

Parallel session 1 1

25 August 2026 11:30 - 13:00

KBG 228

Single Paper

Assessment and Evaluation, Learning and Instructional Technology

AI-Supported Assessment Systems: Design and Evaluation

Keywords: adaptive assessment, artificial intelligence, assessment tool, automatic feedback, feedback literacy, feedback tool, formative assessment, higher education, motivation, rubrics, self-assessment

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Anne Lohuis, Saxion University of Applied Sciences, Netherlands

Prototype Development of an AI Feedback System Trained on a Rubric and Exemplars

Keywords: artificial intelligence, feedback literacy, feedback tool, rubrics

Presenting Author: Takaho Iwata, Kwansei Gakuin University, Japan; **Co-Author:** Hayato Sezaki, The University of Tokyo, Japan; **Co-Author:** Bunich Otaki, Shizuoka University, Japan

Although interest in AI systems for assessment in higher education is growing, GenAI trained solely on rubrics tends to overestimate student performance. This study investigates whether incorporating exemplars improves evaluation accuracy, which refers to closeness to instructor ratings, and how students perceive such feedback. A university course with a presentation task is used as a case. In Study 1, two custom GPT-based models (GPT-5.1) were developed: one trained only on a rubric and another additionally trained on exemplars. They were used to evaluate 61 student presentation slides, compared with instructor ratings as pragmatic reference points. The results suggest providing exemplars improved agreement with instructor evaluations and reduced overestimation. In Study 2, an exploratory survey with five former students examined perceptions of AI-generated feedback. The findings suggest that exemplar-informed AI feedback was perceived as useful and trustworthy. Future research will refine the model by analyzing AI feedback comments and students' feedback literacy.

Evaluating a Learner-Driven, AI-Powered Mathematics Classroom Assessment Tool (Context-AI)

Keywords: adaptive assessment, artificial intelligence, assessment tool, motivation

Presenting Author: Burcu Arslan, Educational Testing Service, Germany; **Co-Author:** Jesse R. Sparks, Educational Testing Service, United States; **Co-Author:** Reginald M. Gooch, Educational Testing Service, United States; **Co-Author:** Ikkyu Choi, Educational Testing Service, United States; **Co-Author:** Matthew Bernacki, University of North Carolina at Chapel Hill, United States; **Co-Author:** Candace Walkington, Southern Methodist University, United States

K-12 mathematics classroom assessments often fail to engage all students because math items rarely reflect students' out-of-school interests or real-world experiences, thereby threatening the validity of interpretations of students' knowledge, skills, and abilities resulting from those assessments. To address this challenge, we developed Context-AI, a learner-driven, AI-powered formative mathematics classroom assessment tool. Context-AI personalizes the context of mathematics word problems in real time based on students' input about their out-of-school interests. It does so while preserving the mathematical structure of items created by subject-matter experts, which form the basis for the generated items (i.e., isomorphic items). We will present Context-AI's theory of change, qualitative insights from teacher co-design focus groups, and our initial evaluation of the quality of learner-driven, AI-generated personalized items in terms of item generation time, readability scores, and safety.

Examining the Effectiveness of a Comparative Feedback Dashboard

Keywords: automatic feedback, formative assessment, higher education, self-assessment

Presenting Author: Tine van Daal, University of Antwerp, Belgium; **Co-Author:** Michiel De Vriendt, KULeuven/imec research group itec, Belgium; **Co-Author:** Davy Mortier, University of Antwerp, Belgium; **Co-Author:** Marije Lesterhuis, UMC Utrecht, Switzerland; **Co-Author:** Wim Van Den Noortgate, KULeuven/imec research group itec, Belgium; **Co-Author:** Anaïs Tack, KULEUVEN/imec research group itec, Belgium; **Co-Author:** Renske Bouwer, Utrecht University, Netherlands

Comparative Judgement (CJ) is a reliable and valid method to assess writing, yet CJ scores lack transparency regarding assessment criteria. Natural Language Processing (NLP)-based feedback presents a promising opportunity but remains unexplored. Despite its lack of transparency, CJ also generates rich comparative feedback opportunities to support students' internal feedback (ranking of essays, benchmarks, access to peer work), but empirical evidence is lacking. This study addresses these gaps by investigating the effectiveness of a comparative feedback dashboard that integrated CJ-derived feedback with automated NLP-based feedback in an undergraduate scientific writing course. Feedback effectiveness was examined in terms of students' attention to (log data), appraisal of (self-report), and elaboration on feedback (think-aloud/interview data). Results indicate that students engaged with the dashboard and positively appraised its different components. The dashboard also supported students' elaboration on feedback. Overall, findings demonstrate the potential of combining human CJ with automated feedback to enhance students' internal feedback.

Parallel session 1 2

25 August 2026 11:30 - 13:00

KBG 224

Single Paper

Assessment and Evaluation, Higher Education

Feedback Literacy and Feedback Uptake

Keywords: feedback, feedback literacy, GenAI, higher education, measurement, quantitative methods, skills assessment

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Henning Fjortoft, Norwegian University of Science and Technology (NTNU), Norway

AI, Expert or Peer? Provider Biases and Feedback Uptake Behavior Among Pre-Service Teachers

Keywords: feedback, GenAI, higher education, quantitative methods

Presenting Author: Lucas Jasper Jacobsen, University of Hamburg, Germany; **Co-Author:** Ute Mertens, IPN - Leibniz Institute for Science and Mathematics Education, Germany; **Co-Author:** Thorben Jansen, Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Germany; **Co-Author:** Weber Kira, Hamburg University, Germany

The EU AI Act requires teachers to use high-risk AI safely, making it essential to assess AI-generated outputs. Feedback is one of the most important outputs, yet we know little about pre-service teachers' perceptions of AI-generated feedback. This study investigates how the perceived source of feedback (large language model (LLM), expert, peer) affects perception and uptake, and whether recognition accuracy and feedback quality moderate these effects. In a randomized experiment with 273 pre-service teachers, participants received written feedback on a mathematics learning goal, identified the source, rated five perception dimensions (fairness, usefulness, acceptance, willingness to improve, affect), and revised the goal accordingly. LLM feedback was rated highest in fairness and usefulness, resulting in the highest uptake (52%). Recognition accuracy moderated perceptions, with LLM feedback falsely ascribed to experts receiving more positive evaluations. Regression analysis showed that only feedback quality predicted uptake. We discuss addressing source-related biases in teacher education.

The role of mathematic skills and economic content knowledge for understanding economic graphs

Keywords: higher education, measurement, quantitative methods, skills assessment

Presenting Author: Petros Bengel, University of Tübingen, Germany; **Co-Author:** Malte Ring, University of Tübingen, Germany; **Co-Author:** Volker Schöer, University of the Witwatersrand, South Africa

Research repeatedly highlights the importance of math skills for students of economics. Specifically, graphs are used frequently in lectures to represent domain concepts. Given their importance as learning aid, it is vital to identify what skillset is required to handle economic graphs and how it is related to general math skills. To achieve this goal, we investigate whether economic graph skills and math graph skills can be meaningfully entangled. To answer this question, we draw from cross-sectional data from 1085 South African university students. By inspecting the dimensionality of items and grouping them into latent traits we find first hints that the skills are separable. In a second step, we will explore group differences to determine how the relative importance of mathematical versus economic content knowledge for graph comprehension varies across students. This will allow us to better understand and develop supporting interventions for learners who struggle with economic graphs.

A Canonical Correlation Analysis of Preservice Teachers' Teacher and Student Feedback Literacy

Keywords: feedback, feedback literacy, higher education, quantitative methods

Presenting Author:Büşra Zeynep Arslan, Gazi University, Türkiye; **Presenting Author:**Yesim Capa Aydin, Middle East Technical University, Türkiye

Feedback literacy is a significant component of preservice teachers' professional development, both as university students and as future educators. Few studies examine the relationship between preservice teachers' teacher feedback literacy and student feedback literacy. To address this gap, a canonical correlation analysis was performed with 699 preservice teachers at four state universities in Türkiye, using the Student Feedback Literacy (SFL) and Preservice Teacher Feedback Literacy (PTFL) scales (Cronbach's alpha: .76–.85 for SFL, .79–.89 for PTFL). The analysis identified three significant canonical functions, with only one accounting for 42.25% of the shared variance. For student feedback literacy (Set 1), all dimensions except readiness were strongly linked to the canonical variate, while all six dimensions of teacher feedback literacy (Set 2) were meaningful. Overall, the results suggest that higher student feedback literacy is associated with higher teacher feedback literacy.

Parallel session 1 3

25 August 2026 11:30 - 13:00

KBG 208

Single Paper

Assessment and Evaluation

Self-Assessment: Accuracy, Processes, and Learning

Keywords: academic achievement, artificial intelligence, grading, higher education, rubrics, secondary education, self-assessment, self-feedback

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Hayden Mueller, Technical University of Munich, Germany

Pessimistic, Optimistic, or Accurate?: Exploring Students' Explanations for Self-Assessment Accuracy

Keywords: academic achievement, grading, self-assessment, self-feedback

Presenting Author:Leire Pinedo, Universidad de Deusto, Spain; **Co-Author:**Javier Fernández, Universidad Autonoma de Madrid, Spain; **Co-Author:**Daniel García-Perez, Universidad Complutense de Madrid, Spain; **Co-Author:**Ernesto Panadero, Dublin City University, Ireland

This study examined the accuracy of higher education students' self-assessment and the reasons they attribute to discrepancies between their own scores and those of an instructor. A total of 112 psychology students wrote an academic essay, self-assessed and graded it, and were interviewed about any discrepancies. Results showed a weak but significant correlation between students' and instructor's scores, along with a general trend toward mild underestimation. Task performance significantly affected accuracy: higher achievers tended to underestimate, while lower achievers overestimated. Interview data revealed that under-estimators explained discrepancies through pessimistic or overly self-critical self-perceptions, whereas over-estimators relied on optimistic impressions and overconfidence to explain the gap with the instructor's score. Overall, inaccuracies were rarely attributed to differences in assessment criteria; instead, they primarily reflected students' internal perceptions and affective biases toward their own performance.

Self-Assessment Accuracy in AI-Assisted versus Traditional Contexts

Keywords: artificial intelligence, grading, rubrics, self-assessment

Presenting Author:Seval Kemal, Dublin City University, Ireland; **Co-Author:**Leire Pinedo, Universidad de Deusto, Spain; **Co-Author:**Darina Scully, Dublin City University, Ireland; **Co-Author:**Ernesto Panadero, Dublin City University, Ireland; **Co-Author:**Kevin Marshall, Microsoft Ireland, Ireland

Self-assessment is a critical component of self-regulated learning, yet students often struggle to judge their work accurately. With the increasing integration of artificial intelligence in education, questions arise about how AI-assisted feedback can enhance evaluative judgment beyond traditional rubric-based approaches. This study investigates whether AI-assisted self-assessment using Microsoft Copilot improves accuracy, and students' perceptions (confidence, trust, and perceived feedback generation ability), compared to rubric-only self-assessment. In a randomized controlled trial, higher education students self-assess a reflective essay using either a rubric alone or a rubric with AI assistance, and complete pre- and post-surveys. Data collection is ongoing, with a preliminary sample of 16 participants. We hypothesize that AI-assisted self-assessment will enhance accuracy and increase students' confidence, trust in their evaluative judgment, and perceived ability to generate feedback, particularly among lower-performing students. This research will contribute to understanding how educational AI tools can support or undermine evaluative judgment development.

How Do Students Learn to Self-Assess? Evidence from Secondary and Higher Education

Keywords: higher education, secondary education, self-assessment, self-feedback

Presenting Author:Beñat Etxeberria, University of Deusto, Spain; **Co-Author:**Leire Pinedo, Universidad de Deusto, Spain; **Co-Author:**Javier Fernández, Universidad Autonoma de Madrid, Spain; **Co-Author:**Carlos Rodríguez Hernández, New Jersey Institute of Technology, United States; **Co-Author:**Ernesto Panadero, Dublin City University, Ireland

Self-assessment is a key component of formative assessment, yet little is known about how students learn to self-assess across educational levels. This study explores how students report learning to self-assess, focusing on the sources they identify and the activities through which this learning occurs. Semi-structured interviews were conducted with 187 students from secondary and higher education. Results indicate that students most frequently report learning to self-assess through personal experience, while teachers and family members are mentioned less often. Personal experience was the dominant source across educational levels, particularly in higher education. Among the participants who reported having received instruction, secondary students mainly referred to learn through self-grading and revision before submission, whereas higher education students more often mentioned the use of self-assessment instruments. The findings suggest that self-assessment is often experienced as implicit rather than explicitly taught, highlighting the need to make instructional support for self-assessment more visible and intentional.

Parallel session 1 4

25 August 2026 11:30 - 13:00

MIN 009

Workshop Format

Assessment and Evaluation

Designing Education for Hybrid Intelligence Development

Keywords: assessment as learning, feedback literacy, GenAI, programmatic assessment

Interest group: SIG 01 - Assessment and Evaluation

Wise-Owl in action: Designing education for hybrid intelligence development

Keywords: assessment as learning, feedback literacy, GenAI, programmatic assessment

Presenting Author:Kelly Beekman, Fontys University of Applied Science, Netherlands; **Co-Author:**Lambert Schuwirth, NewMed, Australia

This hands-on workshop guides participants through designing education and assessment that develops and evaluates 'hybrid intelligence'; the productive integration of human and generative AI (GenAI) capabilities. Participants will explore a framework of six essential elements, that integrate with programmatic assessment(PA)-principles: communicating with AI, synergistically producing superior human–GenAI output, ethical and responsible use, critical thinking and information verification, association-rich knowledge, and agency-balance.

Participants engage in guided activities using 'Wise-Owl,' a GenAI-driven study companion designed to develop hybrid intelligence while supporting subject matter learning. Grounded in research on human learning, feedback literacy, assessment and evaluative judgment, Wise-Owl offers diverse modes and functions scaffolding active learning.

Through interactive demonstrations and design activities, attendees redesign existing teaching and assessment approaches, and leave with tailored AI-integrated approaches applicable to their institutional contexts. This workshop embodies "Opening up assessment" by illustrating how GenAI can openly scaffold learning, shifting assessment from one-time-evaluation to continuous, coached experiences.

Parallel session 1 5

25 August 2026 11:30 - 13:00

Pangea

Effectively Integrating AI and Peer Feedback: Theoretical Perspectives and Empirical Findings

Keywords: artificial intelligence, automatic feedback, feedback, feedback literacy, formative feedback, GenAI, peer feedback, quantitative methods, tertiary education

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Martin Greisel, University of Augsburg, Germany

Discussant: Omid Noroozi, Wageningen University and Research Centre, Netherlands

Feedback is central to learning, yet educators now must choose among feedback from teacher, peer, and generative-AI without clear guidance on when each source works best. Because these sources can be complementary, effective instructional design requires understanding not only their relative strengths but also how learners process them differently and how hybrid arrangements can be combined to maximize learning benefits. This symposium addresses two questions: (1) how feedback characteristics and reception processes differ across teacher, peer, and genAI sources, and (2) how peer and genAI feedback can be orchestrated for optimal impact. Across four studies, we present a social-psychological model distinguishing informational versus relational processing (Contribution 1), empirical comparisons of feedback features and uptake (Contribution 2), and evidence from hybrid designs examining sequencing (Contribution 3) and augmentation (Contribution 4), including whether AI substitutes for or enhances peer-feedback learning. Together, the contributions offer design-relevant insights for deploying complementary feedback sources.

What is the Difference between AI and Human Feedback: The Role of Investment and Judgment Appraisals

Presenting Author: Jennifer Meyer, University of Vienna, Austria; **Co-Author:** Johanna Fleckenstein, University of Hildesheim, Germany; **Co-Author:** Ernesto Panadero, Dublin City University, Ireland; **Co-Author:** Martin Daumiller, University of Freiburg, Germany

Generative AI and its possibilities for providing educational feedback even for complex performances are changing the educational landscape. The present work introduces a theoretical framework discussing source-specific differences in feedback appraisals. We theorize that human feedback sends a social signal that is absent in AI-generated feedback that triggers evolved attentional mechanisms based on the need to belong theory. We argue that these social signals lead to appraisals of investment and judgment, as well as increased personal relevance of the feedback. We call this mechanism the relational pathway. Building on this, we argue that human feedback is not always preferable and the relational pathway can have both favorable and detrimental consequences for feedback processing, depending on situational affordances and learner needs. We offer a conceptual lens on AI as a feedback provider, laying the groundwork for designing studies that clarify mixed findings and explore implications for hybrid human-AI feedback systems.

Peer, AI, and Teacher Feedback: Features, Perceptions and Uptake

Presenting Author: Renske de Kleijn, UMC Utrecht, Netherlands; **Presenting Author:** Renske Bouwer, Utrecht University, Netherlands; **Co-Author:** Golnoush Haddadian, Georgia State University, United States; **Co-Author:** Antoine van den Beemt, Technical University Eindhoven, Netherlands; **Co-Author:** Omid Noroozi, Wageningen University and Research Centre, Netherlands; **Co-Author:** Saba Soleimani, Interdisciplinary transformation University (IT:U), Austria

As prior research has mainly examined peer, teacher, and GenAI feedback in isolation this study compares these three feedback sources in terms of feedback features, student perceptions, and student feedback uptake. The findings show that students blindly preferred the comments from teachers most and that AI feedback was perceived least favorable but affectively supportive. In terms of the actual comments, AI feedback had the least overall. In terms of uptake a substantial number of comments was ignored, but the same holds for elaborative uptake. All in all, the feedback sources differed in volume and in functional emphasis. In conclusion, teacher feedback remains central for metacognitive guidance and self-regulation. These findings suggest that combining feedback sources may support meaningful engagement, learner agency, and sustained writing development rather than relying on a single source.

Combining AI and peer feedback: How different feedback processes interact and complement each other

Presenting Author: Rasmus R. Hansen, Aarhus University, Denmark; **Co-Author:** Ida Bang Hansen, Aarhus University, Denmark; **Co-Author:** Christopher Neil Pillop, Aarhus University, Denmark

With the emergence of GenAI-based chatbots, sought-after personalized feedback has become widely available. Studies have shown differences between AI and human feedback, but few studies have investigated combined effects of these feedback sources. This study presents a case study where undergraduate physics students (N = 20) refined their oral presentations through a feedback process. Students were divided into two groups: one received AI feedback followed by peer feedback—the other group had the order reversed. Through AI chatlogs, surveys, and five stimulated recall interviews, this study seeks to understand how the feedback sources differ, interact, and complement each other. A reflexive thematic analysis lets interviews act as a guide for exploration of chatlogs, and survey data. Results highlight how different feedback sources can help students focus on different aspects of their work, such as structure, correctness, and oral presentation, by differences in modality, interactivity, and students' presuppositions of the feedback providers.

Designing AI to Augment (Not Replace) Peer Feedback: A Comparative Study of Four Feedback Conditions

Presenting Author: Martin Greisel, University of Augsburg, Germany; **Co-Author:** Julia Hornstein, University of Augsburg, Germany; **Co-Author:** Alina Sheppard-Bujtor, University of Augsburg, Germany; **Co-Author:** Ingo Kollar, University of Augsburg, Germany

Providing peer-feedback promotes deep processing, yet students often struggle to produce high-quality comments. Generative AI can generate additional feedback, but may reduce learning by offloading cognitive work. This study examined how AI can augment rather than replace peer-feedback by comparing two strategies: a "best-of" approach (human and AI generate feedback; reviewers select the best parts) and a "feedback-on-feedback" approach (AI evaluates and helps refine peer-feedback). These were contrasted with human-only and AI-only feedback. In two waves, N = 354 pre-service teachers completed a five-phase learning scenario over five weeks. AI augmentation increased the feedback length and the number of suggestions, but did not significantly impact how adequate the receiver perceived the feedback, or the subsequent task performance (preliminary analyses). Effects on feedback features will be presented at the conference. Further research should identify effective human-AI combinations that improve both feedback quality and learning outcomes.

Parallel session 2 1

25 August 2026 14:00 - 15:30

KBG 208

Collaborative Space

Assessment and Evaluation

Student Agency in Programmatic Assessment

Keywords: assessment for learning, feedback, higher education, programmatic assessment

Interest group: SIG 01 - Assessment and Evaluation

Unlocking Student Agency in Programmatic Assessment

Keywords: assessment for learning, feedback, higher education, programmatic assessment

Presenting Author: Kitty Meijer, HU University of Applied Sciences Utrecht; Open University of the Netherlands, Netherlands; **Presenting Author:** Bas Agricola, HU University of Applied Sciences, Netherlands; **Co-Author:** Simone Buys, HU University of Applied Sciences Utrecht, Netherlands; **Co-Author:** Maaike Koopman, HU University of Applied Sciences Utrecht, Netherlands; **Co-Author:** Liesbeth Baartman, Maastricht University, Netherlands; **Co-Author:** Desirée Joosten-ten Brinke, Maastricht University, Netherlands

Programmatic Assessment (PA) is designed to empower students, shifting the focus from summative decision-making to continuous feedback and development. However, despite these design intentions, students often struggle to exert agency within assessment contexts dominated by hierarchical power dynamics. In this session, we introduce a new research project investigating the interplay between PA affordances (opportunities for action) and students' will and ability to exert agency. Participants will engage in a collaborative exploration of two critical challenges: (1) identifying which affordances in PA truly invite agentic behaviour, and (2) addressing the methodological difficulties of measuring agency in assessment contexts. This collaborative space session is relevant for researchers and practitioners interested in the intersection of assessment design, student autonomy, and methodology.

Parallel session 2 2

25 August 2026 14:00 - 15:30

MIN 009

Workshop Format

Teaching and Teacher Education

Teacher Training Evaluation

Keywords: assessment tool, competence assessment, evaluation, formative assessment

Interest group: SIG 01 - Assessment and Evaluation

Transforming Teacher Training Evaluation with the Digital Training Companion App

Keywords: assessment tool, competence assessment, evaluation, formative assessment

Presenting Author: Maria Pannatier, Swiss Federal Institute of Technology, EPFL Lausanne, Switzerland; **Presenting Author:** Johann Groll, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland; **Co-Author:** Sunny Avry, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

This workshop adopts a visible, learning focused, process-oriented approach to examining the development of teachers' digital competence. Addressing the ongoing discrepancy between evidence-based recommendations and current practice, it shifts away from short-term satisfaction metrics towards flexible, context-responsive evaluation. Based on Guskey's framework, the Digital Competence Training Assessment Model (DCTAM) is presented using the Digital Training Companion (DTC), a web-based application that facilitates the design of continuous assessments, feedback loops and sustainability indicators, supporting trainers, teachers, students and decision-makers. The DTC integrates formative, participatory and AI-enhanced evaluation practices that strengthen agency and accountability across professional learning ecosystems. Based on findings from a mixed-methods pilot (2024–2025) and a professional community of teachers (2025–2026), the workshop demonstrates how open, technology-supported assessment can enhance teachers' competence, motivation and classroom application, while providing participants with hands-on experience of using the tool and encouraging reflection on the wider implications of assessment innovation practices.

Parallel session 2 3

25 August 2026 14:00 - 15:30

Pangea

Single Paper

Assessment and Evaluation

Assessment in Transition

Keywords: assessment, assessment design, assessment policy, automatic feedback, feedback tool, GenAI, grading, quantitative methods, secondary education

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Xin Liu, Deakin University, Australia

Assessment in Mathematics and Permitted Tools in a Historical and Future Perspective

Keywords: assessment design, assessment policy, GenAI, secondary education

Presenting Author: Tina Thysgaard, University of Southern Denmark, Denmark; **Co-Author:** Ane Qvortrup, University of Southern Denmark, Denmark

This paper examines how changes in permitted tools have historically reshaped assessment in Danish upper secondary mathematics and considers implications for emerging technologies. Drawing on conceptual history (Koselleck, 2007) and a mathematics-as-discourse perspective (Sfard, 2009), the study employs an abductive approach (Schwartz-Shea & Yanow, 2011) to analyse ministerial orders, reform documents, and other contemporaneous sources tracing the introduction and legitimization of the pocket calculator. Findings indicate that the introduction was accompanied by a shifting emphasis toward mathematical applications, modelling and problem-solving, and an increase in extra-mathematical contexts in exam questions. Drawing parallels to Large Language Models (LLMs), the paper discusses how their integration may similarly reconfigure assessment, requiring students to demonstrate proficiency through orchestrating generative AI tools. The findings highlight how permitted tools signal broader conceptual shifts and invite reflection on future assessment frameworks in mathematics education.

Gradeless Teaching in Upper Secondary Education: Teachers' Experiences of Emerging Challenges

Keywords: assessment, assessment policy, grading, secondary education

Presenting Author: Mathias Hanson, Kristianstad University, Sweden; **Co-Author:** Anders Jönsson, Kristianstad University, Sweden; **Co-Author:** Henning Fjortoft, Norwegian University of Science and Technology (NTNU), Norway

Classroom grading has been shown to influence students in multiple ways, including effects on motivation, learning orientation, and wellbeing. In response, initiatives to reduce or remove grades from assessment feedback have emerged, often referred to as gradeless or grade-reduced approaches. However, such initiatives may also alter other aspects of the teaching–learning process, creating new challenges. This study investigates how teachers experience and respond to these challenges in a Norwegian upper secondary school that implements a gradeless introductory period. During this period, student work is not graded and feedback is exclusively formative. Drawing on semi-structured interviews with six teachers and one school leader, the study explores teachers' perceptions of challenges related to student motivation, students' understanding of performance level, and their ability to plan academic development. The findings highlight the persistent influence of the grading system and tensions between reducing grading in practice and meeting policy requirements to assign final grades.

How Learners Use Digital Support: Identifying Feedback-Use Profiles

Keywords: automatic feedback, feedback tool, quantitative methods, secondary education

Presenting Author: Larissa Altenburger, Leuphana Universität Lüneburg, Germany; **Co-Author:** Michael Besser, Leuphana University Lüneburg, Germany

Digital learning platforms offer various forms of automated feedback, yet learners differ in how they use and perceive this feedback. This study examined learners' feedback use using a person-centered approach. Fifty-five German seventh graders worked on fraction tasks on a digital learning platform providing multiple feedback options (e.g., explanations, complete solutions). Learners' interactions with feedback were recorded and analysed in terms of frequency and duration. Descriptive analyses revealed considerable variability in feedback use across learners. Latent profile analyses identified three distinct feedback-use profiles: help seekers, solution-oriented users, and minimal help seekers. Comparisons of feedback perception across profiles showed that help seekers perceived the feedback more positively than the other groups, with statistically significant differences in perceived competence. The findings highlight substantial heterogeneity in learners' use of digital feedback and underline the importance of supporting learners in actively using feedback. Given the small sample size, results should be interpreted as exploratory.

Parallel session 3 1

25 August 2026 16:00 - 17:30

Pangea

Single Paper

Assessment and Evaluation, Learning and Instructional Technology

Rethinking Assessment in the Age of GenAI

Keywords: assessment design, feedback, feedback literacy, GenAI, higher education, online assessment, programmatic assessment, quantitative methods, self-feedback, transparency

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: David Zamorano, Universidad de Deusto, Spain

AI Literacy of Teachers: Prompt Engineering and Model Selection as Predictors of AI-Feedback Quality

Keywords: feedback, feedback literacy, GenAI, quantitative methods

Presenting Author: Lucas Jasper Jacobsen, University of Hamburg, Germany; **Co-Author:** Julia Pargmann, Universität Hamburg, Germany; **Co-Author:** Weber Kira, Hamburg University, Germany

Feedback is central to promoting pre-service teachers' professional skills. Large language models (LLMs) such as ChatGPT are increasingly used to provide feedback. It is unclear which factors predict feedback quality. Two quasi-experimental studies investigated which prompt characteristics improve LLM feedback and whether prompt quality or model choice has the greater influence. Pre-service teachers formulated learning goals, for which various LLMs generated feedback. In Study 1 (N = 240 feedbacks), different prompts were tested with several LLMs. The best combinations were used again in Study 2 (N = 345 feedbacks) to test their predictive power for feedback quality. Hierarchical regressions show that both the choice of LLM and prompt quality are significant predictors. Just a few targeted principles of prompt engineering are sufficient to generate high-quality feedback. Domain-specific language is particularly effective. The selection of suitable LLMs and prompting skills should therefore be fostered as relevant characteristics of instrumental AI literacy.

Thinking to Learn: Navigating AI Use and Integrity in Higher Education Assessment

Keywords: assessment design, GenAI, higher education, transparency

Presenting Author: Anne-Marie Chase, Swinburne University of Technology, Australia; **Co-Author:** Kelly Galvin, Swinburne University of Technology, Australia;

Co-Author: Sandra Luxton, Swinburne University of Technology, Australia

Generative artificial intelligence (GenAI) is rapidly transforming assessment and evaluation in higher education, introducing new pedagogical and ethical challenges. This topic is critical as educators seek transparent, adaptable, and responsive approaches to assure learning and integrity in diverse, technology-rich contexts. Despite growing interest, there remains a gap in practical frameworks guiding principled decisions about GenAI integration in assessment. This study asks: How can assessment design be opened up to support ethical, context-sensitive GenAI use while preserving cognitive ownership and validity? Drawing on qualitative research with unit convenors and conceptual analysis, we introduce a developmental assessment framework that aligns learning outcomes, constructs, and observable indicators with ethical AI use. Key findings reveal enhanced educator confidence, clearer communication of AI boundaries, and practical recommendations for academic integrity. These insights contribute a transparent, adaptable model for reimagining assessment in the age of AI, supporting both individual learning and collective educational goals.

Developing hybrid intelligence through agency-balanced assessment

Keywords: assessment design, GenAI, higher education, programmatic assessment

Presenting Author: Kelly Beekman, Fontys University of Applied Science, Netherlands; **Co-Author:** Lambert Schuwirth, NewMed, Australia

This paper introduces hybrid intelligence as a foundational educational concept for the generative AI (GenAI) era and proposes an assessment-framework centred on agency-balance between human-AI contributions. Rather than policing AI-use, we advocate stimulating students' capacity for optimal human-AI cognitive-task distribution. While challenging in traditional high-stakes testing-contexts, this approach proves feasible within programmatic assessment (PA) or other longitudinal frameworks. Drawing on PA-theory and human-AI complementarity research, we argue students must demonstrate AI-tool proficiency and metacognitive capacity for optimal task distribution. We identify six essential elements of hybrid intelligence: (1) effective AI communication, (2) superior joint human-AI outputs, (3) ethical AI-use, (4) critical data-literacy, (5) association-rich knowledge networks, and (6) agency balance. We propose assessment-strategies including process-portfolios documenting human-AI role allocation, reflective agency annotations, and comparative demonstrations against AI-only baselines. By explicitly assessing how students engage with or abstain from AI, we reframe assessment to acknowledge distributed intelligence.

Disciplinary differences in using GenAI for self-feedback in asynchronous online higher education

Keywords: feedback, GenAI, online assessment, self-feedback

Presenting Author: Montserrat Martínez Melo, Fundació per a la Universitat Oberta de Catalunya (FUOC), Spain; **Co-Author:** Anna Espasa, Universitat Oberta de Catalunya, Spain; **Co-Author:** Rosa M Mayordomo, Open University of Catalonia (UOC), Spain; **Co-Author:** Teresa Guasch, Open University of Catalonia (UOC), Spain; **Co-Author:** Jose Luna, Universitat Oberta de Catalunya, Spain

Generative Artificial Intelligence (GenAI) is increasingly positioned as a feedback agent in higher education, yet its role as an external reference (comparator) that triggers students' self-feedback remains underexplored, specially in fully asynchronous online programmes. This study examines whether the use of GenAI as self-feedback comparator varies by academic level and discipline. Results, based on an online survey addressed to university students ($n=2,044$), show that 15.9% reported using GenAI as a comparator for self-feedback at any point in a recent assessed assignment. Use did not differ by academic level (undergraduate 15.0% vs master's 18.0%). However, discipline was strongly associated with the use of GenAI for self-feedback ($\chi^2(5, N=1,979)=79.33, p$

Parallel session 3 2

25 August 2026 16:00 - 17:30

KBG 228

Single Paper

Assessment and Evaluation

Fairness and Inclusion in Educational Assessment

Keywords: assessment for learning, assessment policy, assessment quality, digital assessment, fairness, formative assessment, inclusive assessment, large-scale assessment, neurodiversity, qualitative methods, secondary education

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Burcu Arslan, Educational Testing Service, Germany

A Framework for Evaluating Large-Scale Assessments: Language Inclusivity in TIMSS, PISA, and NAEP

Keywords: fairness, inclusive assessment, large-scale assessment, qualitative methods

Presenting Author: Martina Heinle, Technical University of Munich, Centre for International Student Assessment, Germany; **Presenting Author:** Ainhoa González Uriarte, Technical University of Munich, Centre of International Student Assessment, Germany; **Co-Author:** Lisa Teufele, Technical University Munich, Centre for International Student Assessment, Germany; **Co-Author:** Samuel Greiff, Technical University of Munich, Centre for International Student Assessment, Germany

As the number of students with a migration background and the widening competence gap between immigrant and non-immigrant students grows, it is essential that large-scale assessments generate valid data on immigrant and, consequently, mostly multilingual students to identify performance trends and to develop targeted support measures. However, given the wide variation in testing modes and participation regulations across large-scale assessments, the validity of the resulting data remains questionable. Addressing this discrepancy and research gap, we developed a framework to assess the language inclusivity of large-scale assessments. Technical and study reports, as well as background questionnaires of TIMSS, PISA, and NAEP, are analysed to determine language inclusivity. Our results indicate considerable variation, which only partly aligns with established recommendations from language acquisition research regarding the inclusion and assessment of multilingual students. This inconsistency across assessments raises concerns about both the validity of the data produced and the overall fairness of tests.

Personalisation of adaptive environments for fair and inclusive assessment and learning

Keywords: digital assessment, formative assessment, inclusive assessment, neurodiversity

Presenting Author: Maarten van der Velde, MemoryLab, Netherlands; **Co-Author:** Thomas Wilschut, Rijksuniversiteit Groningen, Netherlands; **Co-Author:** Hedderik van Rijn, Rijksuniversiteit Groningen, Netherlands; **Co-Author:** Burcu Arslan, Educational Testing Service, Germany

A key promise of adaptive digital (formative) assessment and learning environments is that they automatically adjust difficulty and pace to students' levels, providing appropriate challenge or support. Although such adaptivity is generally beneficial, specific system design choices can adversely affect some students, leading to systematic misinterpretation of their abilities and threatening construct validity. For instance, using response speed to assess mastery may disadvantage students with reading or typing difficulties when the construct is unrelated, as slower responses may erroneously be attributed to insufficient mastery.

To achieve accurate assessment of students' abilities, such construct-irrelevant variance should be mitigated. Classroom studies indicate that allowing students with spelling difficulties to speak rather than type answers and removing question-level response time limits for students with attentional differences can reduce the influence of irrelevant factors on performance. These examples show that personalising the assessment environment to individual needs can improve fairness, inclusivity, and validity.

How formative assessment is used to inform teacher actions in the classroom

Keywords: assessment for learning, formative assessment, qualitative methods, secondary education

Presenting Author: Janneke Van Der Steen, HAN University of Applied Sciences, Netherlands; **Co-Author:** Tamara Schilt-Mol, Hogeschool van Arnhem en Nijmegen, Netherlands; **Co-Author:** Cees van der Vleuten, Maastricht University, Netherlands; **Co-Author:** Desirée Joosten-ten Brinke, Maastricht University, Netherlands

Over the years, formative assessment has gained interest in educational systems around the world. Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to inform decisions about what next steps in teaching and learning will best suit students' learning needs. However, how teachers come from this collection of information about student learning to a well-informed formative decision about the follow-up usually stays implicit. This multiple case study focused on bringing the invisible processes that are involved in formative decision-making to the surface. Through journal-writing and interviews with teachers from secondary education different pathways were found in which formative decision-making occurs. Furthermore, the outcomes of the current study reveal that formative assessment is not an isolated strategy but always embedded in teachers' existing knowledge and beliefs about the context, content, learning and learners.

The controversy of anonymous assessment in Oslo: Sacrificing learning for perceived fairness

Keywords: assessment policy, assessment quality, fairness, formative assessment

Presenting Author:Gøril Brataas, Department of Scandinavian Languages and Literature, Faculty of education, Norway; **Co-Author:**Ane Aarre, Department of Teacher Education and School Research, University of Oslo, Norway; **Co-Author:**Jo Helge Schei, Department of Teacher Education and School Research, University of Oslo, Norway; **Co-Author:**Øystein Gilje, Department of Teacher Education and School Research, University of Oslo, Norway

Assessment quality is a cornerstone of the educational system. However, it is often unclear how aspects of assessment quality, such as validity, reliability, and fairness, should be understood, operationalized, and balanced against one another. Reporting from a case study of an anonymous assessment initiative in upper secondary schools in Oslo — motivated by concerns about fairness, reliability, and formative assessment— we examine how teachers and students who participated in the initiative perceive assessment quality. Reflexive, collective thematic analysis of semi-structured interview with 30 participants indicates that anonymous assessment is a controversy that alters the conditions under which assessment judgements are formed and used and raises questions about whether learning is being sacrificed for perceived fairness. We argue that attempts to strengthen assessment quality cannot be evaluated solely in terms of consistency; they must also consider implications for formative use and professional judgment.

Parallel session 3 3

25 August 2026 16:00 - 17:30

KBG 224

Single Paper

Assessment and Evaluation, Higher Education

Feedback and Assessment Practices

Keywords: assessment design, assessment quality, feedback, feedback literacy, formative assessment, higher education, peer assessment, programmatic assessment, quantitative methods, self-assessment, sustainable assessment, tertiary education

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Hayden Mueller, Technical University of Munich, Germany

Holistic curriculum design for evaluative judgement: a developmental-ecological framework

Keywords: assessment design, programmatic assessment, self-assessment, sustainable assessment

Presenting Author:Susie Macfarlane, Deakin University, Australia; **Co-Author:**David Boud, Deakin University/University of Technology Sydney, Australia; **Co-Author:**Joanna Tai, Deakin University, Australia

Developing evaluative judgement supports students' capacity to interpret and apply standards and feedback, and graduates to identify quality where it is not explicit (Tai et al., 2018), yet there is limited research examining how educators design curricula to cultivate it. Semi-structured interviews with 20 educators at an Australian university were analysed using the Framework method (Gale et al., 2013). Educators described strategies ranging from individual tasks to integrated, programme-wide curriculum strategies. Informed by Self-Determination and Social-Cognitive theories alongside epistemic and ontological perspectives, the analysis identified four developmental trajectories - affective, disciplinary, epistemic and ontological - and the curriculum mechanisms and learning conditions educators designed to support them. The study generated a developmental-ecological framework, conceptualising evaluative judgement as emerging through integrated, student-centred curriculum ecologies. Future research should move beyond evaluating single-interventions to examine how evaluative judgement emerges across whole-of-programme curriculum ecologies that attend to students' motivational, epistemic and ontological development.

What feedback literacy behaviours do learners do? A global survey of current and former students

Keywords: feedback, feedback literacy, quantitative methods, tertiary education

Presenting Author:Laura Hughes, Deakin University, Australia; **Presenting Author:**David Boud, Deakin University/University of Technology Sydney, Australia;

Presenting Author:Naomi Winstone, University of Surrey, United Kingdom; **Co-Author:**Jaclyn Broadbent, Deakin University, Australia; **Co-Author:**Phillip Dawson, Deakin University, Australia; **Co-Author:**Elizabeth Molloy, The University of Melbourne, Australia; **Co-Author:**Joanna Tai, Deakin University, Australia

Student feedback literacy (FL) has gained prominence as a key ingredient in effective feedback. Existing research has been limited to quantitative methods focusing on students' beliefs and attitudes, or small-scale qualitative intervention studies. How students behave when engaging with feedback remains underexplored. The Feedback Literacy Behaviour Scale (FLBS) is a recently validated questionnaire that can be used to investigate FL behaviours on a broader scale. Current ($n = 230$) and former ($n = 161$) university students from around the world completed the FLBS (English version). Preliminary results indicated that former students who had completed a degree scored higher on seeking, making sense of, and providing feedback information. Using feedback information and managing affect did not significantly differ between groups. Longitudinal research is needed to understand how FL behaviours may change over time. Exploring the connection between achievement, motivation, and FL behaviours may also be fruitful to support student FL.

Teachers' Feedback Practices and Perceptions: Evidence from a National Survey

Keywords: assessment quality, feedback, feedback literacy, formative assessment

Presenting Author:Adrian Parames Gonzalez, University of Deusto, Spain; **Co-Author:**Ernesto Panadero, Dublin City University, Ireland; **Co-Author:**Juan Fernández, Universidad Autónoma de Madrid, Spain

Feedback is widely recognized as a high-impact educational intervention, yet its effectiveness depends on how it is designed, delivered, and used. This study examines teachers' feedback practices and perceptions through a nationwide survey grounded in current feedback research. Data were collected via a voluntary sample of teachers from different educational stages in Spain, with 492 responses available at this stage. A self-report questionnaire comprising 14 items was developed, structured around three key topics: (1) frequency, timing, and opportunities for revision; (2) type of feedback and modes of delivery; and (3) teachers' perceptions and perceived barriers to effective feedback use. The instrument enables a multidimensional analysis of feedback practices and addresses a notable gap in national survey research, particularly in compulsory education. Preliminary results report overall response frequencies, while final analyses stratified by educational level, teaching experience, and gender, will be presented at the conference.

Student Involvement in Assessment Practices in Spanish Universities: A Longitudinal Study

Keywords: assessment design, higher education, peer assessment, self-assessment

Presenting Author:Beñat Etxeberria, University of Deusto, Spain; **Co-Author:**Alaitz Amezcua-Urrutia, Universidad de Deusto, Spain; **Co-Author:**Javier Fernández, Universidad Autónoma de Madrid, Spain; **Co-Author:**Carlos Rodríguez-Hernández, New Jersey Institute of Technology, United States; **Co-Author:**Juan Fraile, Universidad Francisco de Vitoria, Spain; **Co-Author:**Ernesto Panadero, Dublin City University, Ireland

Research on assessment in higher education has increasingly emphasised the importance of student involvement to support learning and transparency. However, evidence suggests that such involvement remains limited in practice. This study provides a nationwide longitudinal overview of student involvement in assessment in Spanish universities by comparing syllabus data from 2015/2016 and 2022/2023. Drawing on a sample of 1,347 syllabi from all the public universities, the study examines how self-assessment and peer-assessment are implemented. Using dichotomous coding and chi-square analyses, results show that both self- and peer-assessment remain marginal practices at both time points. Self-assessment and self-grading are moderately associated but infrequent, with lower prevalence in 2022 than in 2016, while peer-assessment and peer-grading are extremely scarce and decline over time. Despite contextual changes in higher education, assessment practices appear structurally stable, with student involvement occupying a peripheral role rather than being embedded in core assessment designs.

Parallel session 3 4

25 August 2026 16:00 - 17:30

MIN 009

Workshop Format

Assessment and Evaluation

Lego Serious Play in Assessment

Keywords: assessment design, authentic assessment, higher education, inclusive assessment

Interest group: SIG 01 - Assessment and Evaluation

Building Ideas with Lego Serious Play: History Degrees and Final-Year Assessment

Keywords: assessment design, authentic assessment, higher education, inclusive assessment

Presenting Author:Katie Carpenter, University of Leeds, United Kingdom; **Presenting Author:**Carly O'Neill-Barrett, University of Leeds, United Kingdom; **Presenting Author:**Sarah Briggs, University of Leeds, United Kingdom

A project at the University of Leeds explored using LEGO Serious Play (LSP) to design a new 'Public History Project' module, to offer students an alternative to the traditional written dissertation for History undergraduates. LSP is an innovative, inclusive method that facilitates effective reflection and the efficient organisation and sharing of ideas (James and Nerantzi, 2019), making it a suitable method for this type of exploration and co-creation with staff and students. LSP workshops allowed students to express their hopes and concerns about final-year projects through building physical models. It revealed strong emotional responses, both positive and negative, about independent research. The new module aims to increase student choice and engagement by allowing diverse project formats. The research also highlights an additional need: for extra support to address emotional challenges. The findings informed the module design so that the assessment considered students' emotional experiences to foster positive outcomes.

Parallel session 3 5

25 August 2026 16:00 - 17:30

KBG 208

Collaborative Space

Assessment and Evaluation

Feedback in Programmatic Assessment

Keywords: assessment for learning, feedback, feedback literacy, programmatic assessment

Interest group: SIG 01 - Assessment and Evaluation

The Dual Purpose of Feedback in Programmatic Assessment: Growth and Judgement – Tension or Synergy?

Keywords: assessment for learning, feedback, feedback literacy, programmatic assessment

Presenting Author:Judith Gulikers, Wageningen University, Netherlands; **Presenting Author:**Liesbeth Baartman, University of Applied Sciences Utrecht, Netherlands; **Co-Author:**Renske de Kleijn, UMC Utrecht, Netherlands; **Co-Author:**Kelly Beekman, Fontys University of Applied Science, Netherlands; **Co-Author:**Tamara Schilt-Mol, Hogeschool van Arnhem en Nijmegen, Netherlands

Programmatic assessment (PA) aims to optimize learning by embedding assessment for learning within a coherent system that also supports robust high-stakes decision-making. Feedback plays a central role in this ambition, yet it is increasingly clear that feedback in PA serves a dual purpose: supporting student learning and informing summative decisions. While theoretically complementary, these purposes may be experienced as contradictory by learners. Research suggests that when students perceive assessment information as contributing to grades or progression decisions, their engagement with feedback for learning is diminished. This collaborative space invites participants to critically explore what feedback looks like when it serves learning versus decision-making, why fostering the learning function of feedback in PA remains challenging, and whether dual-purpose feedback is feasible in practice without undermining student growth. We intend to identify the inhibiting mechanisms and work towards design principles for feedback in the context of PA.

Parallel session 4 1

26 August 2026 11:30 - 13:00

KBG 208

Single Paper

Assessment and Evaluation, Teaching and Teacher Education

Assessment Design and Implementation

Keywords: assessment design, assessment literacy, assessment quality, authentic assessment, fairness, high school, mixed methods, performance assessment, post-graduate education, programme assessment, secondary education, standardized assessment

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: ZEYNEP DULGER, Marmara University Institute of Educational Sciences, Türkiye

Two years teaching for open-book exams: impact on teaching, learning, wellbeing and performance

Keywords: assessment design, authentic assessment, high school, mixed methods

Presenting Author:Vanessa Scherman, International Baccalaureate, Netherlands; **Co-Author:**Rebecca Chivers, International Baccalaureate, Netherlands; **Co-Author:**Rebecca Hamer, International Baccalaureate, Netherlands

Open book exams (OBEs) are expected to encourage a greater focus on higher order thinking in assessment performance. However, academic literature comprises mostly small-scale studies presenting mixed results and little evidence relevant to implementation in an international multi-lingual high school context. In 2022, an international awarding body recruited 332 schools worldwide to participate in a longitudinal mixed method study involving students, teachers, school management and examiners. Combining data from over 20,000 surveys, 62 focus groups across the four response groups and the grades of 11,000 students[RH1], this study analyses the impact of one- and two-years' preparation for three different types of OBE on school and teaching practices, student learning, epistemological development, wellbeing, test anxiety, exam preparation and performance compared to closed book examinations. Results highlight differences in impact on school and teaching practices by OBE type, as well as the experienced impact variability on student learning, development and performance.

Assessment in Initial Teacher Education: A Syllabi Analysis of In-Person and Online Programmes

Keywords: assessment design, assessment literacy, post-graduate education, programme assessment

Presenting Author:Adrian Parames Gonzalez, University of Deusto, Spain; **Co-Author:**Juan Fraile, University of Francisco de Vitoria, Spain; **Co-Author:**Javier Fernández, Universidad Autonoma de Madrid, Spain; **Co-Author:**Carlos Rodríguez, Tecnológico de Monterrey, Mexico; **Co-Author:**Alaitz Amezuza, University of Deusto, Spain; **Co-Author:**Ernesto Panadero, Dublin City University, Ireland

Assessment practices experienced in initial teacher education shape how the future teachers experience assessment and influence the beliefs and approaches they later apply in their professional practice. This study analyses how assessment practices are implemented in a national sample of 366 syllabi from the Master's in Teacher Education, comparing face-to-face and online programmes. Using document analysis, we examined the assessment practices employed, their weighting in final grades, predominant assessment profiles, and the degree of student involvement. Results showed that online programmes rely predominantly on final written examinations and individual assignments, whereas in-person programmes use a broader and more balanced set of assessment practices. Latent profiles reveal a highly homogeneous assessment model centred on assignments and final exams, with minimal opportunities for student participation. The findings highlight the need to promote assessment literacy adapted to programme modalities and to align assessment practices with student-centred and participatory approaches in teacher education.

Developing a Standardized National Writing Test: Ethical, Conceptual, and Technical Considerations.

Keywords: assessment design, fairness, mixed methods, standardized assessment

Presenting Author:Rosario Escribano, Pontificia Universidad Católica de Chile, Chile; **Presenting Author:**Natalia Avila Reyes, Pontificia Universidad Católica de Chile, Chile; **Co-Author:**Diego Carrasco, Pontificia Universidad Católica de Chile, Chile; **Co-Author:**Carolina Castillo, Pontificia Universidad Católica de Chile, Chile; **Co-Author:**Javiera Figueroa Miralles, Universidad Alberto Hurtado, Chile; **Co-Author:**María Jesús Espinosa, Universidad Diego Portales, Chile

We report on a two-phase funded research program aimed at designing a national-scale writing assessment adhering to psychometric standards while addressing ethical and consequential concerns. Framing writing as a social practice, a research-based sixth-grade assessment was developed using authentic tasks, communicatively focused rubrics, and prior-knowledge scaffolding to reduce unwanted variance and support intended score uses. Design was guided by principles of validity, reliability, and fairness, ensuring that no group is advantaged or disadvantaged by construct-irrelevant factors. Evidence from multiple experimental and quasi-experimental studies conducted during prototype development indicates that task choice was not systematically biased across demographic groups, scaffolding prior knowledge supported performance, pedagogical materials aligned with intended score uses, and rubric dimensions showed no DIF. Final analyses showed that most score variance was attributable to student ability and that scores were invariant across socioeconomic groups. Ongoing work focuses on strengthening measurement precision at higher ability levels while maintaining fairness.

Bridging Coherence and Disciplinary Integrity: Designing Cross-Curricular Performance Standards

Keywords: assessment design, assessment quality, performance assessment, secondary education

Presenting Author:Sofia Kesidou, The University of Sydney, Australia

Educational systems have invested substantially in content standards articulating what students should know, yet performance standards articulating how well students demonstrate learning remain underdeveloped. Existing approaches offer either generic frameworks lacking disciplinary specificity, or subject-specific standards that develop in isolation and fragment curriculum. This design-based research, conducted within New South Wales's K-10 curriculum reform, proposes a tiered framework operating across disciplinary boundaries while respecting domain-specific epistemologies. The framework synthesises assessment theory traditions (Sadler; Tognolini; Masters) with research on intellectual standards, conceptual understanding, epistemic cognition, transfer, and disciplinary literacy. Across two iterative cycles, 52 primary and secondary teachers identified convergent quality dimensions and developed subject-specific translations, revealing both affordances and epistemological tensions in cross-curricular design. The resulting six-dimension framework, through its translation mechanism, provides common language for judging quality of student performance across the curriculum, supporting diagnostic formative assessment while preserving disciplinary integrity.

Parallel session 4 2

26 August 2026 11:30 - 13:00

Pangea

Single Paper

Assessment and Evaluation

Feedback in Digital Learning Environments

Keywords: assessment for learning, automatic feedback, digital assessment, formative feedback, GenAI, higher education, motivation, performance assessment, qualitative methods

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Carly O'Neill-Barrett, University of Leeds, United Kingdom

Erroneous AI-Generated Feedback Engages Learners but Harms Intrinsic Motivation

Keywords: automatic feedback, formative feedback, GenAI, motivation

Presenting Author: Jennifer Meyer, University of Vienna, Austria; **Co-Author:** Marlene Steinbach, IPN - Leibniz Institute for Science and Mathematics Education, Germany; **Co-Author:** Zachary Pardos, University of California, Berkeley, United States

Artificial intelligence-based large language models (LLMs) can automatically provide elaborated feedback, but may generate erroneous ("hallucinated") content whose motivational consequences are poorly understood. In a mathematics tutoring system, this randomized controlled study investigates the effects of LLM-generated feedback containing errors on engagement with feedback and task-related motivation and self-concept. $N = 249$ learners received either all accurate feedback, feedback with hallucinations in 50% or 100% of the instances, or no feedback. Higher hallucination rates increased time spent on feedback but reduced intrinsic motivation. No effects were observed on self-concept. Prior knowledge did not moderate the effects. Highlighting the importance of considering both time on feedback as indicator of momentary engagement and self-reported task motivation, the findings provide initial evidence for the potential motivational risks associated with LLM-generated feedback. Further research is needed to shed more light on the underlying mechanisms that caused the contrasting patterns of results observed.

Assessing skills on screen: Reliability of video-based psychomotor nursing skill assessment

Keywords: digital assessment, formative feedback, higher education, performance assessment

Presenting Author: Greet Leysens, KU LEUVEN, Belgium; **Co-Author:** Wim Van Petegem, KU LEUVEN, Belgium; **Co-Author:** Nathalie Charlier, KU LEUVEN, Belgium

Psychomotor skills are essential in health professional education. However, traditional assessment via direct observation is resource-intensive and stressful. Asynchronous video-based assessment offers a scalable, learner-centred alternative, but its reliability remains underexplored in nursing and midwifery education. In a multi-year research project (2022–2026), we investigate whether assessment of recorded performances provides reliable assessments of novice students' psychomotor skills. First-year students performed three nursing skills after a standardised demonstration. Performances were recorded from multiple camera angles and rated asynchronously by teachers and peers using criterion-referenced checklists. Inter- and intra-rater reliability was analysed across cohorts. Findings inform assessment practices in nursing and midwifery but is also relevant to other domains requiring mastery of psychomotor skills. Future work should explore video-annotation tools and AI-driven feedback to support actionable student feedback and enhance a more time-efficient assessment process.

Learner autonomy in computer-based learning?! A systematic literature review on self-chosen feedback

Keywords: assessment for learning, automatic feedback, digital assessment, formative feedback

Presenting Author: Friederike von Kienitz, Leuphana University of Lüneburg, Germany; **Co-Author:** Fynn Töllner, Leuphana University of Lüneburg, Germany; **Co-Author:** Maike Hagen, Leibniz University Hannover, Germany; **Co-Author:** Michael Besser, Leuphana University of Lüneburg, Germany

Feedback is considered to be one of the most important factors influencing learning and has therefore been subject to extensive research. This research has mainly focused on feedback that is automatically assigned to learners. However, learners often do not use the feedback they receive. Therefore, according to self-determination theory giving learners a choice over their feedback may be a promising approach to foster engagement by giving a feeling of autonomy. Based on these considerations, this review aims at summarizing the current state of research on self-chosen feedback. Our extensive literature search identified 30 articles. The results indicate a strong focus on cognitive rather than affective-motivational variables in past research. However, learners do choose (different) feedback types when given an option. Limitations and implications of the results will be discussed.

Student's experience of feedback personalisation in online higher education: a qualitative analysis

Keywords: assessment for learning, formative feedback, higher education, qualitative methods

Presenting Author: Montserrat Martínez Melo, Fundació per a la Universitat Oberta de Catalunya (FUOC), Spain; **Co-Author:** Anna Espasa, Universitat Oberta de Catalunya, Spain; **Co-Author:** José Luna Luján, Fundació per a la Universitat Oberta de Catalunya (FUOC), Spain; **Co-Author:** Teresa Guasch, Open University of Catalonia (UOC), Spain; **Co-Author:** Rosa M Mayordomo, Open University of Catalonia (UOC), Spain

Feedback personalisation strategies to promote effective dialogic feedback are a key challenge in online higher education. The lack of feedback personalisation (adjustment to students' learning needs for feedback engagement) constrains the possibilities for students to give meaning and value to feedback. The paper shares results of a qualitative research, based on 15 focus interviews with students. In order to gain an in-depth understanding of students' experiences, thematic and content analysis were implemented. Results show a dominant association of feedback personalisation and concrete and written references about individual assignments. Suggestive content is recognized as personalised if providing actionable guidance to their individual texts. Generic or typified feedback perception triggers suspicions, undermine trust and reduces engagement. Students' expression of personalisation as the adaptation to the learners' cognition and motivational needs, suggest a greater cognitive and emotional regulation, greater sense of teacher presence and is linked with the verbalisation of commitment and effort.

Parallel session 4 3

26 August 2026 11:30 - 13:00

MIN 011

Poster Presentation

Assessment and Evaluation, Higher Education

Technology-Enhanced Assessment and Student Engagement, Motivation, and Learning

Keywords: artificial intelligence, assessment, assessment design, assessment quality, assessment tool, competence assessment, digital assessment, evaluation, feedback, formative assessment, higher education, learning analytics, motivation, qualitative methods

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Ligia Tomazin, University of Trento (I), Italy

Understanding Teacher Development in Formative Assessment Through Qualitative Data Triangulation

Keywords: competence assessment, formative assessment, higher education, qualitative methods

Presenting Author: Nienke Boere, HU University of Applied Sciences Utrecht/ Open University, Netherlands; **Co-Author:** Liesbeth Baartman, University of Applied Sciences Utrecht, Netherlands; **Co-Author:** Lisette Munneke, Utrecht University of Applied Sciences, Netherlands; **Co-Author:** Elly de Buijn, Hogeschool Utrecht / OU, Netherlands

This collaborative session explores how multiple qualitative data sources can be triangulated to capture teachers' professional learning over time. Using a qualitative multiple case study in higher professional education, the session focuses on tracing changes in teachers' vocational knowledge of formative assessment practices aimed at fostering students' epistemic agency. This study combines interviews, reflective journals, artefacts, fieldnotes, and group reflections, with each source revealing different aspects of teachers' development. Through small group discussions and collective synthesis, participants address key methodological challenges, including handling contradictions across data sources, constructing nonlinear developmental trajectories, ensuring analytic transparency when working with large and diverse datasets, and considering participants' roles in the triangulation process.

Formative assessment as a predictor of student motivation and metacognition: A multilevel analysis

Keywords: assessment, feedback, formative assessment, motivation

Presenting Author:Rajabu Shafii, Department of Educational Sciences, Switzerland; **Co-Author:**Jean-Louis Berger, Department of Education Sciences, Switzerland

Although formative assessment (FA) fosters learners' academic motivation and metacognition, its effects on students' motivational beliefs and metacognition in a context of large class size, resource-shortage and dominance of traditional assessment practice remain underexplored. Drawing from FA, Expectancy-Value-Cost (EVC), and metacognition theories, this study examined whether students perceived FA practices predict students' motivational beliefs and metacognition at the individual and classroom levels. Data were collected from 1543 students nested in 66 classes from Tanzanian primary schools. A multilevel model revealed that at the individual level, higher perception of FA was associated with higher perceived expectancy, value, cost, and metacognition. At the classroom level, FA was also positively associated with all outcomes, with strong effects at the classroom level. This study suggests that establishing FA as a collective classroom norm, rather than a judgmental tool, exerts a significant positive influence on student motivation and metacognition at both levels.

Technology-Enhanced Assessment: A Systematic Review of Affective-Motivational Effects

Keywords: assessment design, digital assessment, evaluation, motivation

Presenting Author:Arya-Ronahi Dogan, Europa-Universität Flensburg / IPN Kiel, Germany; **Co-Author:**Ecenaz Alemdag, Europa-Universität Flensburg, Germany; **Co-Author:**Marlit Annalena Lindner, Leibniz Institute for Science and Mathematics Education (IPN), Germany

Technology-enhanced assessments (TEAs) gain increasing attention in education. We define TEAs as computer-based assessments that go beyond conventional performance tests with features such as adaptivity, feedback, personalization, immersion, or gamification. Although experimental studies examine the influence of TEA design factors on affective-motivational outcomes, existing reviews and meta-analyses address only specific and limited aspects and mainly focus on cognitive outcomes, while affective-motivational outcomes have hardly been systematically analyzed. Our meta-analysis aims to provide an overview of theoretical frameworks, implemented design factors, targeted affective-motivational outcomes, key moderators, and the emerging role of AI in TEAs. Our systematic literature search yielded 30,390 records for abstract screening, of which 136 papers were included in the full-text screening, which is ongoing. Overall, this study aims to show how TEAs are implemented and how they impact affective-motivational outcomes as a key target of the technology-enhanced factors implemented. Results will be presented at the conference.

Learning Analytics for Conversation-Based Assessment: Structuring Evidence Beyond Performance

Keywords: artificial intelligence, digital assessment, formative assessment, learning analytics

Presenting Author:Adrian Klabunde, Europa-Universität Flensburg / IPN Kiel, Germany; **Co-Author:**Lauritz Schewior, Leibniz Institute for Science and Mathematics Education (IPN), Germany; **Co-Author:**Marlit Annalena Lindner, Europa-Universität Flensburg / IPN Kiel, Germany

Advances in large language models have made conversation-based assessment (CBA) a scalable approach to formative assessment, enabling the integration of diagnosis and instructional support within natural-language dialogue. This study investigates how interaction data from CBA can be transformed into meaningful learning analytics indicators. In a school-based experiment with $N = 230$ secondary school students, an LLM-powered CBA on the carbon cycle was compared with a traditional multiple-choice assessment. We propose a structured framework that combines (a) performance measures reflecting adaptive tutor support, (b) content-related indicators, (c) lexical process indicators derived from student responses, and (d) sentiment-based measures. Preliminary results suggest that composite CBA performance scores are comparable to traditional test scores, while additional indicators provide further insights into student learning. Ongoing analyses examine the validity and diagnostic value of these indicators for formative assessment and instructional decision-making.

The Use of Multimedia in Multiple-Choice Tests: Effects on Examinee Engagement and Performance

Keywords: assessment quality, assessment tool, digital assessment, motivation

Presenting Author:Militsa Ivanova, University of Cyprus, Cyprus; **Co-Author:**Michalis Michaelides, University of Cyprus, Cyprus

While innovative multiple choice (MC) items incorporating multimedia features are increasingly used to enhance assessment quality, their effects on engagement and validity remain unclear. This study examines how multimedia features in MC items influence examinee engagement and performance in computer-based cognitive assessments. The study experimentally compares innovative items incorporating animation, graphics, audio, or video with paired traditional text-based MC items that include text or static images only. Each student completes all item pairs in a within-subject counterbalanced design to control for item and order effects. Examinee engagement is assessed through a multimodal approach combining response time indicators, self-reported effort, and eye-tracking measures of attention allocation. Linear mixed-effects modeling is used to analyze performance and engagement while accounting for participant- and item-level variability. The study aims to advance understanding of how multimedia item formats shape test-taking behavior and inform the design of valid and engaging computer-based assessments.

Parallel session 4 4

26 August 2026 11:30 - 13:00

MIN 009

Poster Presentation

Assessment and Evaluation

Programmatic Assessment and Competency Development

Keywords: academic achievement, assessment as learning, assessment design, competence assessment, feedback literacy, higher education, performance assessment, programmatic assessment, reflection, secondary education, skills assessment

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Charlotte Van Tricht, University of Antwerp, Belgium

Opening Up Assessment-as-Learning by Co-Creation: Assessment Tools for Global Citizenship Skills

Keywords: assessment as learning, higher education, reflection, skills assessment

Presenting Author:Laurie Delnoij, Maastricht University SBE, Ts 53, Netherlands; **Co-Author:**Ainsley Loudoun, Maastricht University, Netherlands; **Co-Author:**Inken Gast, Maastricht University, Netherlands

This poster presents the development and initial implementation of two assessment-as-learning components in a new asynchronous Global Citizenship skills module for international business undergraduates on their study abroad: (1) a recurring self-assessment and (2) a reflective rubric co-created with students. Drawing on experiential and reflective learning theories, these tools support students' awareness, goal-setting, and self-regulated skill development during their study abroad experience. We describe the co-creation process and report survey feedback on the reflective rubric ($N=91$), highlighting high perceived clarity and usefulness alongside a need for clearer differentiation and guidance between performance levels. Student engagement with the self-assessment tool ($N=216$) was high initially and showed selective, repeated use for skills perceived as most relevant or challenging, indicating learner autonomy and opportunities for targeted scaffolding. The poster will invite discussion on how interactive, student-informed assessment tools can open up more learner-centred, developmental, and scalable assessment practices in large, diverse cohorts.

Stimulating reflection in programmatic assessment: zooming out and zooming in

Keywords: feedback literacy, higher education, programmatic assessment, reflection

Presenting Author:Niels van der Baan, Maastricht University, Netherlands; **Co-Author:**Liesbeth Baartman, University of Applied Sciences Utrecht, Netherlands; **Co-Author:**Lianne Loosveld, Maastricht University, FHML, Dept. of Educational Research and Development, Netherlands; **Co-Author:**Jill Whittingham, Maastricht University, FHML, Dept. of Educational Research and Development, Netherlands

Programmatic assessment requires students to ask for feedback on their data points, reflect on this feedback, zoom out and conceptualise this feedback to their overall level of competence. Subsequently, they have to zoom back in to concretise and apply this feedback to subsequent assignments. This requires advanced reflective skills from the student. A coach can support students in this reflective process. Although coaching is frequently used in higher education to promote student reflection and competence development, supporting this process of zooming in and out based on feedback students receive within programmatic assessment places additional demands on coaches. The aim of this project is twofold. First, we will explore students' needs in this process of zooming out and -

in. Second, these insights can be translated into faculty development initiatives for coaches, such as written guidelines and workshops.

Assessing Teamwork Competencies in Upper Secondary and Higher Education: A Systematic Review

Keywords: assessment as learning, competence assessment, higher education, secondary education

Presenting Author:Lynn Van den Broeck, KU Leuven (BE), Belgium; **Co-Author:**Rani Dujardin, KU Leuven (BE), Belgium; **Co-Author:**Katrien Vangrieken, KU Leuven (BE), Belgium; **Co-Author:**Hanne Deprez, KU Leuven (BE), Belgium; **Co-Author:**Jeroen Cloots, KU Leuven (BE), Belgium

Teamwork competencies (TWC) are essential for academic success, employability, and lifelong learning, yet they do not develop automatically through group work. This systematic review examines how TWC are assessed in upper secondary and higher education and whether assessment can serve as an intervention to enhance these skills. Following PRISMA guidelines, a search in Scopus and ERIC identified 2,448 records; from the first 50 full-text papers screened, 15 met inclusion criteria. Most included studies lacked clear definitions and theoretical grounding, conceptualizing teamwork through behaviours such as communication and leadership. Assessment methods relied heavily on self-report surveys and observations, with instruments like CATME and CATS commonly used. Few approaches integrated feedback or process-focused interventions. Current practices remain fragmented, limiting opportunities to leverage assessment for competency development. Future research should adopt theoretically grounded frameworks and embed feedback mechanisms to transform assessment into an active learning intervention.

The student's perspective on generic skills endorsement in higher education

Keywords: academic achievement, competence assessment, higher education, performance assessment

Presenting Author:Sanne ter Meulen, UMC Utrecht, Netherlands; **Co-Author:**Helena Pennings, HAN University of Applied Sciences, Netherlands

Generic skills are increasingly emphasised in higher education, particularly within educational formats such as challenge-based learning and programmatic assessment. However, assessment of generic skills remains complex and context-dependent, with pass/fail often used as a summative outcome. Such grading provides limited opportunities for students to display their development. In an elective minor at the Graduate School of Life Sciences, Utrecht University, an individual endorsement has been introduced alongside the pass decision to formally recognise students' personal and professional achievements. This longitudinal, exploratory, mixed-methods study investigates students' perceptions of the added value and usefulness of such an endorsement. Twelve students complete a questionnaire with validated items at three time points between July 2025 and February 2026. In addition, two semi-structured interviews per student are conducted to obtain deeper understanding of students' perceptions and use of the endorsement. The study aims to contribute to discussions on meaningful assessment of generic skills.

Revised Principles of Programmatic Assessment

Keywords: assessment as learning, assessment design, higher education, programmatic assessment

Presenting Author:Tamara Schilt-Mol, Hogeschool van Arnhem en Nijmegen, Netherlands; **Co-Author:**Liesbeth Baartman, University of Applied Sciences Utrecht, Netherlands; **Co-Author:**Lubberta de Jong, Universiteit Utrecht, Netherlands

Programmatic assessment integrates information from multiple low stakes assessments to support learning and robust decisions about competence. In 2021, a consensus framework for programmatic assessment in health sciences education was published. Increasing implementation across various educational domains has revealed differences in interpretation, contextual challenges, and conceptual ambiguities. This study aimed to critically evaluate and refine the theoretical core of programmatic assessment based on practical experience and expert input. Using a stepwise design, data were collected through a questionnaire, expert meetings, and round table discussions. The findings informed a revised set of seven principles that aims to capture the core of programmatic assessment, enhance conceptual clarity, and guide contextualized design choices. The poster invites discussion about the principles and what constitutes the theoretical foundation of programmatic assessment.

Parallel session 5 1

26 August 2026 14:00 - 15:30

Pangea

Symposium

Assessment and Evaluation

Reimagining Technology-Enabled Assessment and Feedback Processes

Keywords: artificial intelligence, assessment for learning, automatic feedback, feedback, feedback tool, formative assessment, formative feedback, GenAI, online measures, self-feedback, tertiary education

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Kim Schildkamp, University of Twente, Netherlands

Discussant: Kelly Beekman, Fontys University of Applied Science, Netherlands

This symposium brings together papers that provide innovative insights into how assessment and feedback processes can be reimagined using technology. Collectively, the papers examine how digital tools and data-rich environments, including the use of generative artificial intelligence for feedback provision, open up new and largely underexploited opportunities to enhance assessment and feedback practices. Presentations explore how technology can support more timely, transparent, and dialogic feedback; strengthen learner agency and self-regulation; and enable more responsive teaching practices. Drawing on empirical research and design-based approaches across diverse educational contexts, the symposium highlights both the possibilities and tensions involved in rethinking assessment and feedback in technology-enabled learning environments, with implications for practice, policy, and future research.

Technology-Supported Insights into Student Feedback Use and Self-Regulated Learning

Presenting Author:Therese Hopfenbeck, University of Melbourne, Australia; **Co-Author:**Fabienne van der Kleij, Australian Council for Educational Research (ACER), Australia; **Co-Author:**Anastasiya Lipnevich, The Graduate Center of the City University of New York, United States; **Co-Author:**Siv Gamlem, Volda University College, Faculty of Humanities and Teacher Education, Norway, Norway

Personalised feedback is widely recognised as vital for improving student learning. Advances in educational technology, including artificial intelligence, create new opportunities for timely and personalised feedback. However, the impact of feedback on learning is highly variable, shaped by how students perceive, interpret, and act on feedback within self-regulated learning (SRL) processes. This paper presents a conceptual model that explicitly links feedback use and SRL, highlighting the reciprocal relationship between external feedback and students' affective, cognitive, and behavioural responses. The paper also proposes methodological approaches to examine student engagement with feedback in technology-enhanced learning environments. We advocate a shift from retrospective self-report measures to technology-enabled micro-level methods that capture in-the-moment feedback use and SRL processes. Specifically, we introduce a micro-analytic Feedback Self-Regulated Learning (F-SRL) protocol to examine feedback engagement before, during, and after assessment tasks. This approach provides a foundation for personalising and optimising feedback to support learning at scale.

H-AI-FIVE: High-quality AI-Feedback on Instructional explanation Videos for Education

Presenting Author:Ilona Friso-van den Bos, University of Twente, Netherlands; **Co-Author:**Andreas Breiter, University of Bremen, Germany; **Co-Author:**Marieke van Geel, University of Twente, Netherlands; **Co-Author:**Jan Küster, University of Bremen, Germany; **Co-Author:**Sara van der Linden, University of Twente, Netherlands; **Co-Author:**Kim Schildkamp, University of Twente, Netherlands; **Co-Author:**Karsten Wolf, University of Bremen, Germany

Practicing with giving explanations (in explanation videos) is an effective approach for student-teachers to improve their teaching skills (Mukuka & Alex, 2024). However, providing feedback on these videos is resource-intensive. LLM-generated feedback could be a way to support student-teachers and teacher-educators. Therefore, we developed a prototype of a tool where student-teachers can upload their videos and receive AI-generated feedback, based on a multimodal data analysis of the quality of explanation videos. In this pilot, we started with four criteria that are important for the quality of instruction: prior knowledge, cognitive activation, visualization and signaling. In our presentation, we will describe how we used a combination of an educational design research (McKenney & Reeves, 2019) approach with co-creation (Molina León et al. 2022) to develop this tool, as well as present the first evaluation results based on a focus group with teacher-educators and a walkthrough with student-teachers.

Comparing teacher comments to AI feedback: What students prefer and what is better for them

Presenting Author:Ligia Tomazin, University of Trento (I), Italy; **Co-Author:**Anastasiya Lipnevich, The Graduate Center of the City University of New York, United States; **Co-Author:**Luofan Shu, The Graduate Center, CUNY, United States; **Co-Author:**Peter Johnson, The Graduate Center, CUNY, United States; **Co-Author:**Carolina Lopera-Oquendo, The Graduate Center of the City University of New York, Colombia

Providing feedback on student writing is essential for effective instruction but remains challenging due to the time-intensive nature of personalized comments. Advances in artificial intelligence (AI) offer scalable alternatives, yet little is known about how students perceive AI feedback or how it compares with teacher feedback. Across two experimental studies with middle and high school students in Brazil and undergraduate students in the United States, we compared three feedback conditions: teacher comments, AI feedback (ChatGPT), and simulated teacher feedback (AI-generated feedback presented as teacher comments).

Students revised their essays after receiving feedback, and outcomes included writing performance, feedback appraisal, emotional reactions, and engagement. In both studies, writing performance improved across all conditions, with no significant differences by feedback source. However, students generally appraised teacher feedback more positively and, particularly in Study 2, reported greater boredom and skepticism toward AI feedback. These findings support the use of AI alongside teacher feedback.

Parallel session 5 2

26 August 2026 14:00 - 15:30

KBG 224

Single Paper

Assessment and Evaluation

Adaptive and AI-Based Assessment

Keywords: adaptive assessment, artificial intelligence, assessment for learning, automatic feedback, digital assessment, formative assessment, GenAI, measurement validation

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Kitty Meijer, Netherlands

(Large) Language Models for initial item parameter estimation

Keywords: adaptive assessment, artificial intelligence, assessment for learning, GenAI

Presenting Author: Jeroen Donkers, Maastricht University, Netherlands; **Co-Author:** Johannes Steinrücke, University of Twente, Netherlands

Computerized Adaptive Testing (CAT) requires precise item calibration, typically achieved through Item Response Theory (IRT) models such as the Rasch model. However, pre-testing often reveals discrepancies: some items fail model fit requirements or significantly deviate from their intended difficulty. This renders item generation inefficient, as expert time is wasted on finetuning items that are ultimately rejected. In this pilot study, we investigated the potential of text embedding models to predict item parameters prior to empirical pre-testing. Embeddings for 7,065 Rasch calibrated items (in Dutch) were generated using the nomic-embed-text-v2-moe model. Ridge Regression and XGBoost classifiers were trained after categorizing items into three, four or five difficulty categories. Performances (i.e. accuracy) only matched random chance, indicating that embeddings alone do not capture sufficient psychometric information. Consequently, (1) structured linguistic and cognitive-load features of items and (2) language models using raw item text will be assessed for parameter prediction potential.

The Role of Assessment in Adaptivity: A Systematic Meta-Review

Keywords: adaptive assessment, assessment for learning, digital assessment, formative assessment

Presenting Author: Marie Godow, IPN Kiel / Europa-Universität Flensburg, Germany; **Co-Author:** Andreas Lachner, University of Tübingen, Germany; **Co-Author:** Marlit Annalena Lindner, IPN - Leibniz Institute for Science and Mathematics Education, Germany

Adaptivity in education offers a promising approach to addressing classroom heterogeneity by continuously adjusting learning content, instruction, timing, and environment to individual student needs. However, research on adaptivity is characterized by diverse terminologies, such as Adaptive Teaching, Differentiated Instruction, or Personalized Learning, limiting conceptual clarity. This meta-review compares key conceptualizations of adaptivity in education, with a particular focus on assessment practices. Based on a systematic review of 72 theoretical articles, reviews, and meta-analyses, we identified two overarching perspectives: teacher-centered and system-/learner-centered approaches. While both perspectives target similar learner characteristics and adaptive outcomes, they differ markedly in assessment implementation and in the roles assigned to teachers and technology. Teacher-centered approaches rely on informal, teacher-driven assessment, whereas system-/learner-centered approaches use formal, data-driven assessment. The findings highlight the need to integrate teacher-centered traditions with technological innovations to support meaningful assessment practices that enable adaptivity in educational settings.

Engagement in AI-Based Adaptive Dialogs: Process-Oriented Evidence from Science Classroom Settings

Keywords: adaptive assessment, artificial intelligence, assessment for learning, measurement validation

Presenting Author: Marlen Holtmann, IPN Kiel / Europa-Universität Flensburg, Germany; **Co-Author:** Marlit Annalena Lindner, IPN Kiel / Europa-Universität Flensburg, Germany; **Co-Author:** Libby Gerard, University of California, Berkeley, United States; **Co-Author:** Marcia Linn, University of California-Berkeley, United States

AI-based adaptive dialogs embedded in web-based instruction offer opportunities to study behavioral and cognitive engagement during educational assessment. We explore how students engage with AI-adaptive dialogs using detailed process data from a genetics unit implemented in authentic secondary school classrooms. The AI-adaptive dialogs, designed to assess student scientific understanding, are administered in the middle and at the end of the unit. Indicators of engagement include time-on-task and response length, while disengagement indicators include rapid responding, extremely short answers, and off-task behavior. Results show that AI-based adaptive dialogs sustain engagement across timepoints, have low rates of disengagement, and show an increase in knowledge integration. Reduced time-on-task in the second dialog reflects task familiarity and increased efficiency.

Conversation-Based Assessment: Student Perceptions and Efficiency Trade-Offs

Keywords: artificial intelligence, assessment for learning, automatic feedback, formative assessment

Presenting Author: Marlit Annalena Lindner, IPN - Leibniz Institute for Science and Mathematics Education, Germany; **Co-Author:** Lauritz Schewior, Leibniz Institute for Science and Mathematics Education (IPN), Germany; **Co-Author:** Guher Gorgun, IPN Kiel & Europa-Universität Flensburg, Germany

Conversation-based assessment (CBA) is a promising formative assessment approach that captures students' reasoning through interactive dialogue and individualized feedback. Enabled by large language models, CBA can now be implemented at scale, yet empirical evidence on how students experience such assessments remains limited. This preregistered experiment compared student perceptions of CBA with traditional multiple-choice assessment (MCA) in a formative context. After a learning phase on the carbon cycle, 220 adult learners completed both assessment formats in counterbalanced order. Results showed that while CBA required substantially more time than MCA, students perceived it as more motivating, interactive, adaptive, and effective for learning. MCA was consistently rated as more efficient. Notably, two-thirds of participants preferred CBA for learning despite lower efficiency. Qualitative comments highlighted the perceived support and relaxed testing atmosphere in CBA. The findings indicate a trade-off between engagement and efficiency and suggest that CBA holds potential, but requires substantial research.

Parallel session 5 3

26 August 2026 14:00 - 15:30

MIN 011

Demo Session

Assessment and Evaluation

Design Principles for Formative Assessment

Keywords: assessment design, assessment for learning, formative assessment, transparency

Interest group: SIG 01 - Assessment and Evaluation

Design principles to help design a formative assessment plan

Keywords: assessment design, assessment for learning, formative assessment, transparency

Presenting Author: Janneke Van Der Steen, HAN University of Applied Sciences, Netherlands; **Co-Author:** Tamara Schilt-Mol, Hogeschool van Arnhem en Nijmegen, Netherlands; **Co-Author:** Cees van der Vleuten, Maastricht University, Netherlands; **Co-Author:** Desirée Joosten-ten Brinke, Maastricht University, Netherlands

AbstractHow can you design formative assessment consciously and coherently for multiple lessons? In this demo session the design principles that help to design an aligned, coherent and complete formative assessment plan will be presented and demonstrated to equip participants to design their own formative assessment plan. Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to inform decisions about what next steps in teaching and learning to take. Formative assessment is considered as an effective intervention to improve education and student learning, however does not always meet these expectations when formative assessment is not consciously and coherently planned. These design principles for formative assessment plans are the result of multiple cycles of educational design research with the purpose of helping teachers to tie together different formative activities in an effective way.

Parallel session 5 4

26 August 2026 14:00 - 15:30

MIN 009

Workshop Format

Assessment and Evaluation

What works in Programmatic Assessment

Keywords: assessment design, higher education, programmatic assessment, student evaluations of teaching

Interest group: SIG 01 - Assessment and Evaluation

What works in programmatic assessment, and why?

Keywords: assessment design, higher education, programmatic assessment, student evaluations of teaching

Presenting Author: Andrea Oudkerk Pool, Fontys University of Applied Science, Netherlands; **Presenting Author:** Jeroen van der Linden, HAN University of Applied Sciences, Netherlands; **Co-Author:** Marlies de Vos, HU University of Applied Sciences, Netherlands; **Co-Author:** Tamara van Schilt-Mol, HAN University of Applied Sciences, Netherlands; **Co-Author:** Kelly Beekman, Fontys University of Applied Science, Netherlands; **Co-Author:** Liesbeth Baartman, University of Applied Sciences Utrecht, Netherlands

In this workshop, we present the results from the first study of the research project "*Prototypes: what works in programmatic assessment, and why?*," in which we investigate mechanisms underlying the success or failure of programmatic assessment in higher education. Drawing on a multiple case study of six programmes across three Universities of Applied Sciences in the Netherlands, we focus on three interrelated factors: teachers' assessment beliefs, the development of a shared view on learning, and alignment between the intended, enacted, and experienced curriculum.

In the workshop, we first present the theoretical foundations and empirical findings. Participants will then engage in structured, small-group discussions with programme leaders to critically examine the role of these factors in practice. By explicitly adopting a reflexive and critical stance, we aim to foster collective meaning-making and scholarly dialogue around conceptual assumptions, empirical findings, and implications for assessment design.

Parallel session 6 1

26 August 2026 16:00 - 17:30

KBG 224

Single Paper

Assessment and Evaluation

Rethinking Learning and Assessment

Keywords: artificial intelligence, assessment, assessment design, equity, grading, high school, higher education, qualitative methods, workplace assessment

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Bas Agricola, HU University of Applied Sciences, Netherlands

How judging generative artificial intelligence in assessment shapes understanding of knowledge

Keywords: artificial intelligence, assessment, higher education, qualitative methods

Presenting Author: Margaret Bearman, Deakin University, Australia; **Co-Author:** Jack Walton, Centre for Research in Assessment and Digital Learning, Deakin University, Australia; **Co-Author:** Nicole Crawford, Deakin University, Australia; **Co-Author:** Joanna Tai, Deakin University, Australia; **Co-Author:** David Boud, Deakin University/University of Technology Sydney, Australia

Many university students work with generative artificial intelligence (GenAI) when completing assessment tasks. This involves making judgements about GenAI and its outputs, which may influence deeper conceptions of knowledge. Such processes could have unintended effects on learning, leading to the question: How does working with GenAI for assessment shape students' understandings of knowledge?" Twenty-six students were interviewed about their university study alongside tracing or demonstrating their relevant GenAI use. Data were analysed for judgement events, a sociocultural holistic conception of judgement. Twenty-seven events concerned making judgements about knowledge as part of assessment. Findings suggest that students made judgements about knowledge as *either* 'what is' or 'how to' information. Working with GenAI appears to reinforce the idea that these forms of knowledge are separate. Consideration can be given to how assessment tasks and formats can address this artificial binary in student understanding.

What counts as learning? Exploring working students' invisible educational experiences

Keywords: assessment, equity, higher education, workplace assessment

Presenting Author: Mollie Dollinger, Curtin University, Australia; **Co-Author:** Franziska Lessky, University of Innsbruck, Austria; **Co-Author:** Nicole Crawford, Deakin University, Australia

Universities often position themselves as the proprietors of students' employability, offering experiences from work-integrated learning placements, career development learning, and industry-informed curriculum. Yet increasingly across the globe, students are undertaking employment alongside their studies or coming to university with previous rich working experiences. This work-learning, however, goes largely unrecognised by institutional assessment and credentialing systems. In this study, therefore, we draw on Quinn's (2023) recent conceptualisation of 'invisible education' to explore the patterns of equity students from an Australian university as they learn through their formal studies and employment experiences. Our findings reveal three patterns: invisible learning as parallel, invisible learning as affective micro-moments, and invisible learning as pit stops. Furthermore, they highlight the current limitations of university practices to incorporate these valuable experiences. We argue that opening up assessment and evaluation to new possibilities demands rethinking not just how we evaluate, but what we consider worthy of evaluation.

Going gradeless in upper secondary school. A multiple case study of a natural experiment

Keywords: assessment, grading, high school, qualitative methods

Presenting Author: Henning Fjortoft, Norwegian University of Science and Technology (NTNU), Norway; **Co-Author:** Lise Vikan Sandvik, Norwegian University of Science and Technology (NTNU), Norway; **Co-Author:** Iveta Kohanová, Norwegian University of Science and Technology (NTNU), Norway; **Co-Author:** Sverre Tveit, University of Agder, Norway; **Co-Author:** Dan-Anders Normann, Norwegian University of Science and Technology (NTNU), Norway; **Co-Author:** Arne Kristian Karlsen Amdal, Norwegian University of Science and Technology (NTNU), Norway; **Co-Author:** Stine Angvik, Norwegian University of Science and Technology (NTNU), Norway; **Co-Author:** Andrea Seljeseth, Norwegian University of Science and Technology (NTNU), Norway; **Co-Author:** Kari Smith, Norwegian University of Science and Technology, Norway

Grading is often associated with negative impacts on teacher workloads and student learning and well-being. Internationally, teachers are "going gradeless" (GG) to reduce stress or improve learning. Little is currently known about the rationales behind, implementation of, consequences of, and contextual aspects influencing GG. This multiple case study provides empirical evidence on the implementation of GG in seven Norwegian upper secondary schools. Data from interviews with teachers, school leaders, students, and parents were analyzed, yielding five results: (1) feedback became important, (2) teacher assessment literacy was developed, (3) while some students reported reduced stress, others found GG to be a new source of stress, (4) implementation was variable, and (5) student-centered and high trust cultures was a prevailing feature. The study contributes to empirical research on grassroots level assessment reform on grading and assessment, emphasizing the role of feedback, leadership, and school culture in implementing changes in grading practices.

How does assessment in higher education value divergence? Educators' meaning-making about assessment

Keywords: assessment, assessment design, higher education, qualitative methods

Presenting Author: Jack Walton, Centre for Research in Assessment and Digital Learning, Deakin University, Australia; **Co-Author:** David Boud, Deakin University/University of Technology Sydney, Australia; **Co-Author:** Thomas Corbin, CRADLE, Deakin University, Australia; **Co-Author:** Tomohiro Ishida, Aichi University of the Arts, Japan; **Co-Author:** Juuso Henrik Nieminen, Deakin University, Australia; **Co-Author:** Joanna Tai, Deakin University, Australia

Higher education seeks to prepare graduates for uncertain and diverse futures, yet assessment practices remain rooted in industrial notions of standardisation and 'sameness' which often work against this aim. This study sought to explore how educators challenge this status quo by engaging with divergence as a design principle for assessment. Nine interviews were conducted with educators from three Australian universities, from disciplines including engineering, architecture, anatomy, and creative writing. Following a holistic and narrative approach to analysis, we examined how these educators made meaning about their assessment designs. Participants' personal histories and identities as progressive assessment designers were central to this meaning-making. While participants valued divergence in assessment tasks, this did not always extend to assessors' judgment practices. Taken in dialogue with the literature, our data suggest four main problems of divergence—validity, inclusion, preparedness, and distinctiveness—that can be further explored to support progressive assessment design.

Parallel session 6 2

26 August 2026 16:00 - 17:30

Pangea

Single Paper

Assessment and Evaluation

Understanding Peer Feedback: Processes and Outcomes

Keywords: assessment for learning, formative feedback, higher education, mixed methods, peer assessment, peer feedback, secondary education

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Barbara Kempf, University of Leeds, United Kingdom

The Challenge of Integration: How Students Deal with Multiple Peer-Feedback Messages

Keywords: formative feedback, higher education, peer assessment, peer feedback

Presenting Author: Julia Hornstein, University of Augsburg, Germany; **Co-Author:** Martin Greisel, University of Augsburg, Germany; **Co-Author:** Ingo Kollar, University of Augsburg, Germany

Although computer-supported peer-feedback can enhance student learning, little is known about how students integrate multiple peer-feedback messages. In an exploratory study involving $N = 100$ students, we employed hierarchical cluster analysis to identify distinct patterns in how students integrate multiple peer-feedback messages (RQ1). We also examined whether objective feedback features (accuracy, relevance, implementability, structural quality) and perceived feedback quality (subjective adequacy, subjective feedback quality) predict these patterns (RQ2) and whether these patterns relate to peer-feedback outcomes, such as implementation frequency, revision quality, and domain-specific performance (RQ3). Results revealed three integration patterns: confirmation pattern, conflict pattern, and selective integration pattern. Objective feedback features, but not perceived feedback quality, predict which pattern students adopted. The patterns differed only in how frequently students implemented the feedback received. Overall, these findings suggest that integrating multiple peer-feedback messages is a complex process shaped by objective feedback features, offering important implications for designing computer-supported peer-feedback environments.

Promoting dialogic peer feedback: Exploring communicative approaches and revision depth

Keywords: assessment for learning, formative feedback, peer feedback, secondary education

Presenting Author: Amber Van Hoe, VIVES University of Applied Sciences, Belgium; **Co-Author:** Tijds Rotsaert, Ghent University, Belgium; **Co-Author:** Joel Wiebe, OISE, University of Toronto, Canada; **Co-Author:** Jim Slotta, OISE, University of Toronto, Canada; **Co-Author:** Tammy Schellens, Ghent University, Belgium

Research increasingly highlights the collaborative potential of dialogic peer feedback (PF), yet little is known about how students communicate during PF dialogues and how these interactions support learning. This study investigates 60 PF dialogues embedded in a computer-supported collaborative inquiry learning environment with 149 secondary students. We analysed students' use of sentence starters during PF dialogues, their communicative approaches, and the relationship between interaction patterns and subsequent revisions. Sentence starters appeared in about half of all PF messages but did not reliably elicit discussion-based exchanges. Most dialogues were interactive, though typically evaluative and mimicking authoritative teacher feedback. Discussion-based patterns occurred only occasionally. Interactional richness strongly predicted revision depth: interactive and discussion-oriented dialogues led to substantive content revisions, whereas non-interactive authoritative exchanges produced surface-level changes. These findings clarify how communicative approaches shape the learning potential of PF and underscore the need to explicitly foster dialogic interaction within classroom and assessment cultures.

Beyond implementation: Is Revision Quality Influenced by Peer Feedback Features in Higher Education?

Keywords: assessment for learning, higher education, mixed methods, peer feedback

Presenting Author: Maël Pire, UCLouvain, Belgium; **Co-Author:** Florence Van Meenen, UCLouvain, Belgium; **Co-Author:** Franck Verschuren, UCLouvain, Belgium; **Co-Author:** Liesje Coertjens, Université catholique de Louvain (UCL), Belgium

Recent studies have demonstrated that PF features (e.g., type, recurrence) influence whether students implement it or not. However, binary measures of implementation (0/1) do not distinguish between minor and intensive modifications. Furthermore, research accounting for PF relevance when assessing revision quality remains scarce. Hence, the quality of students' revisions and its link with PF features remains poorly understood. The present study relies on a mixed methods approach to examine this. Using a quasi-experimental design, 38 medical students were asked to revise the first draft of their essay. The course teacher evaluated revision quality while providing justifications for his judgements.

Preliminary findings suggest that explanation PF leads to comparatively more high-quality revisions than identification and suggestion PF. Recurring comments (i.e., mentioned by multiple peers) seem to lead less often to high-quality revisions. Further analyses including multilevel cumulative logit modelling as well as qualitative analyses on the teacher's justifications will be presented.

Parallel session 6 3

26 August 2026 16:00 - 17:30

KBG 208

Collaborative Space

Assessment and Evaluation

Action-Based Assessment

Keywords: assessment design, assessment quality, competence assessment, measurement validation

Interest group: SIG 01 - Assessment and Evaluation

Systematically Validating an Action-Based Assessment of Financial Competence for Young Adults

Keywords: assessment design, assessment quality, competence assessment, measurement validation

Presenting Author: Hayden Mueller, Technical University of Munich, Germany; **Presenting Author:** Andreas Kraitzek, Technical University of Munich, Germany; **Co-Author:** Manuel Förster, Technical University of Munich, Germany; **Co-Author:** Baerbel Fuerstenau, TU Dresden, Germany; **Co-Author:** Roland Happ, Leipzig University, Germany; **Co-Author:** Mandy Hommel, OTH Amberg-Weiden, Germany; **Co-Author:** Christin Siegfried, University of Potsdam, Germany; **Co-Author:** Eveline Wuttke, Goethe-Universität Frankfurt, Germany; **Co-Author:** Andreas Emter, University of Potsdam, Germany; **Co-Author:** Johann Faltermeier, OTH Amberg-Weiden, Germany; **Co-Author:** Sebastian Heide, Leipzig University, Germany; **Co-Author:** Jessica Nitsch, TU Dresden, Germany; **Co-Author:** Josseline Stock, Leipzig University, Germany; **Co-Author:** Vinzent Weber, Goethe-University Frankfurt, Germany; **Co-Author:** Elisa Wagner, Leipzig University, Germany

This open-format session presents a research project which develops and systematically validates an action-based assessment of financial competence for young adults. Addressing limitations of traditional knowledge-focused measures, the project uses situational judgment tests (SJTs) to capture cognitive and non-cognitive dispositions as well as problem-solving processes in realistic financial scenarios. The session outlines a comprehensive validation strategy aligned with the Standards for Educational and Psychological Testing (AERA, 2014), drawing on multiple sources of validity evidence. Methods include expert interviews to define domain content, think-aloud protocols, eye-tracking, and emotion recognition to analyze response processes, as well as planned large-scale field studies to examine internal structure and relations to other variables. Initial findings from qualitative (expert interviews) and process data (think-alouds and eye-tracking) will be shared. Participants are invited to collaboratively discuss task design, validation challenges, upcoming field studies, and the assessment's potential use as both an evaluative and instructional tool.

Parallel session 6 4

26 August 2026 16:00 - 17:30

MIN 009

Workshop Format

Assessment and Evaluation

Teacher Professional Development for Effective Feedback

Keywords: feedback, feedback literacy, higher education, programmatic assessment

Interest group: SIG 01 - Assessment and Evaluation

Supporting teachers' professional learning to develop effective and efficient feedback practices

Keywords: feedback, feedback literacy, higher education, programmatic assessment

Presenting Author: Harry Stokhof, HAN University of Applied Sciences, Netherlands; **Co-Author:** Loes vandenbergh, Hogeschool van Arnhem en Nijmegen/HAN University, Netherlands

Physiotherapy Advanced Collaborative Education (PACE) is an innovative educational concept, consisting of a concentric flexible curriculum in which students follow individual learning paths, collecting evidence of their professional development for their portfolios by working on personal learning questions. Although successful for student learning, one of the greatest challenges teachers now face in PACE is to optimize the effectiveness and efficiency of their feedback practices. Teachers feel overwhelmed by the number of individual feedback requests and are concerned with the limited impact of their feedback on students' deep learning. To support teachers to understand and apply emergent evidence-informed theoretical knowledge for their own specific challenges, a participatory design-based research project has been set up. In the workshop we will introduce and let participants explore a conceptual tool that supports teachers to develop actionable knowledge and to design professional experiments that bridge the gap between evidence-informed theory and urgent practical feedback challenges.

Parallel session 6 5

26 August 2026 16:00 - 17:30

MIN 011

Demo Session

Higher Education

AI supported Course Evaluation

Keywords: assessment, evaluation, feedback literacy, GenAI

Interest group: SIG 01 - Assessment and Evaluation

AI Beacon: Agent-supported Cycle for Bidirectional Course Evaluation

Keywords: assessment, evaluation, feedback literacy, GenAI

Presenting Author: Maria Pannatier, Swiss Federal Institute of Technology, EPFL Lausanne, Switzerland; **Presenting Author:** Julius Kooistra, Berner Fachhochschule, Switzerland; **Co-Author:** Johann Groll, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland; **Co-Author:** Lucia Gomez Teijeiro, Bern University of Applied Sciences, Switzerland; **Co-Author:** Branka Hadji Misheva, Bern University of Applied Sciences, Switzerland

While teaching evaluation is essential for quality assurance in higher education, current approaches are often one-directional and offer limited actionable guidance. AI Beacon is an innovative, agentic, AI-powered solution that supports every stage of the evaluation cycle. It helps teachers design meaningful evaluations, guides students to provide holistic, constructive feedback, assists pedagogical advisors to identify key action points, and supports teachers to implement improvements. Designed as a standalone generative AI tool, AI Beacon enables continuous, actionable, and pedagogically sound evaluation. It addresses three persistent implementation gaps: the knowledge gap (limited expertise in designing evaluations and providing constructive feedback), the management gap (time constraints and technological difficulties) and the follow-up gap (difficulty translating data into teaching improvements). The prototype will be completed in spring 2026 and demonstrated within the Digital Training Companion platform, which integrates AI Beacon and uses the Digital Competency Training Assessment Model to inform the evaluation process.

Parallel session 7 1

27 August 2026 11:30 - 13:00

Pangea

Single Paper

Assessment and Evaluation

Student Engagement and Identity in Feedback

Keywords: feedback, feedback literacy, formative feedback, higher education, motivation, qualitative methods

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Lynn Van den Broeck, KU Leuven (BE), Belgium

Patterns of feedback engagement by prospective STEM students after taking a positioning test

Keywords: feedback literacy, formative feedback, higher education, qualitative methods

Presenting Author: Laurie Peeters, KU LEUVEN, Belgium; **Co-Author:** Wim Van Petegem, KU LEUVEN, Belgium; **Co-Author:** Carolien Van Soom, University of Leuven, Belgium; **Co-Author:** Greet Langie, KU LEUVEN, Belgium

Making an informed study choice can present challenges for prospective students. In Flanders, Belgium, entrance exams are only for medicine, dentistry, and veterinary programmes. To support students entering STEM fields, positioning tests are organised. These are formative, low-stakes assessments that provide participants with feedback on the required knowledge to start with the STEM programme of their choice. This study investigates how prospective students engage with this feedback and how this influences their pre-entry preparation. Based on 14 semi-structured interviews with participants, including a longitudinal follow-up, and a focus group with study coaches, this study explores individual differences in feedback engagement through inductive qualitative analysis. Preliminary results emphasise the need for an ecological perspective as individual and contextual factors are related to feedback engagement. The findings of this study will inform the optimisation of the feedback instruments such that prospective students are better supported during the transition to higher education.

Feedback, identity and self: feedback as a transformational practice

Keywords: feedback, higher education, motivation, qualitative methods

Presenting Author: Juuso Henrik Nieminen, Deakin University, Australia; **Co-Author:** Judith Gulikers, Wageningen University, Netherlands; **Co-Author:** Renske de Kleijn, UMC Utrecht, Netherlands; **Co-Author:** Martijn Leenknecht, HZ University of Applied Sciences, Netherlands

Feedback is known to shape student learning in higher education, yet less is known about how feedback shapes students themselves. We conceptualise feedback as a transformative practice that holds the potential of shaping who students are and who they are becoming as they attend higher education. The empirical question then is, how does feedback shape students? To address this question, we conduct a narrative analysis of 57 undergraduate students' stories about feedback and 'who they are', based on interview material gathered from Hong Kong. Three identity narratives are uncovered: 1) student identity, 2) professional identity, 3) neoliberal identity. We also show how feedback may transform students' stories about themselves. Overall, we call for a deeper understanding of not just how students engage with feedback, but why, and how these experiences leave traces on students themselves.

Student sense-making in feedback processes: A triangulated study of process

Keywords: feedback, feedback literacy, higher education, qualitative methods

Presenting Author: Rasmus R. Hansen, Aarhus University, Denmark; **Co-Author:** Ida Bang Hansen, Aarhus University, Denmark; **Co-Author:** Christopher Neil Prilop, Aarhus University, Denmark

This paper examines how undergraduate students make sense of feedback during structured teacher–student interactions. The study conceptualises feedback as a process that solicits sense-making, which can lead to feedback uptake and revisions, as students interpret feedback. Nineteen students completed a written analysis, a ten-minute feedback interaction, and a subsequent revision. Data included screen-recorded interactions and revisions, survey measures of feedback literacy, and interviews with seven participants. Reflexive thematic analysis was employed to discover influential factors in sense-making processes. Four factors shaped sense-making: engagement, self-evaluation, context, and relation. These factors guided triangulation between the interviews and screen-recorded process data. From this triangulation, three levels of sense-making were identified: micro-level interpretation of individual comments, meso-level integration of multiple comments, and macro-level reflection on the overall encounter. The findings show how feedback literacy influences whether students recognise and act upon feedback cues and highlight implications for designing effective feedback interactions.

Parallel session 7 2

27 August 2026 11:30 - 13:00

KBG 208

Single Paper

Understanding Rubrics, Criteria, and Feedback Effects

Keywords: assessment, assessment for learning, assessment quality, assessment tool, feedback, formative feedback, rubrics, self-feedback, transparency

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Anke Swanenberg, Wageningen University and Research Centre, Netherlands

What Do Students Actually Do with Rubrics? The Impact of Level Order, Feedback, and Task Occasion

Keywords: assessment tool, feedback, formative feedback, rubrics

Presenting Author: Leire Pinedo, Universidad de Deusto, Spain; **Co-Author:** Alazne Fernández Ortube, University of Deusto, Spain; **Co-Author:** David Zamorano, Universidad de Deusto, Spain; **Co-Author:** Lucía Barrenetxea-Mínguez, University of Deusto, Spain; **Co-Author:** Iván Sánchez-Iglesias, Universidad Complutense de Madrid, Spain; **Co-Author:** Ernesto Panadero, Dublin City University, Ireland

Rubrics are common tools in educational evaluation, yet little is known about (1) how students engage with them, (2) what order of rubric presentation is best, or (3) what type of feedback is best to accompany a rubric. This study examined how the order of rubric performance levels and four types of feedback (no feedback, rubric-based, process-oriented, and product-oriented) influenced student self-regulatory actions in two landscape analysis tasks. 120 first-year university students think-aloud while using the rubric across five phases: reading the rubric, orally analyzing a landscape, writing the analysis, receiving feedback according to condition, and repeating the oral and written analyses for a second landscape. Results showed that positive actions, particularly planning and rubric review, predominated, while negative actions occurred in early phases. While rubric order and feedback type had no significant effects, task phase significantly influenced the frequency of positive and negative actions, revealing dynamic shifts across phases.

The case for rethinking assessment criteria

Keywords: assessment for learning, assessment quality, rubrics, transparency

Presenting Author: Jack Walton, Centre for Research in Assessment and Digital Learning, Deakin University, Australia; **Presenting Author:** Liesbeth Baartman, University of Applied Sciences Utrecht, Netherlands

In this conceptual paper we revisit dominant assumptions about assessment criteria in higher education. We argue that normative, prespecified criteria constrain both learning and the representation of student achievement. In the literature, we identified critiques concerning transparency, validity, and power in relation to assessment criteria. Drawing on Bernstein's metaphors of reservoirs and repertoires, we propose an alternative conceptualisation of criteria as sociocultural meaning-making resources embedded in disciplinary communities. This reconceptualisation acknowledges the fuzzy boundaries of quality and the negotiated nature of professional standards. We argue that students should not be expected to master entire reservoirs but to gradually grow into disciplinary expertise. Understanding criteria as collective, situated judgements invites new possibilities for assessment design and for articulating what meaningful assessment entails.

Effects of Rubric Order and Feedback Type on Physiological Arousal: An Electrodermal Activity Study

Keywords: assessment, assessment for learning, feedback, rubrics

Presenting Author: David Zamorano, Universidad de Deusto, Spain; **Co-Author:** Héctor J. Pijera-Díaz, University of Jyväskylä, Finland; **Co-Author:** Leire Pinedo, Universidad de Deusto, Spain; **Co-Author:** Alazne Fernández Ortube, University of Deusto, Spain; **Co-Author:** Lucía Barrenetxea-Mínguez, University of Deusto, Spain; **Co-Author:** Ernesto Panadero, Dublin City University, Ireland

Despite the central role of rubrics and feedback in formative assessment, their design-related effects on students' real-time engagement remain underexplored, a gap that multimodal learning analytics and psychophysiological indicators such as electrodermal activity (EDA). This study examined whether rubric order (ascending vs. descending performance levels) and feedback type (no feedback, product-, process-, or rubric-based) affected students' physiological arousal while completing landscape analysis tasks. Eighty-seven first-year university students participated in a controlled laboratory study in which EDA was recorded across five task phases. Results showed no significant effects of rubric order or feedback type on arousal, nor did arousal predict task performance. These findings suggest that, although EDA captures meaningful fluctuations in engagement across task phases, the specific rubric and feedback manipulations examined here did not modulate physiological arousal.

Effects of Teacher Feedback, Rubrics, and Exemplars on Italian Psychology Students' Writing

Keywords: feedback, formative feedback, rubrics, self-feedback

Presenting Author: Ligia Tomazin, University of Trento (I), Italy; **Co-Author:** Paola Venuti, University of Trento (I), Italy; **Co-Author:** Arianna Bentenuto, University of Trento (I), Italy; **Co-Author:** Silvia Perzoli, University of Trento (I), Italy; **Co-Author:** Anna Serbati, University of Trento (I), Italy; **Co-Author:** Federica Picasso, Faculty of Education - University of Genova (Italy), Italy; **Co-Author:** Anastasiya Lipnevich, The Graduate Center of the City University of New York, United States

Despite the critical role of feedback in improving learning and performance, providing high-quality feedback in higher education is a complex and time-consuming task for instructors. Therefore, it is essential to explore alternative feedback approaches that are both effective and sustainable. This experimental study examined the impact of rubrics, annotated exemplars, and teacher comments on college students' writing performance. Participants were 71 second-year master's students in Clinical Psychology (56 women) from an Italian university. Students completed a writing assignment in class and submitted it via Moodle platform before being randomly assigned to one of three feedback conditions: rubric, annotated exemplars, or teacher comments. They then revised their assignments in class using the feedback received. Findings revealed no statistically significant differences among groups ($p = .654$), suggesting that rubrics and annotated exemplars can produce performance improvements comparable to those achieved through teacher feedback.

Parallel session 7 3

27 August 2026 11:30 - 13:00

KBG 224

Single Paper

Assessment and Evaluation

Teacher Assessment Literacy and Student Partnership

Keywords: assessment, assessment as learning, assessment for learning, assessment literacy, formative assessment, measurement validation, peer assessment, performance assessment, secondary education, self-assessment

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Javier Fernández, Universidad Autónoma de Madrid, Spain

Between Support and Selection: Antinomies of Classroom Assessment and Pre-Service Teachers' Beliefs

Keywords: assessment for learning, assessment literacy, formative assessment, performance assessment

Presenting Author: Marius Diekmann, University of Münster, Germany; **Co-Author:** Pia Wensing, University of Münster, Germany; **Co-Author:** Sabine Gruehn, Institute for Educational Science, Germany

Performance assessment is a core aspect of school-based learning, shaped by the tension between supportive and selective purposes. Teachers' beliefs about "good" assessment strongly influence how they navigate these competing demands. While prior research has identified distinct assessment-related belief types among in-service teachers, little is known about such beliefs during teacher education. Using a newly developed questionnaire, we surveyed about 700 pre-service teachers (Bachelor's and Master's programs) at the University of Münster (Germany) regarding their assessment beliefs and conceptions of "good" assessment. Results show strong overall endorsement of formative purposes, with only minor differences between Bachelor's and Master's students. Hierarchical cluster analysis revealed distinct belief patterns, ranging from selective or disciplinary orientations to formative support with varying views on external differentiation. These findings highlight the complexity of assessment beliefs early in teacher education and underscore the importance of reflecting on professional beliefs to foster openness, adaptability, and learner-centered assessment practices.

From Uncertainty to Confidence: An Emotional Journey in Becoming a Teacher Assessor

Keywords: assessment, assessment literacy, formative assessment, self-assessment

Presenting Author: Jana Kratochvilova, Masaryk University, Faculty of Education, Czech Republic; **Co-Author:** Jarmila Bradova, Institute for Research in School Education, Slovakia; **Co-Author:** Katerina Vlckova, Masaryk University, Czech Republic; **Co-Author:** Julie Grombířová, Masaryk University, Czech Republic; **Co-Author:** Jakub Gyönyör, Masaryk University Faculty of Education, Czech Republic; **Co-Author:** Magdalena Tylová, Faculty of Education, Masaryk University, Czech Republic

This study examines how teachers develop assessment literacy through professional learning courses focused on comprehensive, developmental formative assessment within learning communities. It investigates the socio-emotional dimension of this development in relation to changes in assessment practice. The

sample comprised 106 in-service primary and lower-secondary school teachers who participated in ten-month professional development courses. Data were collected through written self-reflections on their development of assessment literacy, plans for ongoing development, perceived impacts on students, knowledge sharing with colleagues, and through a self-assessment scale of assessment literacy development. The findings reveal emotional tensions, identity shifts, and relational dynamics that shape teachers' understanding and enactment of their roles as assessors. The study contributes to current conceptualisations of teacher assessment literacy by foregrounding the socio-emotional dimension of learning to assess and offering implications for the design of teacher professional development.

Formative assessment practices within teachers' assessment literacy: development of an instrument

Keywords: assessment for learning, assessment literacy, formative assessment, measurement validation

Presenting Author:Karin Portmann, Pädagogische Hochschule St. Gallen, Switzerland; **Co-Author:**Robbert Smit, University of Teacher Education St.Gallen, Switzerland; **Co-Author:**Rahel Schmid, St.Gallen University of Teacher Education, Switzerland; **Co-Author:**Felix Kruse, University of Teacher Education St. Gallen, Switzerland; **Co-Author:**Anna-Lena Ullrich, University of Teacher Education St. Gallen, Switzerland

Assessments are integral to effective teaching and learning (Pastore, 2023). Assessment Literacy (AL) is a context-dependent, dynamic, and multidimensional construct comprising knowledge, skills, and beliefs (e.g., DeLuca et al., 2018). Formative assessment practices (FAP) are central to this framework. A measurement instrument that focuses on FAP in line with Wiliam and Thompson (2008) has not yet been established. This study addresses the research question: To what extent do newly developed items measure the construct FAP? The instrument was built from existing (e.g., Lysaght et al., 2017; Veugen et al., 2022) and new items. It was qualitatively piloted through expert review and think-aloud testing and quantitatively tested with teacher education students (n = 213). First explorative results indicate a FAP model with five factors. This instrument may be used in teacher education and professional development to identify areas for improvement.

Reframing Assessment in Schools through Students as Partners in Assessment

Keywords: assessment as learning, assessment literacy, peer assessment, secondary education

Presenting Author:GIORGIA SLAVIERO, Università degli Studi di Padova, Italy; **Co-Author:**Valentina Grion, Pegaso Telematic University, Italy

Recent European surveys (TALIS, 2025) report the widespread use of observational practices and continuous feedback as assessment strategies, while student involvement remains limited. International research (Black & William, 1998; Boud & Soler, 2016) suggests that, beyond summative and formative approaches, assessment can be reconceptualised as a shared and learning-oriented process as framed within the Students as Partners in Assessment (SaPA) approach (Nieminen et al., 2025). This doctoral research adopts a multi-phase mixed-methods design comprising three exploratory phases and an experimental one. The first phase examines assessment practices adopted by Italian teachers through document analysis and selected school-based case studies. The second phase consists of two systematic reviews, international and national, investigating SaPA within K12 contexts. A third exploratory phase explores teacher professional development models for assessment. Building on these data, the study designs and evaluates a self-managed professional learning environment for teachers focused on Students as Partners in Assessment (SaPA).

Parallel session 7 4

27 August 2026 11:30 - 13:00

MIN 009

Poster Presentation

Assessment and Evaluation

Feedback Cultures and Assessment Transformation

Keywords: artificial intelligence, assessment literacy, assessment policy, equity, evaluation, feedback, feedback literacy, formative feedback, GenAI, higher education, mixed methods, post-graduate education, primary education, self-assessment, student evaluations of teaching

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Burcu Arslan, Educational Testing Service, Germany

SET in motion: Sparking a culture of constructive feedback

Keywords: evaluation, feedback, higher education, student evaluations of teaching

Presenting Author:Niels van der Baan, Maastricht University, Netherlands; **Co-Author:**Marloes Nederhand, Erasmus University Rotterdam, Netherlands; **Co-Author:**Bas Giesbers, Maastricht University, Netherlands; **Co-Author:**Bjorn de Koning, Erasmus University Rotterdam, Netherlands

Student Evaluation of Teaching (SET) in higher education can contribute to improving educational quality. However, various issues can hinder SET's potential to contribute to an educational quality culture. For example, response rates on online SET questionnaires are low, which hinder the generation of reliable SET results. In addition, feedback provided in SET is not always constructive and can even be harmful. These issues can be detrimental for establishing an educational quality culture and realizing SET's full potential. The aim of this project is twofold. On the one hand, we scale up a previously successful innovation to increase the response rates on SET. On the other hand, we will develop a new innovation on how to provide constructive feedback in SET. The implementation and effectiveness of this approach will be studied to strengthen the innovation and to offer practical recommendations for realizing SET's full potential in the educational quality assurance.

Exploring the role of space in international doctoral students' feedback conversations

Keywords: feedback, formative feedback, higher education, post-graduate education

Presenting Author:Xin Liu, Deakin University, Australia; **Co-Author:**Joanna Tai, Deakin University, Australia; **Co-Author:**Margaret Bearman, Deakin University, Australia; **Co-Author:**Mollie Dollinger, Curtin University, Australia

Feedback is not only a discursive practice, but a spatial one. For international doctoral students who have relocated to pursue their studies, spatial contexts may carry particular weight, since supervisory feedback conversations unfold within social and material environments that differ from those of their home contexts. Drawing on constructivist grounded theory, we analysed two rounds of interviews involving 13 international doctoral students. Findings suggest that the social nature of informal spaces fostered incidental feedback conversations and feedback relationships to enhance learning. The symbolic meanings embedded in these spaces facilitated emotions-sharing, trust-building, and sociocultural learning, but their exposed and uncontrollable nature occasionally impeded feedback conversations. By contrast, formal spaces, with more structures and resources, supported more productive task-specific feedback conversations. However, institutional constraints often limited access to suitable formal spaces. This study offers insight into the ways that feedback conversations could be improved through consideration of formal and informal spaces.

Beyond Grades: Futures Thinking for Transforming Assessment in Education

Keywords: assessment policy, equity, evaluation, self-assessment

Presenting Author:Katrin Saks, University of Tartu, Estonia

Traditional grading systems have been widely criticized for fostering extrinsic motivation, surface-level learning, and inequity. This study explores how eliminating grades can transform educational practices by developing future-oriented scenarios for a gradeless school. Using futures thinking, systems thinking, and exponential thinking, the research examines systemic, technological, and socio-cultural conditions that enable learner-centered, equitable, and sustainable assessment models. Two futures workshops with diverse stakeholders generated scenarios emphasizing democratic participation, personalized learning, and AI-supported feedback. Scenario 1 prioritizes community-based governance and adaptability in uncertain socio-political contexts, while Scenario 2 focuses on equity, autonomy, and technological innovation. A STEEP+V analysis highlights key drivers, including social fragmentation, technological ethics, economic inequality, and environmental sustainability. Findings suggest that gradeless schooling is not merely a pedagogical shift, but a systemic transformation that requires aligning innovation, equity, and democratic resilience to create education models that prioritize growth over grades.

Assessment Literacy and Classroom Assessment Policy implementation in Chilean Primary Schools

Keywords: assessment literacy, assessment policy, mixed methods, primary education

Presenting Author:Stefania Gutiérrez Martinetti, Kings' College London, United Kingdom

Classroom assessment involves the collection and use of evidence to inform teaching and learning, requiring teachers to continuously develop their assessment literacy. Although widely recognised as essential, assessment literacy remains conceptually contested. This study conceptualises it as a multidimensional, contextually situated construct that shapes classroom assessment and potentially influences how educational policies are interpreted and implemented. While existing research has examined its relationship with teacher practice and professional learning, its role in policy implementation remains under-explored, particularly in the Chilean context. Adopting an interpretivist mixed-methods approach, this study explores how the assessment literacy of Chilean educational stakeholders influences the implementation of educational policies. Data collection includes semi-structured interviews and the Approaches to Classroom

Assessment Inventory. A pilot study has been completed to refine data collection instruments. At the time of submission, data analysis is ongoing, and emerging analytical insights will be presented in the conference poster.

Learning with, by, and through GenAI feedback: Novel teacher feedback literacies in a GenAI world

Keywords: artificial intelligence, feedback, feedback literacy, GenAI

Presenting Author: Christopher Neil Prilop, Aarhus University, Denmark

The rapid expansion of generative AI (GenAI) in higher education is transforming how students engage in feedback processes. While GenAI enables more frequent feedback interactions, such interactions often fail to stimulate meaningful sense-making, as GenAI systems are designed to provide quick fixes rather than support co-constructive dialogue. Thus, despite increased access to feedback, GenAI may lead to “senseless” feedback interactions and further devalue human feedback practices. Building on teacher feedback literacy, this conceptual paper argues that teachers require new feedback literacies to design feedback processes that genuinely foster students’ sense-making and feedback uptake. Extending Salomon and Perkins’ notion of technology as an intellectual partner, the paper introduces a GenAI teacher feedback literacy framework for designing learning *with*, *by*, and *through* GenAI feedback. The framework outlines key teacher competencies needed to scaffold dialogic feedback interactions, support lasting cognitive and metacognitive learning, and design a reconfigured feedback ecosystems.

Parallel session 7 5

27 August 2026 11:30 - 13:00

MIN 011

Poster Presentation

Assessment and Evaluation, Teaching and Teacher Education

Assessing Competencies and Understanding Teacher Practices

Keywords: adaptive assessment, assessment, competence assessment, evaluation, feedback literacy, grading, large-scale assessment, measurement validation, motivation, peer feedback, primary education, quantitative methods

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Alex Kozulin, Achva College and Feuerstein Institute, Israel

Validating the Radboud Computational Thinking Test. An unplugged assessment for upper-primary pupils

Keywords: competence assessment, measurement validation, primary education, quantitative methods

Presenting Author: Jo van Schaik, Behavioural Science Institute, Radboud University Nijmegen, Netherlands; **Co-Author:** Rosanne Sluiter, Radboud University Nijmegen, Netherlands; **Co-Author:** Viola Siegmeier, Behavioural Science Institute, Radboud University Nijmegen, Netherlands; **Co-Author:** Noortje Janssen, Radboud University, Netherlands; **Co-Author:** Carlijn Veldhorst, Behavioural Science Institute, Radboud University Nijmegen, Netherlands; **Co-Author:** Ard Lazonder, Radboud University, Netherlands

There is broad agreement that computational thinking (CT) is a foundational problem-solving competence with cross-curricular benefits, yet the field lacks validated assessments to test these claims. Because many existing CT tests rely on programming tasks, it remains unclear whether they measure genuine CT skills or programming experience. To address this, we developed the XXXX Computational Thinking Test (RCTT), an unplugged paper-and-pencil assessment targeting abstraction, decomposition, and algorithmic thinking. This preregistered validation study (N=550 upper-primary pupils) uses confirmatory factor analysis, classical test theory, and a 1PL model, alongside convergent, discriminant, and fairness testing. The project aims to deliver a validated, publicly available CT assessment to support evidence-based CT education and research.

Longitudinal Patterns of ICT-Related Teacher Collaboration: Evidence from the ICILS Study

Keywords: assessment, evaluation, large-scale assessment, quantitative methods

Presenting Author: Sahsenem Öz, University of Jyväskylä, Finland; **Co-Author:** Asko Tolvanen, University of Jyväskylä, Finland; **Co-Author:** Kaisa Leino, University of Jyväskylä and Finnish Institute for Educational Research, Finland; **Co-Author:** Raija Hämäläinen, University of Jyväskylä, Finland

Using the longitudinal ICILS teacher data (2018 and 2020), this study aims to identify collaboration subgroups for ICT learning among 1,246 Finnish in-service teachers, the factors predicting group membership, and the stability of these groups over time. The analyses involved latent class analysis (LCA) and latent transition analysis (LTA). LCA results revealed three collaboration groups at timepoint 1 (High–Collaborator, Observer–Discussor, and Low–Collaborator) and four groups at timepoint 2, including an emergent Non–Observer group. ICT self-efficacy, participation in formal learning activities, support for student collaboration, time availability, and age predicted membership in classes. LTA results showed that classes remained largely stable, except for one class, which showed a moderate membership shift over time. Finally, several pandemic-related factors were associated with membership in the High–Collaborator class at the post-pandemic time point, including longer online teaching engagement and limited access to readily available online resources.

Latent Profiles of Estonian Teachers’ Beliefs on Grading, Motivation, and Self-Regulated Learning

Keywords: assessment, grading, motivation, quantitative methods

Presenting Author: Liia Ots, Tallinn University, Institute of Educational Sciences, Estonia; **Co-Author:** Kati Aus, Tallinn University, Institute of Educational Sciences, Estonia; **Co-Author:** Piret Oppi, Tallinn University, Estonia; **Co-Author:** Eve Kikas, Tallinn University, Estonia

To create change in assessment and evaluation practices, we must first understand the belief systems that maintain teachers’ teaching practices in classrooms. This study examines the interplay between teachers’ beliefs about grading, their skepticism toward self-regulated learning (SRL), and their views on motivational strategies (controlled vs. autonomous). Based on a large-scale survey of 3,853 general education teachers in Estonia, a Latent Profile Analysis (LPA) was conducted to identify distinct subgroups. Preliminary results reveal three unique teacher profiles, demonstrating that views on assessment are not isolated but are interconnected with broader pedagogical beliefs about control and autonomy. Multinomial logistic regression revealed that a Master’s degree was a significant predictor of membership in the autonomy-supportive profile, whereas lower qualification levels were associated with less supportive beliefs.

One Size fits all? Are Pre-service Teachers’ Feedback Messages Skill-Sensitive?

Keywords: adaptive assessment, feedback literacy, peer feedback, quantitative methods

Presenting Author: Alina Sheppard-Bujtor, University of Augsburg, Germany; **Co-Author:** Martin Greisel, University of Augsburg, Germany; **Co-Author:** Julia Hornstein, University of Augsburg, Germany; **Co-Author:** Ingo Kollar, University of Augsburg, Germany

While peer-feedback is recognized as an effective pedagogical approach, not every learner benefits from it. This study investigates adaptivity within the peer-feedback context, defined as the alignment of the feedback messages’ make up and content on the one hand with the recipient’s skill level on the other. Analyses will pertain to whether and how adaptivity is reflected in feedback feature constellations—for example the proportion of explanations and solutions—within the feedback messages. And how this influences subsequent outcomes like the acceptance and implementation of the feedback by its receiver. Pre-service teachers generated a task solution, provided feedback to two random peers, and in turn received feedback to revise their initial task solution. The processing of data is still ongoing. Results may inform the design of scaffolds that support students in the generation of adaptive feedback.

Parallel session 8 1

27 August 2026 14:00 - 15:30

Pangea

Symposium

Assessment and Evaluation

The e-portfolio paradox: Integrating assessment for and of learning to foster student agency

Keywords: assessment for learning, assessment tool, higher education, programmatic assessment

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Bas Agricola, HU University of Applied Sciences, Netherlands

Organiser: Bas Agricola, HU University of Applied Sciences, Netherlands

Discussant: Juuso Henrik Nieminen, Deakin University, Australia

e-Portfolios are widely recognized as high-impact pedagogical practices, with potential to foster agency and make the process of learning visible. Implementation research however confirms a chronic failure to fully realize this promise. There is a conflict between the portfolio as an assessment product (summative) versus a learning process (formative). The portfolio often reverts to a mere digital repository that undermines its developmental function. Programmatic assessment requires students to expose their struggles and respond honestly to critical feedback. Yet, if this vulnerability is penalized due to

summative purposes, students engage in strategic documentation, reducing rich learning trajectories to fragmented snapshots of success. This symposium addresses the central question: How can e-portfolio systems and their associated feedback practices maximize student agency within Programmatic Assessment? In three papers this conflict is investigated across three interconnected levels: the design of the assessment system, the learner's strategic response, and the facilitative dialogue provided by coaches.

How different e-portfolio components stimulate student learning in programmatic assessment

Presenting Author: Bas Agricola, HU University of Applied Sciences, Netherlands; **Co-Author:** Kitty Meijer, HU University of Applied Sciences Utrecht; Open University of the Netherlands, Netherlands; **Co-Author:** Lieke Ceelen, Hogeschool Utrecht (University of Applied Sciences Utrecht), Netherlands; **Co-Author:** Liesbeth Baartman, University of Applied Sciences Utrecht, Netherlands

Programmatic Assessment (PA) relies on longitudinal data points to foster learning, with e-portfolios being essential tools for managing this process. However, it remains unclear how specific e-portfolio configurations support assessment for learning. This exploratory multiple case study investigated which e-portfolio components support feedback processes and how students perceive them. We employed a two-phase design involving interviews with curriculum and e-portfolio designers, followed by student focus groups across Dutch higher professional education programmes. Results identified five critical components: learning outcomes, data points, feedback, reflection, and overview. While many programmes aimed for student autonomy, but students often struggled with feedback overload or passive engagement due to ineffective designs. The study highlights a paradox: fostering genuine student agency in PA requires directive e-portfolio configurations rather than open-ended freedom. Consequently, optimizing e-portfolio design is a prerequisite for effective assessment for learning.

Designing for Student Agency: Institutional Choices in e-Portfolio Implementation

Presenting Author: Desirée Joosten-ten Brinke, Maastricht University, Netherlands; **Co-Author:** Joyce Moonen - Van Loon, Maastricht University, Netherlands; **Co-Author:** Ilse Sistermans, Maastricht University, Netherlands; **Co-Author:** Liesbeth Baartman, Maastricht University, Netherlands; **Co-Author:** Liesbeth Kester, Maastricht University, Netherlands

Electronic portfolios are widely promoted as tools to integrate assessment for and of learning by supporting feedback, reflection, and student agency. However, in higher education practice, e-portfolio use often shifts towards compliance-oriented documentation, undermining their developmental potential. This paper presents findings from an institution-wide e-portfolio implementation at Maastricht University, examining how contextualised institutional and pedagogical design choices shape student agency. Drawing on a comparative analysis of eight programmes across four faculties using PebblePad, combined with interviews with students and staff, the study explores how varying e-portfolio design choices influence learners' experienced agency. Preliminary findings indicate that student agency is not an inherent feature of e-portfolios but an outcome of intentional design choices and support structures. The paper contributes to ongoing debates on balancing educational steering and autonomy in portfolio-based assessment.

Coaches Use of E-Portfolios to Support Student Learning in Programmatic Assessment

Presenting Author: Andrea Oudkerk Pool, Fontys University of Applied Science, Netherlands; **Co-Author:** Kelly Beekman, Fontys University of Applied Science, Netherlands

Programmatic assessment in higher education aims to promote student-centred learning by supporting longitudinal development and strengthening students' self-regulated learning (SRL). E-portfolios make learning processes visible, aggregate feedback, and support progress decisions. Coaches, longitudinally involved with students, are expected to use e-portfolio information to guide learning and support SRL development. However, e-portfolios do not automatically support SRL and little is known about how coaches actually use them in practice. This exploratory qualitative study investigates how coaches use e-portfolios to support student learning within a programmatic assessment context, involving semi-structured interviews with faculty coaches at a Dutch University of Applied Sciences. Preliminary findings suggest coaches predominantly use portfolios for monitoring completion rather than scaffolding learning, with accountability pressures often privileging evaluative over developmental purposes. The findings will inform e-portfolio design and faculty development to foster meaningful student agency.

Parallel session 8 2

27 August 2026 14:00 - 15:30

KBG 208

Single Paper

Assessment and Evaluation, Motivational, Social and Affective Processes, Teaching and Teacher Education

Assessment, Judgement, and Learning in Professional Contexts

Keywords: assessment design, assessment of learning, authentic assessment, feedback, formative assessment, higher education, mixed methods, programmatic assessment, qualitative methods, summative assessment, sustainable assessment, workplace assessment

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Mollie Dollinger, Curtin University, Australia

Feedback with feelings: expressing judgements about performance in clinical environments

Keywords: feedback, formative assessment, qualitative methods, workplace assessment

Presenting Author: Margaret Bearman, Deakin University, Australia; **Co-Author:** Joanne Hilder, Southern Cross University, Australia; **Co-Author:** Damian Castanelli, University of Melbourne/Monash Health, Australia; **Co-Author:** Elizabeth Molloy, The University of Melbourne, Australia; **Co-Author:** Chris Watling, Royal College of Physicians and Surgeons Canada, Canada; **Co-Author:** Robyn Woodward-Kron, The University of Melbourne, Australia; **Co-Author:** Rola Ajawi, University of British Columbia, Canada

Feedback can evoke strong emotions but often feelings are treated as if they are something that the individual learner must resolve. This study takes a broader view of feelings as embodied, relational and social. A focused ethnography of feedback in intensive care medicine and surgical education traced the role of feelings in feedback in specialty training. In this context, feedback was mostly related to day-to-day work. Judgements about performance often appeared to be a matter of feeling; tacitly sensed as much as explicitly expressed. Feedback associated with formal workplace assessment could intensify feelings if the trainee was blindsided by a previously unknown concern. Supervisors often made judgements *about* feelings, directing trainees to express (or not) certain kinds of emotions. These findings suggest explicit engagement with feelings in feedback may be useful. Collective discussion and reflection about emotions may help all parties to better understand the role they play in learning.

Dilemmas in implementing the learning function of programmatic assessment: a multiple case study.

Keywords: assessment design, feedback, formative assessment, programmatic assessment

Presenting Author: Judith Gulikers, Wageningen University, Netherlands; **Co-Author:** Liesbeth Baartman, University of Applied Sciences Utrecht, Netherlands

Programmatic Assessment (PA), originating in medical education, is increasingly adopted in Dutch Vocational Education (VET) as institutions move from single exams to continuous evidence collection across learning trajectories. PA integrates learning and qualification decisions, using a backbone of learning outcomes to guide feedback and defer high-stakes decisions until patterns emerge. While principles of PA are established, contextualized design choices remain challenging—particularly balancing formative growth with summative accountability. This qualitative multiple-case study explored how four VET programs (beauty, marketing, media design, pedagogy) implement PA's learning function. Data from documents and stakeholder interviews revealed five recurring themes: strong learning vision, gaps between design and practice, complex/multiple backbones, need for electronic systems, and incomplete feedback loops. Four dilemmas emerged: product vs. process focus, documenting vs. dialogic feedback, learning from mistakes vs. passing data points, and freedom vs. structure. Findings suggest simplifying design and explicitly embedding formative cycles to strengthen learning within PA.

Reconceptualizing assessment of transdisciplinary learning

Keywords: authentic assessment, higher education, programmatic assessment, sustainable assessment

Presenting Author: Anke Swanenberg, Wageningen University and Research Centre, Netherlands; **Co-Author:** Judith Gulikers, Wageningen University, Netherlands; **Co-Author:** Yvette Baggen, Wageningen University, Netherlands; **Co-Author:** Perry den Brok, Wageningen University, Netherlands

Addressing complex global challenges in education requires transdisciplinary learning, where students integrate academic knowledge with societal perspectives. However, assessing transdisciplinary student learning is challenging; it is unpredictable and uncertain in nature; crosses disciplinary and societal boundaries; is collaborative; and has a dual focus on process and product. Existing assessment approaches, rooted in predetermined outcomes and individual accountability, are insufficient, highlighting the need for new assessment approaches. This contribution employs a theoretical synthesis approach to develop an integrated conceptual understanding of how assessment can support transdisciplinary learning. The synthesis draws on conceptual, empirical, and case-based literature that examine assessment practices in transdisciplinary learning. The goal is to reconceptualize assessment in transdisciplinary learning by exploring its purpose (*why*), the learning it wants to make visible (*what*), and the appropriate methods and tools (*how*). Bringing them together into a coherent assessment framework will support assessment designs that capture the complexity of transdisciplinary learning.

Assessing for Survival: A Longitudinal Study of Novice Medical Teachers' Conceptions

Keywords: assessment of learning, higher education, mixed methods, summative assessment

Presenting Author: Javier Fernández, Universidad Autónoma de Madrid, Spain; **Co-Author:** Lucía Barrenetxea-Mínguez, University of Deusto, Spain; **Co-Author:** Ernesto Panadero, Dublin City University, Ireland; **Co-Author:** Eneko Balerdi, Universidad de Deusto, Spain

This longitudinal study explores the evolution of novice medical teachers' conceptions of teaching and assessment during a transition to Problem-Based Learning. Using questionnaires alongside interviews, data were collected from teachers across three occasions over four years. Results reveal that while teaching conceptions remained stable, conceptions of assessment for accountability and certification significantly declined. Qualitative analysis showed that teachers perceived a validity gap between the formative goals of PBL and summative assessment, leading them to delegitimize this aspect of the assessment process to resolve conflict. The findings highlight the importance of aligning active methodologies with both formative and summative assessment processes.

Parallel session 8 3

27 August 2026 14:00 - 15:30

KBG 228

Single Paper

Assessment and Evaluation, Higher Education

Cognitive and Social Processes in Feedback and Self-Assessment

Keywords: artificial intelligence, assessment, assessment design, authentic assessment, feedback tool, formative assessment, higher education, motivation, online measures, peer assessment, peer feedback, reflection

Interest group: SIG 01 - Assessment and Evaluation

Chairperson: Gøril Brataas, Norway

Cognitive Feedback Processing of Online Self-Assessment Instruments and Self-efficacy

Keywords: feedback tool, higher education, motivation, online measures

Presenting Author: Charlotte Van Tricht, University of Antwerp, Belgium; **Co-Author:** Bo Matton, University of Antwerp, Belgium; **Co-Author:** David Gijbels, University of Antwerp, Belgium; **Co-Author:** Vincent Donche, University of Antwerp, Belgium

Online self-assessment instruments with automated, personalized feedback are increasingly prevalent to support students in the transition from secondary to higher education. However, it remains unclear if and how students cognitively process and use this type of feedback. Recently, calls for more empirical evidence have been made. Theoretically, four responses are present: (1) attention allocation, (2) affective processing, (3) interpretation, and (4) decision-making. Notably, self-efficacy is expected to influence feedback processing. This study aims to untangle the interactions between feedback responses and self-efficacy in online feedback reports. Cued recalls based on eye-tracking and gaze replays of 33 students reading the feedback aim to provide an in-depth understanding of this feedback processing. Results indicate that students refer to all four responses, and that more internal than external attributions are made. Students with high self-efficacy mention fewer intentions to change than students with low self-efficacy.

Students' Help-Seeking from AI During Problem-Solving: Role of Cognitive Support and Prior Knowledge

Keywords: artificial intelligence, assessment design, formative assessment, online measures

Presenting Author: Ecenaz Alemdag, IPN Kiel / Europa-Universität Flensburg, Germany; **Co-Author:** Marlit Annalena Lindner, IPN Kiel / Europa-Universität Flensburg, Germany

We explored university students' help-seeking from AI during problem-solving and examined whether cognitive support embedded in problems affected their help-seeking and subsequent problem-solving. In an online between-subjects experiment, 210 students solved probability problems either with or without representative tables (Table+/Table-) as cognitive support, with optional access to an AI-based chatbot. In the post-test, they completed problems without any support. Most participants (80%) used AI at least once, requesting instrumental (51.2%) and executive (31.6%) help. Moreover, tables did not significantly influence help-seeking behavior overall. However, high-prior-knowledge students in the Table+ condition solved problems more efficiently and answered the post-test more accurately than their counterparts in the Table- condition, for whom increased help-seeking was negatively correlated with post-test performance. Finally, non-users of AI for help-seeking achieved higher post-test accuracy than help-seekers. Overall, despite its prevalent use, help-seeking from AI might not be beneficial for learning without sufficient prior knowledge and instructional support.

Student motivations to engage with ePortfolio as authentic assessment: A multi-institutional study

Keywords: authentic assessment, higher education, motivation, reflection

Presenting Author: David Randall, Adelaide University, Australia; **Co-Author:** Christopher Deneen, University of South Australia, Australia; **Co-Author:** Rebecca Marrone, University of South Australia, Australia

The potential of ePortfolios as assessment has long been recognised, but this has not translated into an understanding of how to overcome the well-known issues in successfully implementing them as authentic assessment in higher education. ePortfolios require sustained effort, reflection, and learner agency to be successful, but there is little data exploring students' perceptions of ePortfolios and how they intersect with motivation, a key aspect of self-regulated learning. This paper examines student motivation to engage with an ePortfolio assessment at two universities in the UK and Australia. Data was analysed from semi-structured interviews of 40 undergraduate students through the lens of expectancy-value theory to uncover comparative themes across the institutions. Findings highlight the complex and often competing tensions students experience while engaging with ePortfolio as both assessment and learning. Implications of these findings include mapping out conditions under which ePortfolio assessment enhances or undermines student self-regulation and motivation.

Unpacking Peer Feedback Processes with Think-Aloud Protocols and Epistemic Network Analysis

Keywords: assessment, formative assessment, peer assessment, peer feedback

Presenting Author: David Zamorano, Universidad de Deusto, Spain; **Co-Author:** Ernesto Panadero, Dublin City University, Ireland; **Co-Author:** Maryam Alqassab, Leiden University, Netherlands

Peer feedback is assumed to promote learning, yet the cognitive processes students activate while providing and receiving peer feedback remain largely invisible. Leaving open the question of whether the intended learning cognitive processes we design for are truly activated. This study addresses this question and, at the same time, tackles a methodological challenge: capturing cognitive processes in a way that is practical and meaningful while remaining faithful to their complexity. Forty-two undergraduate students acted as both peer feedback providers and recipients, using standardized peer solutions and feedback. Cued-retrospective think-aloud protocols were coded using a hierarchical coding scheme, and Epistemic Network Analysis (ENA) was used to model the dynamic connections between cognitive processes. Results showed that engagement during both provision and reception of peer feedback was dominated by reading, followed by quick comparisons and surface-level judgments. Criteria-based evaluations and consideration of alternative responses were rare, especially during feedback reception.

Parallel session 8 4

27 August 2026 14:00 - 15:30

MIN 009

Workshop Format

Assessment and Evaluation

Quality of Student Feedback Requests

Keywords: assessment tool, feedback, higher education, quality assurance

Interest group: SIG 01 - Assessment and Evaluation

Exploring the quality of students feedback requests in academic writing with comparative judgement

Keywords: assessment tool, feedback, higher education, quality assurance

Presenting Author: Leonie Brummer, Utrecht University, Netherlands; **Co-Author:** Renske de Kleijn, UMC Utrecht, Netherlands; **Co-Author:** Sanne Geeraerts, Utrecht University, Netherlands; **Co-Author:** Monika Donker, Utrecht University, Netherlands

The act of feedback seeking can allow for more students agency in the learning process. The variety of feedback requests requires a context- and conceptualization-dependent approach to evaluate the quality of a request, extending the scarcity of research on feedback-seeking in educational contexts. By

means of comparative judgement, the context-dependent nature alongside the conceptualization of 'a good request' will be considered. The various backgrounds and expertise of attendees will be utilized. During this workshop, attendees will conduct a comparative judgement for a selection of authentic feedback requests. By discussing and ranking the feedback requests in small groups, a relevant discussion on what is relevant for (the quality of) feedback requests will be evoked (15-20 minutes). The plenary discussion will take on board the input from the small groups (15-20 minutes). This discussion will provide framework from which individuals evaluate the feedback requests and can inform instructional practices regarding feedback seeking.

Parallel session 8 5

27 August 2026 14:00 - 15:30

MIN 011

Roundtable

Assessment and Evaluation

Assessment Practice and Teacher Learning

Keywords: assessment, assessment as learning, evaluation, grading, higher education, primary education, reflection

Interest group: SIG 01 - Assessment and Evaluation

Holistic Scoring in Practice: Towards Usable Guidelines

Keywords: assessment, evaluation, grading, higher education

Presenting Author: Anne Lohuis, Saxion University of Applied Sciences, Netherlands

Holistic scoring is increasingly adopted in higher education as programs move toward integrated and authentic assessment practices. Although guidelines for these assessment practices exist, they are often generic or tailored to more analytic approaches, leaving educators with limited support for holistic scoring. To address this gap, Lohuis et al. (2025a) designed guidelines informed by literature and refined them through feedback from teachers and educational advisors (Lohuis et al., 2025b). These revised guidelines provide a strong foundation for practice. However, important feedback points, such as the need for visual representation and a clear distinction between design and implementation, have not yet been incorporated. In this round table, the current version of the guidelines is presented, and participants will be invited to provide input on the visual design and discuss priorities for follow-up research. The session aims to generate actionable insights to enhance usability by future users and advance holistic scoring practices.

Exploring the use of assessment information to foster teachers' professional learning

Keywords: assessment as learning, evaluation, primary education, reflection

Presenting Author: Linda Jakobs, HAN University of Applied Sciences (UAS), Netherlands; **Co-Author:** Tamara Schilt-Mol, Hogeschool van Arnhem en Nijmegen, Netherlands; **Co-Author:** Helma Oolbekkink-Marchand, HAN University of Applied Sciences (UAS), Netherlands; **Co-Author:** Desirée Joosten-ten Brinke, Maastricht University, Netherlands

While teachers frequently collect assessment information, leveraging this information for their own professional learning, remains underutilized. This study examines how primary school teachers use such information to foster their professional growth, focusing on an approach that emphasizes personal responsibility and reflective practice. Based on exploratory research among 20 teachers and three months of audio diaries and supplementary interviews, this roundtable invites the SIG 1 community to participate in a pre-analytical collaborative meaning-making process. The session takes place before the formal analysis and consists of an introduction to the theoretical framework, followed by a showcase of exemplary narrative fragments based on teachers' pseudonymized reflections. Participants engage in open dialogue and use their expertise to identify patterns and tensions in the excerpts. This approach aims to sharpen the analytical lens of the research so that the subsequent formal analysis authentically reflects the yields from everyday practice.