

ANNUAL WORKSHOP

13th May 2026

Henderson Lecture Theatre,
XFi Building, University of Exeter

Are we assessing the right skills?

Collaboration, learning, and work in a changing world

The 2026 workshop brings together Economics educators, researchers, and industry professionals to explore how assessment can better reflect the skills students need in a rapidly changing world. Discussions will focus on curriculum design, learning activities, and assessment strategies that develop teamwork, adaptability, critical thinking, and real-world problem solving.

Programme

10.00 – 10.45	Registration and Coffee
10.45 – 11.00	Introduction to the workshop – <i>Loukas Balafoutas (Head of the Department, Exeter)</i> SWEETS Committee Welcome
11.00 – 12.30	Panel: Are we assessing the right skills? Collaboration, learning, and work in a changing world <i>Chair: Arpita Ghosh (Exeter)</i> Panel members: Parama Chaudhury (UCL), Aakriti Mathur (Bank of England), Helena Fornwagner (Exeter)
12.30 – 13.30	Lunch

13.30 – 14.30	<p>Presentation Session 1</p> <p><i>Chair: Mike Peacey (Bristol)</i></p> <p>From Derivation to Applied Reasoning in Final-Year Macroeconomics <i>Magdaly Okolo (Bath)</i></p> <p>Humanising Assessment through Reflective Journals in Economics <i>Anne Pass (Reading)</i></p> <p>Chain of Thought Economics: Critical Thinking Toolkit for Economics Learners <i>Lily Samkharadze and Yue Yin (Bristol)</i></p> <p>When ChatGPT Takes the Midterm: Student Learning Through Critical AI Engagement <i>Hiren Nisar (Exeter)</i></p>
14.30 – 14.50	Coffee Break
14.50 – 15.50	<p>Presentation Session 2</p> <p><i>Chair: Anastasia Papadopoulou (Bristol)</i></p> <p><i>Confirmed presenters:</i></p> <p>Do Take-Home Assessments Overstate Engagement and Learning in the Age of Generative AI? <i>Robert Riegler and Tim Burnett (Aston)</i></p> <p>Incentivising Engagement: The Impact of Low-Stakes Grading on Formative Assessment Use and Academic Performance <i>Fareena Noor Malhi (Bath)</i></p> <p>Assessment Design and AI Literacy <i>Ahmed Pirzada (Bristol)</i></p> <p>Using Video Retention Analytics to Study the Transition from School-Based Learning to University Economics <i>Rubens Morita (Exeter)</i></p>
15.50 – 16.00	Final Remarks and Close
16.30-	Pub (self-paid)

Poster Session Presentations (presenters underlined):

From Silence to Engagement: Understanding Cultural and Pedagogical Influences on International Students' Participation in Western Classrooms

Yunzi He (Bristol)

The Gender Gap in Multiple Choice Question versus Constructed Response Question Examinations

Imran Shah (Bath)

Assessing Economic Thinking, Not AI Outputs.

Aliya Kenjegalieva (Bath)

Skill building starts in induction week: lessons from an intervention designed to equip economics students for their academic and professional lives

Olha Matvietts and Chloe Tiennot (Exeter)

From Oil Shocks to Inflation Spirals: An Interactive Mentimeter Simulation (Teaching Adaptive Expectations and Monetary Policy)

Samuel Odewumni, Debbie Du Preez, Mario Du Preez, and Rubens Morita (Exeter)

Teaching Akerlof's Market for Lemons Through a Large-Class Simulation: An Economics Education Case Study

Jingya Zeng and Rubens Morita (Exeter)

Abstracts of the presentations:

1. From Derivation to Applied Reasoning in Final-Year Macroeconomics

I examine how final-year macroeconomics teaching can be redesigned to strengthen students' applied research and disciplinary reasoning without abandoning technical rigour. The study is on a year-long undergraduate macroeconomics unit in which Semester 1 is taught through a topic-based sequence of lectures, online workshops, and seminars, and assessed through coursework before the exam in Semester 2. The project begins from a familiar tension in advanced macroeconomics teaching. Structural macroeconomics is attractive to students who are confident in mathematics, creating a preference for solving technical problems over discussing interpretation and critique. Yet macroeconomics also requires students to interpret economic events, connect theory to evidence and recognise model limitations. Preliminary evidence showed that students performed better on the research-led and applied component than on the derivation-heavy component, indicating that scenario-based questions may provide a productive way to combine technical competence with applied reasoning. Assessment evidence also points to a more specific pedagogical issue. In the coursework, students were required to research a real-world event, map the events to a model, and reflect on the model shortcomings. Performance was strongest on the research-led component and weaker as students moved into model mapping and extensions, suggesting that the central challenge might be the transition from reading to model-based explanation and critique. The proposed intervention redesigns the workshop sequence and the coursework to address this gap directly. The project asks whether a more structured scaffolding can improve students' ability to move from evidence to model, from model to critique, and from critique to policy reasoning.

2. Humanising Assessment through Reflective Journals in Economics

Inspired by the need to future-proof learning outcomes and equip students with "soft" skills, I replaced traditional essay-style coursework with a reflective learning journal. This assessment was trialled across two distinct settings: a large undergraduate Econometrics module and a small postgraduate Macroeconomics module. Over six weeks, students made regular entries in an electronic journal, shifting the evaluative focus from the final research output to the research process itself. My aim was to foster genuine engagement and grant students greater agency in their learning journey. Despite the logistical challenges of scaling this for larger cohorts, the pedagogical benefits—particularly in terms of authentic student voice—outweighed the drawbacks.

3. Chain of Thought Economics: Critical Thinking Toolkit for Economics Learners

We have developed and are currently trialling an innovative teaching approach, the Chain of Thought (CoT) Economics framework, which aims to enhance economics learners' problem solving and critical thinking skills. Inspired by the Chain of Thought prompting technique used in machine learning, the framework guides students through a structured, step by step reasoning process using an infographic that illustrates how to break down and analyse an economics problem. To put this structured thinking into practice, we introduce CoT style prompts that translate the framework into clear, sequential steps students can apply in any problem-solving setting, including seminars and assessments. These prompts are easily customisable across a wide range of economics questions. The framework also provides an important bridge to emerging AI supported learning practices. As AI becomes embedded in everyday study and work, prompt writing is rapidly becoming a key employability skill. However, as is now well recognised, the value students gain from AI depends heavily on how they engage with it, yet further research is needed to develop practical and teachable methods for supporting effective engagement. The CoT Economics framework can become both a cognitive support and a communication tool, helping students engage with AI more effectively and more responsibly. By modelling the step-by-step logic of large language models, the framework helps students set clear objectives, organise information and communicate instructions precisely. Over time, as students learn to think and prompt in more organised steps, they begin to use AI in ways that strengthen rather than weaken their critical thinking.

4. When ChatGPT Takes the Midterm: Student Learning Through Critical AI Engagement

The rapid adoption of GenAI presents a dual challenge for economics educators: while tools like ChatGPT offer efficiency, they risk fostering cognitive atrophy and undermining independent analytical skills. In a landscape where more than 95% of students self-reporting using ChatGPT, the pedagogical focus must shift from banning GenAI to "grading the thinking" behind its use. As a result, we designed an AI-critical assessment to move students from passive consumption to active auditing. In this multi-stage coursework task, students critiqued an AI-generated economics essay before producing their own refined work and a reflective piece on the process. Our findings from 40 participants (~20 percent response rate) reveal a distinct "Surface vs. Substance" model. While students recognized AI's strengths in structure (69% agreement) and relevance (70%), they identified failures in substantive quality, with 46% noting a lack of appropriate evidence and 40% citing a deficiency in analytical depth. Engaging in the "revised essay" seemed to be effective: 82% of students reported being confident in assessing AI content for future academic work, and 69% reported gaining a deeper understanding of AI's limitations. Despite these gains, students highlighted some friction points, with 62.5% requesting more explicit pre-assignment scaffolding to manage the complexity of the task. We conclude with four strategic recommendations for economics instructors.

5. Do Take-Home Assessments Overstate Engagement and Learning in the Age of Generative AI?

This study examines the use of short, confirmatory, in-person assessments to estimate the extent of generative AI (genAI) use in remote assessments, and its capacity to drive meaningful student engagement with assignments. Rapid advances in genAI tools (e.g. ChatGPT, Gemini, Copilot) have raised concerns about the validity of traditional take-home assessments, such as essays and projects, and their effectiveness in supporting learning and evidencing achievement of intended learning outcomes. In econometrics, genAI can replicate core disciplinary skills, including justifying methodological choices, writing code, interpreting output, and producing mathematical proofs. While this undermines the integrity of assessment, it may also reduce student learning and engagement with the subject. We present a case study of several 10–12-minute in-class tests administered shortly after submission of take-home assignments across several econometrics modules. Tests were based on the take-home content, requiring students to interpret identical results or explain lines of code. The study investigates (a) how take-home correlates with the in-class test performance, b) how micro-tests affect the assessment mark distribution, and (c) provides preliminary evidence on how repeated micro-testing affects engagement in subsequent modules. Early results highlight limited correlation between the two assessment components and substantial performance differences. Many students struggled to demonstrate, in-person, the knowledge and understanding suggested through their recent take-home submissions, suggesting that take-home assessments may overstate learning outcomes. Overall, while micro-testing highlights disparities between remote work and in-person understanding, it may also offer a constructive mechanism to support engagement and preserve the use of take-home assessments.

6. Incentivising Engagement: The Impact of Low-Stakes Grading on Formative Assessment Use and Academic Performance

This study examines the impact of attaching low-stakes grade weighting to otherwise ungraded formative assessments on student engagement and academic performance in higher education. In the UK context, a sector-wide shift toward reducing the number of assessments to alleviate student workload has inadvertently limited opportunities for students to receive and act on constructive feedback prior to high-stakes summative assessments. Even where formative tasks are provided, student engagement with these opportunities is often low. To address this, we evaluate a pedagogical intervention in which previously ungraded formative assessments are assigned a 10% weight based on completion and engagement. Using administrative data from a higher education unit, we analyse the association between this intervention and student engagement, as well as performance in the summative assessment and overall unit outcomes. Engagement is measured through submission and participation metrics, while performance is captured through summative scores and final grades. The analysis controls for prior academic achievement (first semester average), age, and gender. We find that introducing low-stakes grading is associated with increased engagement with formative tasks and improved academic performance. These findings suggest that modest incentive structures can enhance student interaction with feedback without substantially increasing assessment burden, offering a practical approach to maintaining effective feedback loops in a reduced-assessment environment.

7. Assessment Design and AI Literacy

To what extent does engagement with generative AI tools improve learning outcomes in undergraduate economics, and does assessment design matter? This paper documents a striking reversal in the association between AI engagement and perceived usefulness across two successive cohorts taking the same course at the University of Bristol. In the cohort where

relevant tasks were unassessed and less scaffolded, higher AI engagement is associated with lower perceived productivity. In contrast, in the cohort where a similar task was formally assessed and scaffolded, we find that higher AI engagement predicts higher perceived employability and learning effectiveness. These results suggest that assessment design determines whether AI functions as a shortcut or a productive learning tool, with direct implications for how universities should respond to the rise of generative AI.

8. Using Video Retention Analytics to Study the Transition from School-Based Learning to University Economics

A persistent challenge in undergraduate economics education is supporting students in the transition from secondary-school learning, which often emphasises memorisation, to university-level study, where analytical reasoning and problem-solving are central. This paper uses video learning analytics as a low-cost pedagogical diagnostic to identify which forms of instructional content attract disproportionate replay and interprets these patterns in light of the UK transition from A-level to undergraduate economics. Drawing on audience retention data from eight YouTube-hosted instructional videos (N = 800 per-position observations), we document systematic patterns of repeated viewing concentrated around segments containing explicit concept definitions, labelled diagrams, optimality conditions, and structured summary cards. We formalise this pattern using two indices: the Concept Attention Index (CAI) and the Pedagogical Attention Index (PAI), and estimate OLS regressions with video fixed effects and richer multi-category content coding. The grand pooled CAI is 1.267; audience retention during concept-coded segments is, on average, 26.7 per cent higher than elsewhere in the same video. An integrated scene-sequence and subtitle analysis across all eight videos, combining timestamped visual content with verbatim transcript data, produces one substantive window correction (Video~5, shifted to 89-91%); the scene-aware grand-pooled CAI is approximately 1.206, and all fixed-effects estimates remain highly significant ($p < 0.001$). Approximately 65-69% of statistically significant attention shocks coincide with scene transitions, providing evidence that visual-card onset is the proximate trigger of concentrated replay. A permutation test confirms the observed CAI lies more than 12 standard deviations above 1000 random placebo draws. These patterns are consistent with classification-seeking behaviour inherited from secondary education, in which students disproportionately return to content that can be labelled, recorded, and retrieved; what we term classification-friendly pedagogical anchors. The approach is low-cost, scalable, and transferable across institutions and disciplines.

[Abstracts of the poster presentations:](#)

1. From Silence to Engagement: Understanding Cultural and Pedagogical Influences on International Students' Participation in Western Classrooms

International students frequently remain silent in higher education classrooms, particularly in first-year technical courses where demanding content intersects with new linguistic, cultural, and pedagogical expectations. To address this issue, the study examines the motivations underlying students' silence and how they negotiate participation norms within Western-style academic environments. Focusing on Phase 1 of a mixed-methods project, this paper draws on surveys mapping perceived participation barriers and semi-structured interviews exploring students' lived classroom experiences in economics and mathematical modules. Quantitative patterns are summarised descriptively, while qualitative accounts are analysed thematically to uncover how students weigh linguistic, cultural, and disciplinary risks when deciding whether to speak. Although the empirical sites are geographically specific in China (SWJTU-Leeds Joint School) and the UK (School of Economics,

University of Bristol), the observed dynamics reflect broader East Asian communication patterns common across many Western-style classrooms, making the issues examined highly relevant to economics programmes with international cohorts. Preliminary findings indicate that silence among international students is often purposeful and strategic, aligning with insights from Kajan et al. (2025), who argue that Chinese students' silence reflects culturally informed self-protection rather than disengagement. Participants across both sites reported tensions between wanting to contribute and worrying about misinterpreting technical concepts, misusing specialised terminology, or breaching implicit participation norms. Early thematic patterns also suggest that local pedagogical cultures shape students' judgements about when and whether it feels safe to speak, echoing wider East Asian participation tendencies documented in Western educational settings. These insights offer a grounded understanding of the motivations and constraints informing classroom silence within economics education and clarify how cultural expectations and institutional contexts interact to shape participation. Phase 1 thus provides a strong foundation for Phase 2, which will develop and evaluate context-sensitive pedagogical strategies aimed at fostering more inclusive engagement across diverse international learning environments.

2. The Gender Gap in Multiple Choice Question versus Constructed Response Question Examinations

Success in university examinations can be crucial for students' future employment outcomes. Exam format influences performance asymmetrically across genders. This paper thus analyses the gender gap in student performance across two assessment formats: multiple-choice questions (MCQs) and constructed response questions (CRQs). Using data from undergraduate and postgraduate modules at a large UK university across the 2020/1 to 2022/3 academic years, we examine whether gender grade discrepancies arise in each exam format. We leverage variation within academic years, modules and degree programmes, comparing grades in online MCQ and in-person CRQ assessments. Both parametric linear and non-parametric local constant regressions yield aligned inference: we do not find that women perform differently from men in online MCQ tests. However, women significantly outperform men in CRQ exams by 0.2 standard deviations. There is no heterogeneity in this relationship across ethnicities, however this overperformance is more pronounced at the undergraduate level. Quantile regressions further show this outperformance stems from lower grade quantiles. A 2023/24 robustness check using an in-person exam containing both MCQs and CRQs confirms the observed gender gap in baseline results stems from question format rather than mode of delivery. Our findings underscore the need consider exam format to equitably accommodate diverse strengths.

3. Assessing Economic Thinking, Not AI Outputs.

The rapid growth of GenAI tools such as ChatGPT and Copilot is reshaping not only how students learn, but also what it means to demonstrate knowledge and skill in higher education. This raises a critical question: are we still assessing the right skills in a world where collaboration increasingly includes AI? Focusing on the PGT programmes, this project explores how assessment practices can evolve to better reflect contemporary forms of learning, collaboration, and work. While GenAI offers new opportunities to support learning, it also challenges traditional notions of authorship, academic integrity, and individual performance. Using a mixed-methods approach, the study investigates how students engage with GenAI in their studies and assessment preparation, drawing on survey data to examine patterns of use, perceived learning benefits, concerns around fairness, and impacts on skill development. Building on these insights, the project examines how coursework can be redesigned to prioritise the assessment of higher-order skills, including economic reasoning, critical thinking, and reflective engagement with AI-assisted processes. Pilot assessment redesigns move beyond evaluating final outputs alone, instead emphasising transparency, process, and the ability to critically interpret and build upon AI-generated content.

4. Skill building starts in induction week: lessons from an intervention designed to equip economics students for their academic and professional lives

As economics education increasingly prioritises the development of transferable skills such as

teamwork, adaptability, critical thinking, and real-world problem solving, there is growing value in embedding these competencies from the very start of the student journey. This paper presents a practice-based case study of the design of induction week activities for undergraduate economics students, focusing on how early-stage interventions can support both engagement and skills development. Drawing on transition pedagogy and belonging literature, we reflect on a series of induction activities implemented in 2024 and subsequently refined in 2025. These include business school-level sessions and department-led initiatives. Our experience highlights how different formats can be used not only to deliver information but also to develop key competencies actively. In particular, collaborative, task-based activities—such as the Digital Escape Room challenge—foster teamwork, adaptability, and problem-solving by requiring students to work together, navigate uncertainty, and engage actively with real institutional contexts. At the same time, structured in-class interventions, such as short, low-stakes community-building exercises embedded within the Economics Culture Induction, support the development of critical engagement, build confidence, and increase students' willingness to participate in academic discussions. A central insight from this work is the importance of embedding the development of social and collaborative skills throughout induction week, using a combination of interactive and structured activities. The paper contributes to the SWEETS agenda by offering a practical framework for integrating core competencies into induction design, demonstrating how skills development can begin from day one of economics education.

5. From Oil Shocks to Inflation Spirals: An Interactive Mentimeter Simulation (Teaching Adaptive Expectations and Monetary Policy)

This paper presents a classroom simulation designed to help introductory economics students understand how supply shocks, adaptive expectations, and central bank intervention interact to generate inflation persistence. The activity was implemented in a first-year undergraduate economics module at the University of Exeter using Mentimeter, with approximately 16–22 students responding at each stage. Students were guided through a sequence of questions in which they acted as workers negotiating wages under adaptive expectations. The simulation combined The Phillips Curve, the WS-PS framework, and the Aggregate Demand model show how a positive bargaining gap generated by an oil shock can trigger an inflation spiral. Accuracy rates across seven stages never fell below 90%, and students correctly identified that closing the bargaining gap stabilises inflation without reducing it, and that a sustained negative gap is required to bring inflation back to target — the distinction between stabilisation and disinflation that students most commonly struggle with. These results suggest that the sequential, interactive format helped students internalise the recursive logic of adaptive expectations, though correct responses during a guided classroom exercise do not necessarily demonstrate independent conceptual understanding. The paper also proposes an extension that introduces a credibility parameter blending the inflation target with lagged inflation, allowing instructors to contrast adaptive and rational expectations within the same simulation.

6. Teaching Akerlof's Market for Lemons Through a Large-Class Simulation: An Economics Education Case Study

Teaching adverse selection effectively remains a challenge in introductory economics, as students often struggle to internalise the behavioural mechanisms underlying The Market for Lemons. This paper presents a scalable, technology-enhanced classroom simulation designed to make the concept intuitive through direct student participation. Implemented with approximately 80 first-year undergraduate students at the University of Exeter, the activity combines role-play under asymmetric information with real-time data collection using Mentimeter. Students are assigned roles as buyers or sellers of high- and low-quality goods and make decisions under uncertainty. The resulting data closely replicate theoretical predictions: buyers' willingness to pay converges to the expected value under uncertainty, while high-quality sellers disproportionately exit the market, generating a clear adverse-selection spiral. The introduction of institutional remedies—such as warranties and pre-purchase inspections—produces measurable shifts in behaviour, illustrating how information frictions can be mitigated.