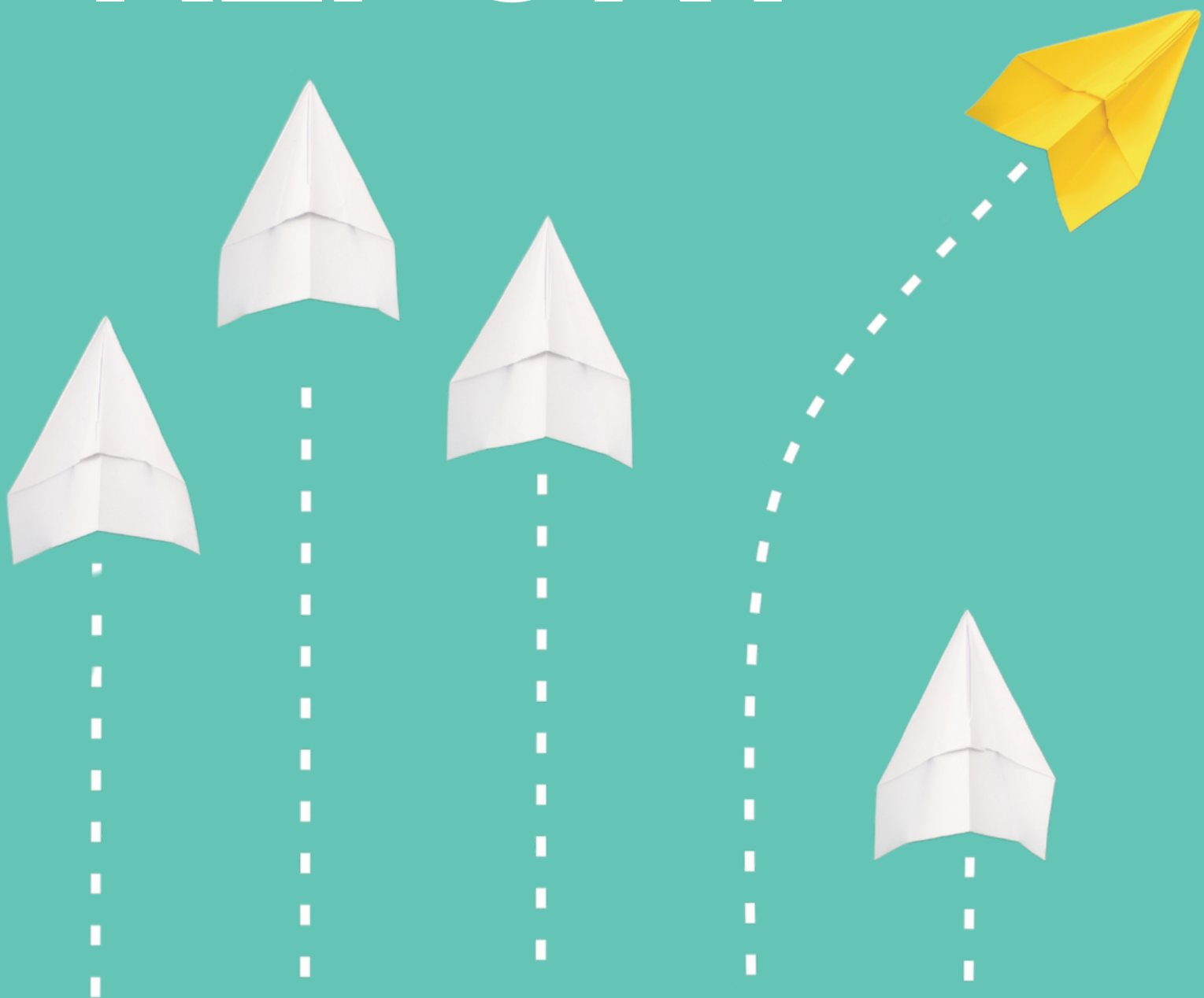


PANEL REPORT



EXECUTIVE SUMMARY



SESSION 2

The second meeting of the AI Edify Learning and Teaching Innovation Panel focused on assessment and non-exam assessment (NEA) in the age of AI.

Across the discussion, a number of tensions emerged. Current JCQ and exam board guidance was described as lengthy, inconsistent, and sometimes misleading—particularly in relation to “citing AI.” Detection tools were widely rejected as unreliable, raising concerns not only about accuracy but also about intellectual property and pupil–teacher trust. NEA models, still rooted in pre-AI assumptions, were seen as increasingly unfit for purpose. At the same time, there was caution that reform must not lead to overly controlled assessments that limit creativity or exacerbate inequities.

Alternatives were explored, including viva-style oral assessment, reflective methodologies, progressive disclosure models, and subject-specific approaches. These conversations consistently pointed back to a central question: what should qualifications measure in a world where AI is embedded in everyday practice?

External Guests:

The following guests were in attendance to listen and to join in the discussion. However it is important to note that the views expressed in this report are solely those of the panel members .

Laura Craven (Director of Compliance, Cambridge OCR and Chair of the JCQ Malpractice Group.

Connor Toms (Computer Science Subject Advisor, Cambridge OCR),

Joanna Williamson, (Research Fellow, Ofqual).

CHAIRS

Bianca Farthing

Founder - AI Edify,
Executive Director, Initiative for Conscious AI

Luke Ramsden

Deputy Head, St Benedict's School

Will Van Reyk

Deputy Head, Director of Innovation, St Paul's Girls' School

MEMBERS

Chris Loveday

Vice Principal, Barton Peveril Sixth Form College

Clare Jarmy

Director of Innovation, Haberdashers' Elstree Schools

Gary Henderson

Director of IT, Millfield School

James Flanagan

Deputy Head Academic at Magdalen College School

Jane Basnett

Director of Digital Learning, Downe House School

Jenny Powlesland

Assistant Headteacher (Teaching & Learning), Leighton Park School

Jess Payne

Deputy Head (Teaching & Learning), St Helen's School

Joanna Hughes

Group Director of EdTech, Berkhamsted Schools Group

Joe Cozens

Head of AI Governance, Clifton High School

John Doherty

Director of Digital Learning, Strategy and Innovation, Wycombe Abbey School

Luke Harris

Head of Computing and Digital Strategy, Bloxham School

Natasha Fenwick

Director of Enterprise and Partnerships, Queen's College London

Rebecca Markham-Jones

Assistant Head, Charterhouse School

Robin Street

Director of Innovation, North London Collegiate School

Steve Birtles

Head of Digital Teaching and Learning, Eton College

Tom Scott-Soundy

Head of Digital Learning & Innovation, Alleyn's School

Tom Wade

Assistant Head (Digital Strategy and Learning), Haileybury College

KEY THEMES

1. Guidance as Ambiguous, Burdensome, and Sometimes Misleading

The panel reflected on the current state of JCQ and DfE guidance, which was regarded as both essential and problematic. Schools depend on these frameworks to anchor policy, yet participants described them as lengthy, inconsistent, and not always aligned with classroom realities. The burden of interpretation falls heavily on schools, leading to uneven practice and a sense that responsibility has been devolved without sufficient clarity.

The treatment of referencing illustrates this tension. JCQ materials state "Where AI tools have been used as a source of information, student acknowledgement must show the name of the AI source used and the date the content was generated." Several contributors considered this unhelpful as it seems to imply that AI information carries a similar weight to that of conventional sources. By citing AI as a "source of information" it implies stability, traceability, and credibility — qualities that generative AI outputs do not reliably provide. Without clearer differentiation, there is a risk that pupils are encouraged to treat unstable outputs as equivalent to established sources.

Disclosure requirements were also questioned. If AI is embedded within everyday platforms such as Google Search, should pupils be expected to declare its use each time? Some participants suggested that this expectation risks trivialising the act of disclosure and undermining its intended purpose.

Detection software prompted a particularly careful exchange. The majority view was that such tools should not be used. Participants stressed that they are unreliable against current large language models and risk creating more ethical problems than they solve. While one perspective noted that outputs from detectors can give staff something tangible when concerns about authorship arise, this limited benefit was judged to be heavily outweighed by the risks. False positives remain common, unfairly undermining trust and confidence, while evidence also shows that detection systems discriminate against pupils whose first language is not English. Beyond these immediate harms, reliance on flawed tools risks lending institutional legitimacy to technology that lacks validity, transparency and fairness, diverting attention from more constructive approaches to academic honesty.

What unsettled several panellists most was the fact that plagiarism detection tools are included in JCQ guidance at all. Even without an explicit endorsement, their presence in the document was interpreted by many as a signal of approval. For some, this risks pushing schools towards practices that may do more harm than good, by legitimising unreliable methods and shifting attention away from more proportionate, relational approaches.

The discussion returned to communication. Suggestions included developing parent-facing guidance to complement school policies and reduce misunderstanding, and a broader call for national guidance that is concise, realistic, and adaptable across contexts. The panel recognised that a single framework may not account for subject-specific or institutional differences, and that future documents will need to balance shared principles with contextual flexibility.

2. Approaches to Regulating AI Use

The panel reflected on the varied ways schools are currently setting boundaries around AI use. Some have introduced traffic light systems, where “green” denotes permitted use, “amber” partial or guided use, and “red” prohibition. These were seen as offering clarity and a straightforward message for pupils. At the same time, participants noted that such systems can flatten what is in reality a more complex and context-specific issue. The interpretation of what counts as “appropriate” use often depends as much on the judgement of individual staff as on a written policy, and this variation can lead to inconsistency across subjects and departments.

Others reported taking a different stance by imposing a blanket ban on AI in NEA. This approach was described as the most straightforward and “safest” interim measure, reducing ambiguity and potential risk of malpractice while national guidance remains unsettled. Yet panellists also considered the potential drawbacks. A ban might create a sense of fairness by removing disparities between pupils with access to more advanced or paid-for AI tools and those limited to free versions. However, it could equally limit opportunities for pupils to develop responsible practices in using technologies that are already embedded in higher education and the workplace.

Alongside these approaches, one panellist suggested an alternative A–D model to categorise work as analogue, basic, controlled, or AI-enabled. Unlike traffic lights, this model was not intended to establish a hierarchy of value but to describe different modes of working. While not presented as a settled solution, it pointed to the possibility of frameworks that avoid binary distinctions and better reflect the range of practices schools are encountering.

3. NEA: Methodology over Product

The panel devoted considerable time to the issue of NEA, reflecting a shared sense that current models were designed for a pre-AI world and are becoming increasingly unsustainable. Traditional NEA is primarily concerned with the quality of the final product. In an environment where AI can generate polished outputs that are difficult to distinguish from human work, panellists questioned whether this emphasis still offers a fair or reliable picture of pupil learning.

A central thread of the discussion was the value of shifting the focus towards methodology and process. Several participants argued that students should be required to explain how they arrived at their work, including whether and how they used digital tools. Oral assessment and viva-style conversations were repeatedly highlighted as effective means of doing so. These were described as low-tech but powerful interventions, enabling teachers to test not just recall but depth of understanding. Panellists noted that such practices already exist in other frameworks: the International Baccalaureate makes use of oral elements. Therefore the recommendation is that teachers discuss NEA directly with pupils as part of the authentication process.

In addition to viva-style checks, a range of monitoring strategies were considered. These included requiring pupils to log progress at key stages, meeting regularly with a teacher or mentor to review drafts, and undertaking elements of NEA in supervised classroom time. Sampling techniques and the possibility of occasional unannounced vivas were also suggested as ways to build confidence in authenticity without creating excessive administrative burdens. The shared emphasis was on building points of dialogue and reflection into the process, rather than treating coursework as a single, isolated submission. Yet these approaches were not without concern. Allocating lesson time to supervised assessment could reduce the scope for broader teaching, and frequent check-ins may increase workload for staff. The panel acknowledged that a balance would be difficult to strike: too little oversight risks undermining confidence in the validity of coursework, while too much could transform it into a form of controlled assessment. This prospect caused unease, with several contributors warning that the very features that make coursework valuable — creativity, independence, and reflection — could be eroded if reform leans too far towards surveillance.

Differences between subjects also featured in the discussion. In some areas, such as the sciences, methodology is already a key focus, while in the humanities originality and interpretation are more central. This raised the question of whether a single model of NEA reform is realistic, or whether subject-specific approaches will be required. The group also recognised that universities often expect coursework to reflect authentic student voice, suggesting that reforms at school level will need to be compatible with wider academic practice.

Overall, the panel converged on the view that NEA must evolve in order to remain credible. While no single model was agreed upon, there was strong support for approaches that prioritise process, make use of dialogue, and strengthen the relational elements of assessment. At the same time, caution was expressed that reforms should not simply replicate controlled assessments under a new guise. The challenge will be to preserve creativity while ensuring authenticity, and to design systems that are both practical for teachers and meaningful for pupils.

4. Authenticity, Equity, and Trust

The discussion of coursework naturally led to broader questions about authenticity and the extent to which schools can be confident that submitted work reflects a pupil's own efforts. Towards the end of the session, the group was reminded that these concerns are not new. Long before the emergence of generative tools, teachers had to consider whether a sibling or parent had contributed disproportionately to a pupil's NEA. AI introduces new mechanisms, but the underlying principle remains the same: questions of authorship have always existed, and authenticity still rests on professional judgement, familiarity, and dialogue.

The importance of proportionality was a consistent theme. Teachers emphasised that approaches to authentication must be fair, workable, and rooted in trust. Overly rigid systems risk creating a climate of suspicion, damaging the relationships that are central to effective learning. Instead, strategies such as regular check-ins, structured conversations, and teacher familiarity with a pupil's work were presented as more sustainable ways of maintaining confidence in authorship. Mentoring arrangements, in which a member of staff follows the development of a pupil's work across time, were also suggested as ways to embed relational oversight.

Equity concerns added further complexity. Several contributors noted that if NEA reform involves increased in-class supervision or tightly structured check-ins, pupils with special educational needs or disabilities (SEND) could be disproportionately disadvantaged. Such students may require flexibility in how they work, and tighter restrictions could inadvertently narrow the accommodations available to them. At the same time, disparities in access to technology were raised as a pressing issue. Pupils with subscriptions to more advanced AI platforms are likely to be at an advantage compared to those limited to free tools, creating unevenness within classrooms. For some, this made a case for blanket bans in the interests of fairness; for others, it underscored the need to teach all pupils how to use AI responsibly, so that inequities are addressed through education rather than prohibition.

Another thread concerned the risks of replacing relational approaches with technological surveillance. Reliance on detection tools was seen as undermining trust between teachers and pupils, particularly given the risk of false positives. Several panellists argued that if authenticity checks are to command confidence, they must be grounded in dialogue and professional judgement, not outsourced to unreliable software. Taken together, the conversation suggested that authenticity in an AI era will require both continuity and change. The continuity lies in the enduring role of teacher judgement, familiarity, and relationship-building. The change lies in adapting assessment processes to acknowledge the presence of AI, without allowing it to destabilise equity or erode trust.

5. Subject-Specific and Future Challenges

The panel also reflected on how AI intersects differently with assessment across subjects. It was acknowledged that the issues raised by NEA are not uniform: in the sciences, where methodology and reproducibility are central, the challenges of AI may be less about authorship and more about how students justify their processes. In contrast, in the humanities, where originality and interpretation are highly valued, the risk of over-reliance on generative tools can look quite different. Several participants suggested that future guidance may need to move away from generic frameworks and instead develop subject-specific approaches that recognise these distinctions.

This raised a wider question about what qualifications should represent in a world where AI is unavoidable. Some argued that assessments should prepare students for the kinds of environments they will encounter beyond school, including workplaces where AI is already embedded. Others emphasised that qualifications also have a role in safeguarding academic integrity and ensuring that the distinctive contributions of pupils are visible and valued. The tension between preparing for the future and protecting established standards ran throughout the conversation.

The discussion also touched on more speculative concerns. Questions were raised about the long-term implications of technologies that may one day go beyond generative AI, including implantable devices. While these scenarios remain hypothetical, their inclusion reflected a broader awareness that guardrails must be designed to adapt over time. Guidance that responds only to today's tools risks becoming obsolete as technologies evolve.

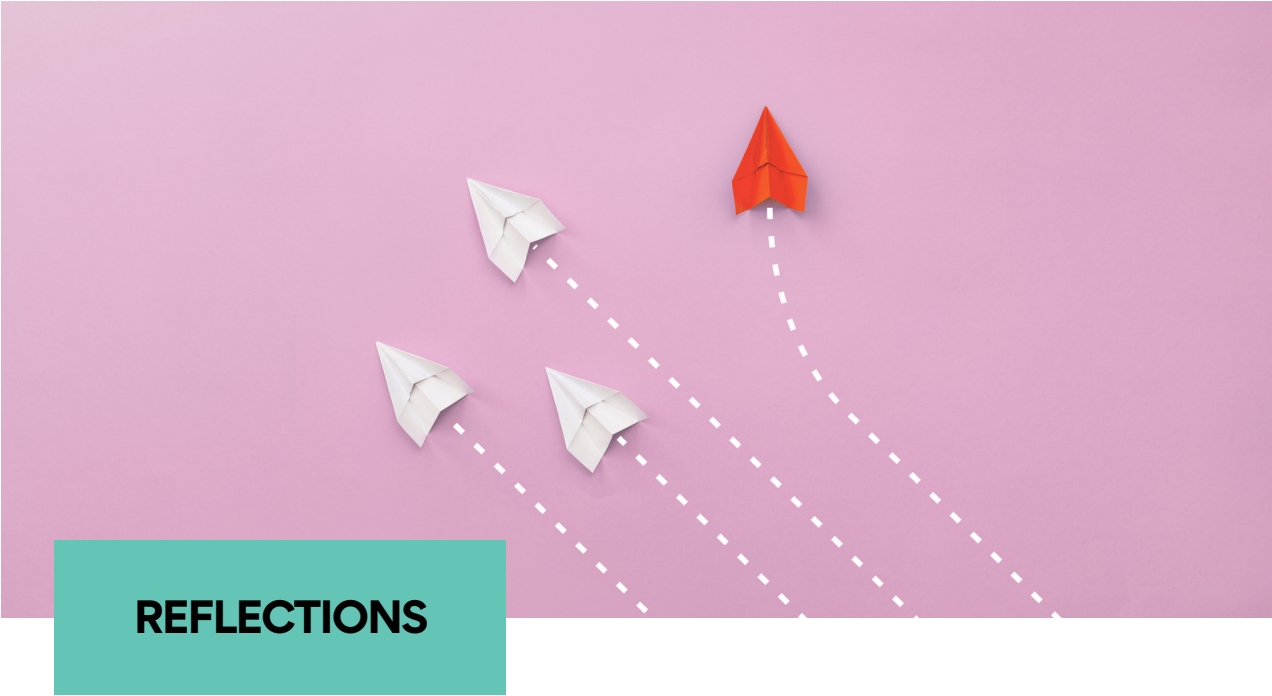
Although no consensus was reached, the discussion made clear that assessment reform cannot rely on blanket measures alone. Subject-specific nuance, forward-looking flexibility, and adaptability to technological change will all be essential if qualifications are to remain both fair and meaningful.

Takeaways, Guidance, and Advice

The session produced a range of insights for schools and policymakers who are grappling with the role of AI in assessment. While consensus was not always possible, several clear messages emerged.

- Avoid detection software. Although some staff value its ability to provide a tangible starting point when challenging pupils about authorship, the risks of false positives and the ethical concerns about uploading work to third-party platforms were judged by most to outweigh the benefits. Its inclusion in JCQ guidance was read by several panellists as an implicit endorsement, which they felt was unhelpful and potentially misleading.
- Clarify language around referencing AI as a “source of information”. Current messaging risks encouraging pupils to treat AI outputs as stable sources when they are not. Clearer differentiation is needed between acknowledging the use of AI and citing conventional references.

- Prioritise process over product. NEA should give weight to methodology, reflection, and explanation of choices. Approaches such as oral assessment, structured check-ins, progress logs, and mentoring relationships can help ensure authenticity while maintaining trust.
- Build relational rather than surveillance-based systems. Familiarity between teachers and pupils, alongside proportional checks, were seen as more reliable than technological monitoring.
- Attend to equity. Bans may create fairness by levelling access, but they may also limit the development of responsible skills. Similarly, tighter structures could unintentionally disadvantage pupils with SEND. Schools must weigh equity concerns carefully in designing local policies.
- Acknowledge subject differences. One-size-fits-all frameworks are unlikely to capture the realities of AI across disciplines. Subject-specific approaches will be necessary if qualifications are to remain both rigorous and relevant.
- Remember continuity as well as change. Concerns about authorship are not new. Long before generative AI, teachers asked how they could be sure a sibling had not done the work. The challenge today is not to reinvent authenticity from scratch but to adapt long-standing principles to new tools and contexts.



REFLECTIONS

This session underlined the difficulty of adapting assessment practices to an environment where AI is already embedded in everyday tools but not yet fully accounted for in national frameworks. A consistent theme was the tension between clarity and oversimplification. Traffic light systems and blanket bans both provide straightforward rules, yet risk flattening the complexity of subject-specific contexts and individual teacher judgement.

NEA attracted particular scrutiny. While panellists agreed that the current model which was designed for a pre-AI world, is increasingly unsustainable, there was caution that reforms must not tip into controlled assessments that narrow creativity. The most promising suggestions emphasised process over product: oral questioning, reflective commentary, and regular dialogue between staff and pupils. These were seen as proportionate ways of safeguarding authenticity without undermining trust.

Across all themes, there was recognition that questions of authorship pre-date AI. As one was observed, teachers have long asked how they can be sure that work was not completed by a sibling or parent. The principle remains the same: authenticity rests on professional judgement, familiarity, and proportionate oversight. AI complicates these judgements but does not remove the need for them.



FINAL THOUGHTS

CONCLUSION

The discussions in this session pointed to a future where assessment cannot remain in its current form. Existing models, built for a pre-AI landscape, struggle to safeguard authenticity and risk becoming detached from the realities of how pupils learn and work. Reform is needed, but panellists cautioned against solutions that erode creativity or impose overly rigid controls.

Across the debate, several priorities emerged. National guidance must be clearer, more consistent, and free from implied endorsements of unreliable tools. NEA should place greater emphasis on methodology, reflection, and dialogue, supported by proportionate checks that preserve trust. Equity must be a guiding principle: reforms must not disadvantage SEND pupils or widen gaps between those with and without access to advanced technologies. Subject differences should also be recognised, with frameworks that allow flexibility rather than assuming uniform solutions.

GOT ANY QUESTIONS?

Email:
Bianca.Farthing@aiedify.com

Website:
www.aiedify.com

