



A reflective conversation on assessment in the age of generative AI using the traffic light framework at the University of Leeds

Chrissi Nerantzi¹, John Palfreyman¹ & Maria Pavlopoulou¹

Abstract

This paper presents a reflective conversation on assessment in the age of Generative AI (AI from now on), framed by the University of Leeds' traffic light system for AI use in assessment. The dialogue brings together three perspectives: two postgraduate students, one of whom is also an academic, and a further academic, exploring how the system's Red, Amber, and Green categories shape practice and perceptions. While the framework offers clarity, our discussion reveals its interpretive complexity, particularly around Amber, and questions its exclusive focus on assessment rather than learning. We argue for transparency, modelling, and process-oriented approaches that integrate AI use throughout the learning journey. Experiences with Amber and Green assignments illustrate opportunities for ethical, creative, and authentic engagement with AI, alongside challenges such as maintaining critical voice and academic integrity. We conclude that while the traffic light system served an initial purpose, it now appears outdated. A more flexible framework is needed. One that connects learning and assessment, acknowledges nuance, and positions human judgement and collaboration at the heart of AI-enabled education.

Keywords

genAI, AI, reflective conversation, assessment, traffic light system

¹ University of Leeds, Leeds, UK

Corresponding Author: Chrissi Nerantzi, Professor in Creative and Open Education, School of Education, University of Leeds, Leeds, LS2 9JT, UK

Email: C.Nerantzi@leeds.ac.uk

Context

On the 16th October 2025, during the Assessment in the Age of Generative AI event hosted by the [Media and Learning Association](#), participants explored ethical AI use in education. The event featured guidelines from the [TaLAI Erasmus+ project](#) and other European initiatives. As part of this, three individuals from the University of Leeds came together as part of a panel discussion to reflect on their experiences with assessment and generative AI (Nerantzi, 2025; Rajagopal, 2025):

Maria: A secondary school teacher of Greek and MA Digital Education student in the School of Education supervised by Chrissi.

Chrissi: Maria's tutor and dissertation supervisor; also, a student on the Leeds Executive Leadership Apprenticeship (LELA) course.

John: Chrissi's tutor on the LELA course and advocate for responsible AI use in the Business School.

The conversation of these individuals, structured around the University of Leeds' (2025) *traffic light system* for AI use in assessment based on the panel discussion (Nerantzi et al., 2025), is discussed in this paper by these individuals as co-authors. The article adopts a dialogic and reflective approach. It is organised around a reconstructed conversation that took place during the panel and presented in three parts that broadly follow the 'Red', 'Amber', and 'Green' categories of the University of Leeds' traffic light system for AI use in assessment. The conversation has been curated and edited by the authors to foreground shared questions, points of tension, and moments of sense-making that emerged through dialogue.

The conversational sections are followed by a collective synthesis, presented in the section *What we have learnt thanks to this conversation*. This section represents a reflective and relational interpretation developed by the authors through revisiting the conversation, their dialogic notes, and their shared experiences as educators and learners. The analysis is embedded, interpretive, and oriented towards surfacing insights that may resonate with others navigating assessment and learning in the age of generative AI. Readers are invited to engage with the conversation not as a linear argument, but as a situated scholarly dialogue that opens up questions, possibilities, and ongoing inquiry.

Part 1: understanding the traffic light system or clarity vs. complexity

John: The AI traffic light framework for assessment has been in place since 2023 and is used across all faculties at the University of Leeds. It provides a framework for how AI can be used in assessment. Module tutors define the colour of the assessment based on the traffic lights specifying if the particular assessment is Red, Amber or Green. As of August 2025, only about half of UK higher education institutions have such a framework.

If the traffic light coding is Red it means AI usage is banned in a particular assignment. If it's Amber, AI usage is allowed, within rules set by module leader/tutor and each tutor may have different rules. If it's Green, students are taught how to use AI responsibly and encouraged to use AI tools in the assessment to support their learning.

In the Business School, for example, last academic year (2024/25), around 85% of taught modules across the school, and that included undergraduate and postgraduate modules, had Red assignments. The University of Leeds goal is to move all assessments to Green as soon as possible. I'm supporting colleagues to develop undergraduate and postgraduate modules to Green.

Chrissi: In what way is the traffic light system useful to you, Maria, as a postgraduate student on our MA in Digital Education?

Maria: From my point of view, as a student, the University of Leeds and how I have experienced the traffic light system, I can say that it has been useful as it has helped me reduce uncertainty around academic integrity. It made me think more profoundly about how I use AI tools, not just whether I can. It seems to mirror real-world expectations of ethical AI use and fosters students' digital literacy. As a teacher, I can see it being a model to help my own students understand issues around responsible AI use when they get started. Overall, it builds a shared understanding perhaps between academic staff and students. I have been wondering if it will remain useful going forward. What about you John?

John: The traffic lights have been around for a while now, so I think it's ready for a rework. It's better than having nothing, and good to have when we were all new to AI but it's also challenged. Many students and staff don't fully understand it. The traffic light framework is a bit complex. And more importantly, assessments themselves need to evolve. The classic essay isn't ideal anymore, especially if it can be written in seconds by AI. What is your perspective Chrissi?

Chrissi: The traffic lights are a framework that appears simple and comprehensive, but it isn't. There is no point in having Red if we all need to learn how to use AI. Amber is vague and interpreted very differently. It is a very broad grey area. Green is clear: go for it. But what do we really mean by this? Definitely not copy and paste.

Also, the traffic lights have a focus on assessment. It makes me wonder, what about learning with AI? If an assignment is Red, can I as a student still use AI to learn? Do I as a student need permission to use AI in my learning? The focus of the traffic light framework seems to be on assessment, but if students learn how to use AI for learning, wouldn't that be a useful step forward? I am teaching on a range of modules and we invite students to use AI in their learning, not to replace learning or critical engagement. If we allow students to use it in their learning and then tell them not to in their assignments, is this right? What are we assessing and what is the purpose of assessment?

I am a student and an educator at the moment. Emerging research suggests that as educators we seem to blame students for cheating and then spend a lot of effort in policing the use of AI (Gorichanaz, 2023). But we also use AI ourselves and don't always tell them. What we really need is transparency, to learn how to use it to support learning and to reimagine assessment.

What is the purpose of assessment? What is its value? If our assessment criteria align with what a machine can do in seconds, what's the point? I think it would be better to focus on: human judgement, decision-making and critical reflection.

John, from a teaching perspective, how can tutors make sure students actually understand and apply the policy correctly?

John: Great question! I increasingly feel that the traffic light system is often misunderstood. This is made worse by the hype and misinformation around AI. It needs simplification. Maybe an infographic, podcast, or animated video would help. But I am not sure this is enough or even useful where we are now.

The learning climate is key. Students need to feel safe asking questions if they don't understand whether it's about the policy or how to use AI in their learning without fear of looking silly in front of their peers.

Part 2: Amber assignment experiences or shifting focus from assessment to learning

Chrissi: Let's have a look at Amber assignment specifically. On our MA in Digital Education modules, we use the Amber category for assignments. How has this worked for you Maria?

Maria: The Amber category has worked really well for me as it encourages AI use in a reflective, supportive way rather than as a text generator. In my *Designing Digital Education* module, I used ChatGPT and MojoAI to help shape my *Greek GenAI Academy for Educators* project, which was later featured in the University's student-led open book *Learning with AI* (University of Leeds, 2025). The procedure helped me think more critically and creatively, linking design decisions to pedagogical issues while keeping transparency about how I used AI. It's been a great method to boost my digital literacy and confidence in using these tools ethically.

Being transparent about AI use has really changed my mentality towards learning. It rendered me more responsible of my learning process, as I have to think why I am using AI, if and how I really need it to help me and what its limitations are. Thus, I feel pushed to keep ownership of my own work, by always staying transparent for it. My tutors' transparency – for instance, when using AI to design course material or even assessments – has fostered trust and formed an open space of sharing ideas towards AI's practical educational value on HE education.

Moreover, reflection plays a key role, as by thinking of my own decision-making regarding AI's educational use, I understand a lot for me as a learner. I can judge when I over-rely on it or how I could gain more confidence on its proper ethical usage. Back in 2024, in Open Education Week Live (GenAI Literacy Student-led Panel), I presented my three-step pathway towards GenAI literacy. Reflection was an important step, as without it, experimentation alone remains superficial. However, sometimes AI output can be proven inaccurate or just unhelpful. Although stressful at the beginning, these moments are the most valuable, forcing me to rethink my learning approach and engage critically with the learning material and/or task. Hence, a learning opportunity emerges through what seems a failure. Working with AI for your learning as a HE student is not about getting the best outputs, but rather about asking the smartest questions and forming better judgement, critical awareness and resilience as a learner.

Chrissi: It is wonderful to hear Maria that being transparent about use of AI as an educator and student has made a difference to you and you recognise the value of reflection and the importance of questions.

Going back to the traffic light system we have, my interpretation of Amber is that students can use AI in any way that supports their learning, but they don't have to. It is very

flexible and could even be seen as Green by some colleagues. In our *Education in a Digital Society* module, we use Amber too. It is a Problem-Based Learning (PBL) approach, where students engage with AI-enabled curriculum design and learn how to use AI responsibly (Nerantzi, 2025).

John, you are my tutor on the course I am doing and a colleague. From a tutor's point of view, what kind of guidance helps students use AI responsibly in Amber assignments?

John: I am not a big fan of Amber, but I can see its usefulness. Clarity is key; students need clear, practical, and understandable examples of where AI can be used and where it can't.

Chrissi: First of all, just mentioning AI only in the context of assessment is very limiting, I feel. It ignores how students can use and are using AI in their learning.

Modelling always helps, I know this from my life as an academic developer and teacher educator. Not just asking students to experiment with AI, but doing it together. Building it into the learning process. Learn from mistakes. Collaborate and co-create.

Maria: In fact, I experienced working with AI as a form of collaboration and co-creation in my MA studies. Even mistakes were a big part of collaboration with tutors, like you Chrissi, as well as with peer students. At a HE education level, using AI engages us all in a fruitful dialogue, that constantly challenges us to test ideas, refine and foster skills, essential to learning in general.

Chrissi: Documenting AI use can also be really useful. It is part of the learning process and reflection. Too often we focus exclusively on the output while in assignments, students should evidence how AI was used and what it enabled. Demonstrating its value for learning is key.

John: I've incorporated reflective practice into the assessment to capture this learning process, asking students to reflect on their GenAI usage, explaining what they did, what was good, what did not work so well and what they would do differently next time.

Part 3: Green assignment experiences and what comes next or reimagining assessment

Chrissi: For tutors, what does good assessment design look like in the Green category? How can you assess both the students' knowledge and their AI use?

John: I started using AI Green in assessments in late 2024. Since then, I have led seven postgraduate modules that are Green and supported colleagues in doing the same, including building flagship undergraduate modules.

I ensure assessments are real-world and case study-oriented and always include a reflective element on AI usage, especially learning from when things go wrong.

Most students embrace AI and use it to boost critical thinking and submission quality. I get a lot of positive feedback students say it's a breath of fresh air and that it doesn't feel like cheating. My industry contacts are big fans of this approach too.

Chrissi: Assessment often focuses on the final product – a polished piece of work. But learning is messy, and the process is often forgotten. I like us to focus more on *assessment while learning*, assessment as learning. That requires investing in resources, learning communities, and trust.

As educators we need to model learning, to be learners ourselves and be open about it. A good assessment diet has purpose and sparks curiosity. Students should want to engage with it because it excites them, and they see its value.

If the assessment can have a life beyond the module, even better. Can we design assignments that are useful for students and relate to their personal or professional contexts or even for a wider community? Open educational resources come to mind. While we say the process matters, there is also something special when we create something of value.

Maria, what potential challenges would arise if you had a Green-light assignment?

Maria: A Green-light assignment would be really exciting but also quite demanding, since AI would play a major role in the final outcome. The biggest challenge would be maintaining my own critical voice and not permitting AI to take over the creative part of the process. It would also mean staying very transparent and careful about important issues, like bias, authorship, and accuracy. Balancing creativity with academic rigour could be a challenge, but it would also be an amazing chance to explore how academic students and AI can genuinely collaborate in learning.

Chrissi, how did you use AI in the Green assignment for the *Strategy* module?

Chrissi: John, that was in your *Strategies for Business Leaders* module on the LELA course. As a student, I felt trusted. Learning and assessment were entangled practices. I was encouraged to experiment and learn freely.

I didn't outsource thinking. I wanted to maximise learning. I fully documented my AI use. AI didn't do the thinking for me. This was important. I edited any outputs based on my own ideas and text. I also avoided using AI until I had read widely about what I wanted to learn and write about.

In that Green assignment, I used AI in a range of ways, which I documented in a dedicated appendix. The AI technologies I used were Secure Copilot, Napkin AI, and NotebookLM and the tasks were:

- Seeking clarification
- Summarising
- Reducing word count
- Applying a matrix from the literature to a specific strategy to identify patterns
- Visualising some of my ideas
- Generating a podcast interview based on the assignment

These strategies supported my learning in complementary ways and were deliberately used as part of an iterative sense-making process rather than as efficiency shortcuts. Seeking clarification and summarising helped me check my understanding of key concepts and arguments, while reducing word count supported critical decision-making about what really mattered in relation to the assessment criteria. Applying a matrix from the literature and visualising ideas enabled me to explore patterns, relationships, and implications more deeply, strengthening conceptual understanding. Generating a podcast

interview based on the assignment formed part of this reflective process by requiring me to listen to my thinking and further reflect on my thoughts. These strategies helped me refine my critical voice and consolidate learning rather than replace engagement with reading, writing, or analysis.

The approach I took, I mean how I worked with AI, resembled an alternative uses creativity test (Guildford, 1971) which I often use in our creativity module where I ask my students to come up with as many different and novel uses of an everyday object, for example a toothbrush or an envelope. This enabled me to explore AI affordances. What else can I do with AI to support my learning? I think my curiosity was driving me to experiment, explore and discover. It wasn't a challenge. It was liberating.

What we have learned thanks to this conversation

We have found our reflective conversation valuable and are grateful to the Media and Learning Association for this opportunity (Nerantzi et al., 2025; Rajagopal, 2025). In revisiting the dialogue across the three parts of the paper and reflecting collectively on our dialogic notes and lived experiences as educators and learners, we identified a number of shared insights that cut across roles, disciplines, and moments in the conversation. The synthesis below does not represent a linear summary of individual contributions, but a collective process of sense-making that foregrounds recurring questions, points of agreement, and productive tensions that surfaced through the dialogue.

1. Clarity vs. complexity

A central thread running through the conversation was the tension between the apparent simplicity of the traffic light system and its lived complexity in practice. Across the discussion of Red, Amber, and Green assignments, all three voices returned to the ambiguity of interpretation, particularly in relation to Amber, which emerged repeatedly as a space of uncertainty and uneven practice. At the same time, concerns about Red assessments recurred at different points in the conversation, raising questions about the contradiction between banning AI use in assessment and the growing expectation that students develop ethical and responsible AI practices. While these concerns were articulated from different positionalities, together they point to a broader, shared unease with frameworks that appear clear on the surface but lack nuance in application.

2. Shifting focus from assessment to learning

Another insight that emerged across the conversation was a shared recognition that the traffic light system is framed almost exclusively around assessment, with insufficient attention to how students learn with AI. This surfaced in reflections on students' use of AI for sense-making, experimentation, and confidence-building, as well as in discussions about modelling and transparency. Rather than positioning AI solely as something to be controlled at the point of assessment, the conversation repeatedly returned to the importance of making learning processes visible, encouraging dialogue about AI use, and creating spaces where both students and educators can learn together. Seen collectively, these contributions highlight a shift from policing AI use towards supporting learning relationships grounded in trust, openness, and care.

3. Reimagining assessment

The final insight focuses around the need to rethink assessment design itself in an AI-enabled context. Drawing on examples discussed in relation to Amber and Green assignments, the conversation surfaced a shared desire to move away from an exclusive focus on polished outputs and towards assessments that value process, reflection, and authenticity. Across roles and experiences, there was agreement that assessment can be both rigorous and relational, and that process-oriented designs can support deeper learning, critical engagement, and ethical AI use. This reimagining also brought forward questions of purpose and value, with contributors emphasising the potential of assessments that connect to students' personal, professional, or wider societal contexts and have relevance beyond the immediate requirements of a module.

From our conversation and the above synthesis, we can see that it may be the right time to move away from the traffic light system and explore alternative approaches that more closely align with the needs of educators and students in AI-enabled education. In this respect, it is useful to respond to Amiel et al. (2025), who caution that “we must avoid the trap of seeing technological fixes (AI included) as complete solutions to complex educational challenges”. We share this concern and position our work deliberately against technologically deterministic narratives. In this study, technology does not act as a driver or solution in itself. Instead, it is approached as a relational element within pedagogical designs grounded in trust, care, and inquiry.

By foregrounding sound pedagogical principles and creating opportunities that place humans at the heart of exploration, dialogue, and learning, we seek to demonstrate how educational practices can remain responsive to individual learners, local communities, and global challenges. From this perspective, AI is not framed as an answer to complexity, but as a catalyst for reflective practice, critical questioning, and shared meaning-making; therefore, AI use is learning-oriented, not tool-oriented. Such an orientation aligns with values that privilege openness, relationality, and ethical responsibility, recognising education as a situated and collective endeavour with implications not only for learners but also for wider society and the planet.

What are some of the educational challenges we experience in a VUCA¹ world (Taskan et al., 2022)? And if AI is already part of our everyday lives and this is how students perceive it (Attewell, 2025), what does this mean for higher education, learning, teaching and assessment? What is our collective responsibility? How can we place human connections and collaboration at the heart of it all?

While the traffic light system may have served a purpose at a moment of urgency, when there was a need to establish basic rules around AI use in assessment, the framework now appears increasingly limited. It no longer reflects the complexity, flexibility, and nuance of how AI is, or could be, integrated into learning and assessment. A more flexible approach is needed: one that connects learning and assessment more seamlessly and recognises the relevance and value of AI across the learning journey, rather than enforcing a hard stop at the point of assessment.

¹ VUCA standards for Volatility, Uncertainty, Complexity, and Ambiguity

We plan to engage in a collaborative inquiry, also considering the [seven principles of the TaLAI project](#) to explore what alternative frameworks might better support ethical, relational, and pedagogically grounded engagement with AI in higher education.

Acknowledgements

We would like to acknowledge that this article has been written based on detailed dialogic notes put together by the authors in preparation for the event. After the event, notes were further expanded and inserted into Secure Copilot. The output was further critically reviewed and developed by the authors to fully reflect their personal and collective thinking.

References

- Amiel, T., Cox, G., & de la Higuera, C. (2025, October 9). *Open education principles: Resisting the metrics of AI black boxes*. UNESCO. <https://www.unesco.org/en/articles/open-education-principles-resisting-metrics-ai-black-boxes>
- Attewell, S. (2025). *Student perceptions of AI*. Jisc. <https://www.jisc.ac.uk/reports/student-perceptions-of-ai-2025>
- Gorichanaz, T. (2023). Accused: How students respond to allegations of using ChatGPT on assessments. *Learning: Research and Practice*, 9(2), 183-196. <https://doi.org/10.1080/23735082.2023.2254787>
- Guilford, J. P. 1971. *The nature of human intelligence*. McGraw-Hill
- Nerantzi, C. (2025). *Module companion: EDUC5272M Education in a digital society*. School of Education. University of Leeds. doi.org/10.63560/moco8986
- Nerantzi, C., Palfreyman, J., & Pavlopoulou, M. (2025, October 16). AI traffic lights and embracing responsible use of AI in assessment. Assessment in the age of generative AI event, part of the TaLAI Project. *Media and Learning*. <https://media-and-learning.eu/event/assessment-in-the-age-of-generative-ai/>
- Rajagopal, K. (2025, November 1). Event report assessment in the age of generative AI. *Media and Learning*. <https://media-and-learning.eu/subject/artificial-intelligence/event-report-assessment-in-the-age-of-generative-ai/>
- Taskan, B., Junça-Silva, A., & Caetano, A. (2022). Clarifying the conceptual map of VUCA: A systematic review. *International Journal of Organizational Analysis*, 30(7), 196-217. <https://doi.org/10.1108/IJOA-02-2022-3136>
- University of Leeds. (2024). *Categories of assessment*. <https://generative-ai.leeds.ac.uk/ai-and-assessments/categories-of-assessments/>
- University of Leeds. (2025). *Learning with AI: A student edited collection*. <https://leeds.pressbooks.pub/learningwithai/>