AFL): Developing Self-regulated Learning through Evaluative Judgement <i>Sima Caspari-Sadeghi, Johan Bredberg, Brigitte Forster-Heinlein, Moritz Proell, Khaled Ben Latief Jemaie</i>	Qualifications to Support Curriculum for Wales – The new National 14-16 Qualification Suite <i>Oliver Stacey</i>
10:15	Informing the Future, Learning from the Past: Reflections from 25 years of empirical research on Leaving Certificate assessment in Ireland <i>Paula Lehane, Michael O’Leary, Gillian O’Connor</i>	Maintaining test integrity through integration of human judgement and AI systems <i>Sarah Hughes</i>	Is two years of teaching for Open Book Exams enough to change teacher practice? <i>Rebecca Hamer, Rebecca Chivers, Vanessa Scherman</i>	Catering to All Learners: Adapting Digital Assessments for Students with Hearing and Visual Impairments <i>Guri A. Nortvedt</i>	Stability of Item Response Theory (IRT) equating based on different sample sizes for functional skills mathematics <i>Zeeshan Rahman, Richard Harris</i>	Data literacy for educators: A model for transforming data and information into instructional knowledge and practice <i>Todd Milford, Victor Glickman</i>	Weak Signals and Wild Cards: An anticipatory framework for the validation of emergent immersive, interactive, adaptive and sensor-based assessments <i>Vahid Aryadoust, Bryan Maddox</i>
10:45	Empowering Student Agency: A Didactical Model for Designing Learning and Assessment in Technology-Rich Classrooms <i>Øystein Gilje</i>		What you say is what you get? An Experimental Study on the Role of Test Introduction in Test-Takers’ Motivation, Anxiety and Performance <i>Daphné Van Looy, Amélie Rogiers, Carolien Frijns, Maarten Vansteenkiste, Johan van Braak</i>	Enhancing measurement in adaptive learning systems with response time data <i>Maria Bolsinova, Matthieu Brinkhuis, Bence Gergely, Abe Hofman</i>		Transforming teacher feedback: A checklist for supporting student agency in oral reading <i>Karianne Megard Gronli, Bente Rigmor Walgermo, Per Henning Uppstad, Erin Margaret McTigue</i>	
11:15 - 11:45	<p align="center"><b>Coffee Break</b></p> <p align="center">Coffee is served outside Holland Hall, in Amsterdam Lounge (Ground Floor) &amp; Outside meeting rooms (Second Floor)</p>						

# Programme

## Saturday, 08 November 2025

11:45 - 12:30

**Keynote Speech**  
Chair: Deborah Chetcuti, AEA Europe Council Member  
Room: Holland Hall

**Who Tries and When in the Digital Age: Measuring and Modeling Test-Taking Effort through Process Data in Large-Scale Assessments**

Militsa Ivanova, University of Cyprus

12:30 - 13:00

**Closing Ceremony including Poster Award & Accreditation Awards**  
Chair: Elena Papanastasiou  
Room: Holland Hall

**Parting comments to delegates at the end of the Closing Ceremony**  
Damian Murchan, President AEA Europe

13:00 - 14:00

**Lunch**  
Room: Hotel restaurant - "The Verne" - first floor and Rotterdam 2 and 3 Halls - second floor

Social Events

# Social Events

# About The Hague

Welcome Cocktail

Date: 05 November, 2025  
Time: 19:00 – 20:30  
Where: NH Hotel

Welcome Cocktail is the first social gathering between all conference delegates and it will take place at the Venue Hotel. It will be a relaxing evening during which delegates will have the opportunity to talk to colleagues and peers, while enjoying local drinks and ample canapés.



Conference Dinner

Date: 07 November, 2025  
Time: 20:00 – 22:00

We will gather in the lobby of the Venue Hotel to travel to the conference dinner location. Once there, participants will be treated to extensive buffets featuring the finest local specialties, complemented by a selection of drinks.



The Hague, known as Den Haag in Dutch, is a dynamic city located on the western coast of the Netherlands. It serves as the administrative and political heart of the country, hosting the Dutch government, Parliament, and the official residence of the King. Renowned as the “International City of Peace and Justice,” The Hague is home to numerous international organizations, including the International Court of Justice and the International Criminal Court.

Rich in history and culture, the city features stunning architecture, from the medieval Binnenhof complex to modern landmarks like the Peace Palace. Visitors can explore world-class museums, such as the Mauritshuis, which houses masterpieces like Vermeer’s Girl with a Pearl Earring. The city’s seaside district, Scheveningen, offers sandy beaches, a lively boardwalk, and a vibrant cultural scene.

With its blend of historical significance, international influence, and coastal charm, The Hague is a unique destination that embodies a perfect mix of tradition and modernity.

More information on what to do in The Hague when you are there can be found here: <https://denhaag.com/en>



### Top Attractions in The Hague

**The Binnenhof** One of the oldest parliamentary buildings still in use, the Binnenhof is the political heart of the Netherlands. This Gothic-style complex surrounds the Hofvijver lake and is open for guided tours, offering insights into Dutch history and governance.

### The Peace Palace

This iconic building is the symbol of international justice, housing institutions like the International Court of Justice. Visitors can explore the landscaped gardens and visitor center to learn about the palace’s role in promoting global peace.

### Mauritshuis Museum

Located near the Binnenhof, this exquisite museum is a treasure trove of Dutch Golden Age masterpieces, including Vermeer’s Girl with a Pearl Earring and Rembrandt’s The Anatomy Lesson of Dr. Nicolaes Tulp. The elegant 17th-century building itself is a work of art.

### Escher in Het Paleis

This museum, dedicated to the works of M.C. Escher, is housed in a former royal palace. It showcases the artist’s mind-bending lithographs, woodcuts, and drawings, offering a surreal journey through optical illusions and mathematical art.

### Panorama Mesdag

This unique museum features a 19th-century cylindrical painting by Hendrik Willem Mesdag, depicting a 360-degree view of the Scheveningen seaside. It immerses visitors in the coastal landscape of the time.

# About The Hague

## Madurodam

A favorite for families, Madurodam is a miniature park that showcases famous Dutch landmarks and cities in stunning detail. It's a fun way to learn about the Netherlands' history, architecture, and innovations.



## Scheveningen Beach

A lively seaside district with wide sandy beaches, Scheveningen is perfect for a day by the water. Highlights include the iconic pier, with its restaurants, shops, and a Ferris wheel, as well as water sports and vibrant nightlife.

## Grote Kerk (The Great Church)

This imposing medieval church, also known as St. James' Church, is famous for its striking architecture and unique carillon with 51 bells. It's often used for concerts and exhibitions.

## Haagse Bos and Huis ten Bosch Palace

One of the oldest forests in the Netherlands, the Haagse Bos offers a serene escape in the city. It's also home to Huis ten Bosch, the royal family's residence, which adds a touch of regal charm.

## The Gemeentemuseum (Kunstmuseum)

This modern art museum boasts an impressive collection, including works by Piet Mondrian, fashion exhibits, and decorative arts. Its sleek Art Deco design by architect Berlage is a masterpiece in itself.

## Other Highlights

**Noordeinde Palace** A working palace of the Dutch royal family, Noordeinde Palace is surrounded by luxury shops and charming cafes, making it a delightful area to explore.

## The Hague Market (De Haagse Markt)

One of the largest multicultural markets in Europe, this bustling market is a hub for fresh produce, exotic spices, and diverse street food.

## Lange Voorhout

A historic tree-lined avenue filled with cafes and galleries, Lange Voorhout hosts seasonal events like antique fairs and winter markets.

## Nearby Excursions

Delft

A short trip from The Hague, Delft is famous for its blue pottery, charming canals, and connection to the painter Johannes Vermeer.

## Clingendael Park

Known for its stunning Japanese Garden (open in June), this park offers beautiful walks and serene landscapes.

## Why Visit The Hague?

The Hague combines the vibrancy of a modern city with a profound sense of history and culture. Its mix of attractions, from world-class art to seaside relaxation, ensures there's something for everyone. Whether you're drawn to its political significance, coastal beauty, or artistic treasures, The Hague promises a memorable experience.

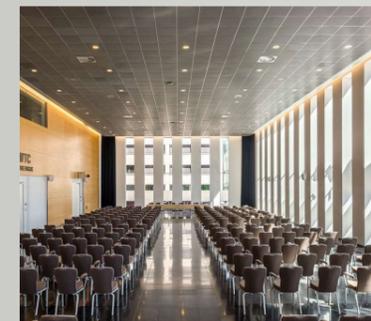
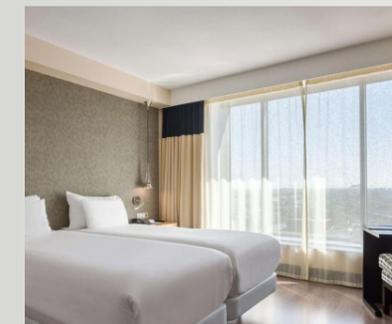
# Venue

## NH Den Haag

NH Den Haag is a modern and stylish hotel located in the heart of The Hague's business district, offering easy access to both the city center and key corporate hubs. Just a short walk from public transport links, the hotel provides a convenient base for business travelers and tourists alike.

The spacious and contemporary rooms feature comfortable furnishings, free Wi-Fi, and stunning city views. Guests can enjoy a delicious breakfast buffet, an on-site restaurant serving international cuisine, and a cozy bar for evening drinks. Additional amenities include a well-equipped fitness center and flexible meeting spaces for events and conferences.

Address: Pr. Margrietplantsoen 100, 2595 BR Den Haag, Netherlands



# Sponsors



we take care of every detail

for your conference needs

## Host Institution



## Coordinator



Easy Conferences Ltd has been in business since 1992 and has been specializing in the complete coordination and organization of conferences and all related activities. Through the development of its own online registration software, the company has expanded its operations outside Cyprus. We have extensive experience in organizing events ranging from 20 to 3000 participants for physical, hybrid or online participation. We consult, manage and assist in every step of the process of any event and we deliver top professional services throughout.

Our services extend from digital support, media promotion, conference website development and management, to the management of all conference related activities, complete interaction with suppliers and participants, online/onsite registration with secretariat, technical equipment and 24/7 phone help line. We are adaptable and extremely flexible as we are aware of the unique requirements and budget restrictions of each conference. Our services may be provided on an all-inclusive or on an a-la-carte basis.

## FLEXIBLE SOLUTIONS TO SUIT YOUR SPECIFIC NEEDS

Easy Conferences can provide organizers with a complete paper submission evaluation system at [www.easyacademia.org](http://www.easyacademia.org). We also have our own, custom-made one-stop-shop Conference Management System, [www.easyconferences.org](http://www.easyconferences.org), which offers the ability to sign up and within minutes register for the conference and its extra activities. Participants can book accommodation, airport transfers, social activities for themselves and their accompanied persons and pay instantly online.

Our extensive experience and personal attention to every participant's needs, backed by our team members' unrivaled expertise in their field, as well as the selection of the right partners, has resulted in our impeccable track record that is our guarantee for perfectly organizing any conference or event.

**Please visit our website, [www.easyconferences.eu](http://www.easyconferences.eu) for more information on our services, a list of upcoming and past events, as well as referrals from our customers.**

P.O.Box 24420, 1704,Nicosia, Cyprus  
 +357 22 591900  
 +357 22 591700  
[info@easyconferences.eu](mailto:info@easyconferences.eu)

[www.easyconferences.eu](http://www.easyconferences.eu)  
[www.easyconferences.org](http://www.easyconferences.org)



## Host Institution

## Sponsors



## Coordinator



aea-europe.net

e-mail: admin@aea-europe.net

Join our social media  
community



[https://twitter.com/AE Ae\\_2000](https://twitter.com/AE Ae_2000)



<https://www.facebook.com/Associationforeducationalassessmenteurope>



<https://www.instagram.com/aea-europe>



<https://www.linkedin.com/company/aea-europe>