Pearson

School Educator Guide

Redefining Formative Assessment in a Generative AI Era

Practical Strategies for Future-Ready Assessment

Designed for school educators, this guide provides actionable strategies to evolve formative assessment in the GenAI era.

Building on our Assessment Evolved report, our step-by-step, concrete examples support teachers to protect learning integrity and to build critical thinking, AI literacy and future-ready skills.

Contents

Building integrity in AI use

The emergence of Generative Artificial Intelligence (GenAI) has radically changed the landscape of assessment in education. Large Language Models (LLMs) like ChatGPT can now produce high-quality essays, and quickly provide answers to multiple-choice questions and practice problems. While the media warns of academic dishonesty, there is a real opportunity to enhance **learning integrity** and build AI literacy in our schools. The focus of our research and associated recommendations is on 'learning integrity' - ensuring that in a GenAI world, students genuinely engage in the process of learning rather than offloading their thinking to an LLM. Learning integrity is critical for developing knowledge, skills and understanding.

The rise of GenAI gives educators a powerful opportunity to evolve formative assessment activities. Many are already thoughtfully exploring its potential. For those just starting this journey, we hope this guide will provide practical recommendations for rethinking and reimagining what formative assessment looks like in a GenAI world.

Our research

Our research focused on secondary education (US grades 6–12/ UK Key Stages 3–5), as these groups are most likely to experiment with and adopt GenAI tools for learning. We surveyed more than 500 educators in the United Kingdom and United States and interviewed a range of global experts in AI and education (for more details on our methods and results, see our Assessment Evolved report). We have used that data, underpinned by existing research, to create this guide.

What this guide is for

We know formative assessment does not exist in a vacuum, and that making changes to one part of the broader teaching and learning system will inevitably have ripple effects elsewhere. However, we purposefully focus on formative assessment because the challenges around maintaining learning integrity in the face of AI are more pressing here. The topic of how summative assessment will or should evolve in an age of AI, though important, is beyond the scope of this guide. Our intention for this guide is to provide educators with practical steps for reflecting on and evolving their assessments.

A snapshot of teacher perceptions

Given how quickly GenAI is evolving, educators, understandably, have different perspectives, levels of confidence and experience with using it. Some educators are taking a more active approach by embedding AI into their formative assessment, while others remain resistant to any AI use in education. Our survey participants generally fell into one of four "personas" based on their pattern of beliefs and practices.

School educator profiles

Protective Skeptics	30%, n=153

"[With AI] I think students would not be learning to the best of their ability which would impact their knowledge moving forward. Feeding their assignments into a machine in return for answers is extremely concerning to me as an educator."

This educator is most protective of formative assessment and sees the risks that students' GenAI usage poses. That said, they could accept AI's role in education when used appropriately.

- **Negative** sentiment toward AI
- They perceive students' AI familiarity as **higher** than their own.

Proactive Innovators	49%, n=247

"AI presents a wealth of exciting possibilities for meaningfully integrating into formative assessment, largely by enhancing personalization, efficiency."

This educator is aware of the risks and drawbacks of AI, but they are willing to experiment and are most likely to incorporate AI into the assessment itself.

- **Positive** sentiment toward AI
- They perceive students' AI familiarity as **lower** than their own.

Staunch Traditionalists	7%, n=36

"Generative AI is entirely destructive to higher education... I see nothing positive about it whatsoever. We, as a culture, are careening headlong into disaster."

This educator is more concerned and resistant to AI. They are least likely to adapt their assessments and see little promise in AI altogether.

- **Negative** sentiment toward AI
- They perceive students' AI familiarity as **higher** than their own

Cautious Explorers	14%, n=69

"I believe it can be helpful in some cases, however, I believe it can only give surface level understanding, [it could cause] loss of authentic insight and causes for academic integrity concerns."

This educator hasn't yet made their mind up on AI's role in assessment. While they are open to its possibilities, they have concerns that need to be addressed.

- **Positive** sentiment toward AI
- They perceive students' AI familiarity as **equivalent** to their own.

This guide aims to meet you wherever you see yourself on this spectrum, whether that's further refining your AI-enabled assessments, or exploring its possibilities for the first time. The reflection questions throughout this guide are tailored for different starting points.

Choose the set that best fits where you are now and use them to reflect on your future practice.

From problematic to productive GenAI use

A focus on learning integrity means encouraging usage of GenAI that is helpful (and not harmful) for learning. Our expert interviewees broadly agreed on what constitutes 'productive' vs 'problematic' use of GenAI for formative assessment.

It can be helpful to consider these indicators of productive versus problematic use when reflecting on your existing formative assessments. ⊙

Features of problematic student use of GenAI	Features of productive student use of GenAI
Students are confused about how, where, when and why GenAI should be used	Students are clear on how, where, when and why GenAI should be used
AI is used to bypass steps of the learning process to go straight to output	GenAI is used as an aid to process or refine ideas
Students outsource their core learning to GenAI	Students show evidence of ideation/originality, even if GenAI is involved
Students cannot explain their thinking or conclusions	Students can fluently communicate their process, ideas, or results, including the role of GenAI in the final product

" These generative tools – they're designed to be a conversation. They're not actually just a oneway – do this, get the result, on you go. We know from using them that you get the best result through iteration. So, there's something to explore around reflection or understanding, getting the student to work with AI on understanding how they iterated their ideas. How have they shaped their ideas? Where did their stimulus come from? Digging into their process, asking them questions about their process."

Rachel Evans
Director of Digital Transformation
Girls' Day School Trust

Testing the limits of your formative assessment

Exploring GenAI's capabilities firsthand is a practical way to demystify its affordances and limitations. If you are new to using this technology, you can start by inputting one of your current assessments into an LLM like ChatGPT. This process can help to highlight both the opportunities and challenges GenAI presents for your formative assessments, sparking ideas for innovation while exposing potential vulnerabilities. **You could use the following prompt (substituting your course details when relevant)** →

> I am a **[grade/key stage]** student in **[name of your course]**. I have received the following assignment:
>
> **[insert your assignment directions here]** as part of **[name of unit/topic]**
>
> Generate the answers/a response to this assessment.

Optional

- In the prompt, ask the GenAI to honor page limit or any other formatting directions
- Experiment with adding constraints e.g., ask for step-by-step reasoning
- Provide an example of a model assignment (do not include any identifiable student data)
- Add instructions for tone

Reflection questions for

Staunch Traditionalists

- What did you learn about AI's capabilities through this exercise?
- Would the AI-generated output receive a passing grade/mark in your course?

Protective Skeptics

- Did the tone or reasoning of the AI output align with what you would expect from your students on this assessment?
- Which aspects of this assessment could remain the same even if students do use AI?

Cautious Explorers

- How could the AI output be used to help students understand different levels of performance (e.g., average vs. excellent work)?
- Imagine redesigning this assessment for a context where AI use is *expected*. What would you change?

Proactive Innovators

- What discussion questions could help students to critically evaluate this output?
- How could you modify one aspect of this assessment to require more personal reflection or authentic application of skills?

Models for rethinking assessment

Considering this new GenAI impacted landscape, thought leaders in AI and education have constructed various frameworks for thinking about the role GenAI can or should play in assessment.

Broadly, we can think of assessment in the age of GenAI through the lens of the Swiss Cheese Model,[1] where multiple, overlapping layers of assessment are ultimately effective, even though each individual assessment may have weaknesses (i.e., holes in a slice of Swiss cheese).

We know that no single assessment comprehensively captures students' learning and is completely resilient to GenAI misuse. But requiring students to demonstrate their learning via a variety of assessment types creates safety nets and ensures more opportunities for students to practice and demonstrate their learning.

With this in mind, we wanted to highlight the following three frameworks:

Framework 1: The Two Lane Approach[2] outlines two different but complementary approaches (i.e., lanes) for assessing student learning in a GenAI era:

Lane 1: summative assessments (Assessments *of* Learning) which happen in person, without GenAI assistance. The purpose of this lane is to provide an accurate and trustworthy judgement of students' actual learning.

Lane 2: formative assessments (Assessment *as* Learning) where students and teachers may intentionally integrate GenAI to varying degrees in accordance with the course/unit goals. Lane 2 assessments are seen as preparation to succeed on Lane 1 assessments. Should students misuse GenAI in their Lane 2 assessments, it is unlikely that they will adequately pass assessments in Lane 1.

Framework 2: The Stoplight/Traffic Light Model[3] comprises three levels for AI use in assessment:

Red: No AI use is permitted.

Yellow: AI permission can be granted if the student details their intentions and rationale for use.

Green: AI use is encouraged in the assessment but requires a conversation with the instructor.

Framework 3: The AI Assessment Scale[4]
details five levels of AI integration ranging from:

1 **No AI:** You must not use AI at any point during the assessment. You must demonstrate your core skills and knowledge.

2 **AI Planning:** You may use AI for planning, idea development and research. Your final submission should show how you have developed and refined these ideas.

3 **AI Collaboration:** You may use AI to assist with specific tasks such as drafting text, refining and evaluating your work.

4 **Full AI:** You may use AI extensively throughout your work either as you wish, or as specifically directed in your assessment. Focus on directing AI to achieve your goals while demonstrating your critical thinking.

5 **AI Exploration:** You should use AI creatively to solve the task, potentially co-designing new approaches with your instructor.

While each framework takes a slightly different approach, they agree that:

- Attempting to keep GenAI out of assessments entirely is both unrealistic and limits students' preparation for the realities of their future education and careers.

- At the same time, there are real risks to learning integrity for students who completely outsource their learning to GenAI.

- Different assessment types serve different purposes depending on the learning goals and context. This has implications for whether/ how GenAI should be used in the context of that assessment.

- It is good practice to include a variety of assessment types and formats that allow students to demonstrate their knowledge and skills in a variety of ways.

In the next section, we provide detailed guidance on how you might evolve your assessments, based on these points of expert consensus. ⊝

Reflection questions for

Staunch Traditionalists

- Does thinking about AI integration along a continuum or spectrum make it feel more manageable?

- Which of these models align best with your own philosophy?

Cautious Explorers

- What protective measures (e.g., traffic lights, Swiss Cheese layers) would give you confidence to try integrating AI into one of your assessments?

- How could sharing or discussing these models with students be used to develop a common understanding of responsible AI use?

Protective Skeptics

- How could you use one or more of these models to communicate clear boundaries around AI use in your classroom?

- What concerns do you have that aren't addressed by any of these models?

Proactive Innovators

- How could you use these models to start a conversation with colleagues who might be more hesitant or skeptical of AI?

- Where could you draw on any of these models to refine or scale a sustainable approach to integrating AI in your assessments?
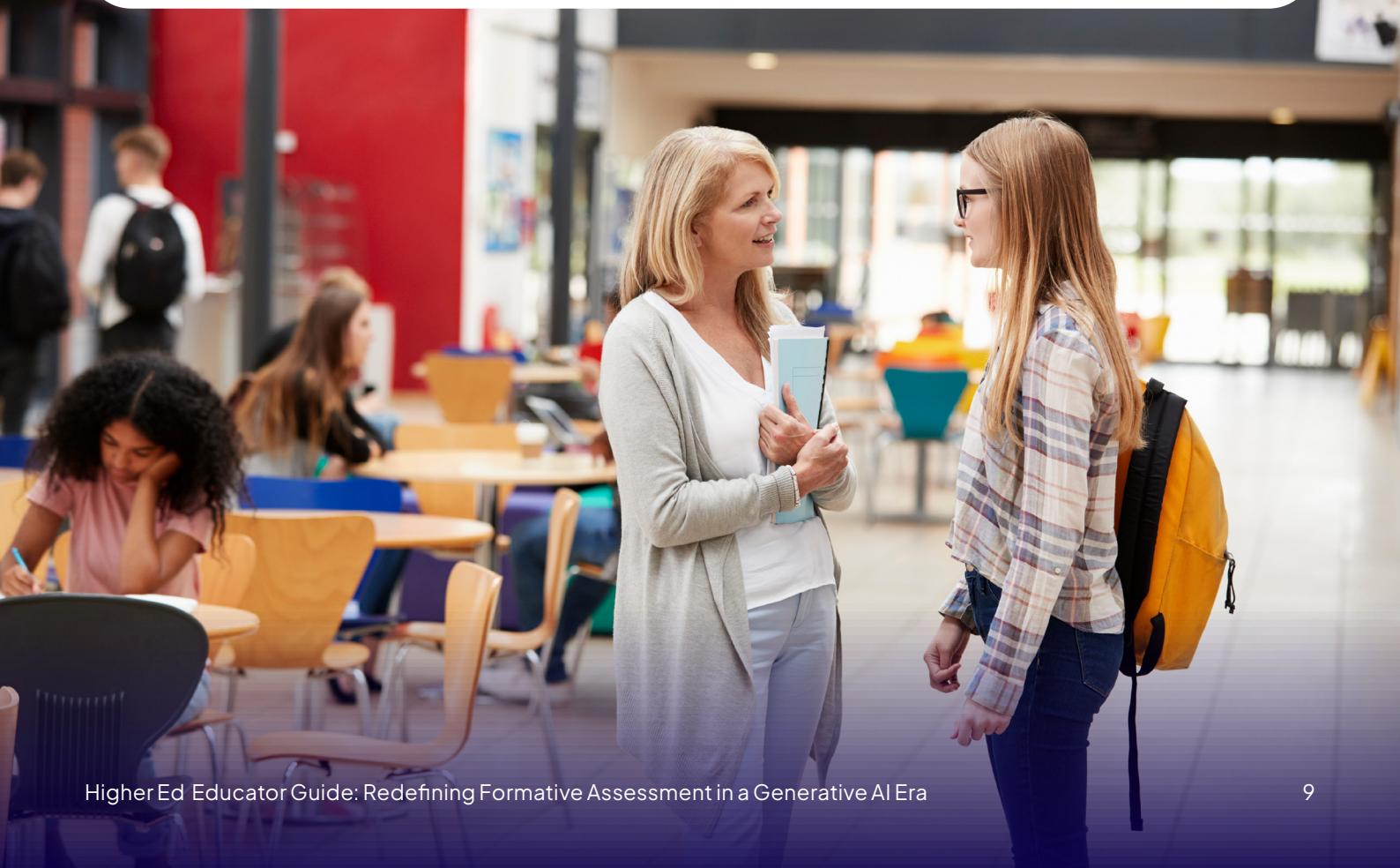
Reimagining formative assessment activities

Redesigning formative assessment to preserve learning integrity in the age of GenAI is no small task, especially given the many other challenges educators are facing.

While rethinking formative assessment may initially require more time and effort for educators, this process will reinforce learning integrity and help to build genuine student understanding in the long run. The strategies outlined here, informed by our educator survey and expert interviews, are designed to be realistic and flexible, can be adopted gradually and will ideally build on educators' existing practices. As we outline in our Assesssment Evolved report, we also recognize that substantial change cannot come without considerable support from both school leaders and policy-makers.

#1: Have a clear purpose in mind

Is your goal the assessment *of* learning (i.e., Lane 1 of the Two-Lane Approach) or is your goal assessment *for/as* learning (Lane 2)?

Will student performance contribute to their final grade for the subject, or is it a means to provide evaluative feedback? What works well for one purpose won't necessarily work well for another, so it's important to be clear on your purpose from the outset.

For assessments of learning, particularly those that entail high stakes for individuals, the validity of interpretations is paramount, which implies a need for a controlled, secure assessment environment to rule out GenAI misuse and other forms of misconduct.

For this reason, many Lane 1 assessments are given in supervised environments (e.g., in-person or remote-proctored).

However, assessment for/as learning allows for a more open assessment environment, with students completing some or all of the assessment unsupervised. Here, GenAI usage may be allowed, invited, or even required, depending on the subject and/or particular assessment targets.

> "
>
> Before we even talk about assessments, we need to reconsider what is being assessed and why?"
>
> **Pat Yongpradit**
> Chief Academic Officer of Code.org
> and Lead of TeachAI

#2: Clearly articulate what you are trying to measure or develop

Once you are clear on the assessment purpose, you should re-examine the assessment target(s): what combination of knowledge, skill or other attributes are you trying to measure and/or develop?

Clarifying what we aim to target is more significant than ever with the rise of AI. Often in school settings, these targets are expressed in the form of academic content standards, such as this Grade 5 Writing Standard from the Common Core State Standards in the US: "Write narratives to develop real or imagined experiences or events using effective technique, descriptive details, and clear event sequences."

A version of this standard that accounts for the reality of GenAI might be: "Use GenAI as a drafting tool in the process of writing narratives to develop real or imagined experiences or events."

> "
>
> We assign essays as a structure, a construct, when we are not assessing a kid's ability to write an essay – we want to know their ability to be argumentative, to draw conclusions, to connect evidence, to have a voice. But what we do is force those different types of learning into a construct that was essentially created for school."
>
> **Amanda Bickerstaff,**
> Chief Executive Officer
> AI for Education

The OECD's AI Literacy Framework[5] outlines how the distinct knowledge, skills and attitudes required to thrive in an AI-saturated world can be helpful in reframing traditional assessment targets. For example, one of the 'Creating with AI' Competences in the framework focuses on using *AI systems to explore new perspectives that build on original ideas*. This competence draws from knowledge about AI's capabilities and limitations; for example "AI excels at pattern recognition and automation but lacks emotions, ethical reasoning, context, and originality." It also weaves in human creativity skills (eg.,the ability to "collaborate with AI to create and refine original ideas while considering issues of ownership, attribution and responsible use") and fosters attitudes of innovation and adaptability.

Reframing the assessment target(s) also requires adjusting learning target(s). That is, students must be explicitly taught how to use GenAI responsibly and ethically for creative writing and for any other academic tasks.

While there are moments when it is important to assess what students can accomplish independently, educators should also consider structured activities (like formative assessment) that help to develop the skills to use GenAI responsibly and effectively. For example, you could explore AI capabilities with students by performing "benchmarking" tests of different AI models to compare outputs or having AI provide an answer to an assignment and then conduct a "tear-down" class discussion. Changing classroom assessments in isolation won't be enough to prepare students to live and work in an AI-powered world though, and in time, learning standards, curricula, and instructional practices will need to change as well.

Example competence from the OECD AI Literacy Framework

Creating with AI competences

Use AI systems to explore new perspectives and approaches that build upon original ideas

Learners experiment with AI to expand their thinking, generate new ideas, or consider alternative viewpoints. They stay accountable for the final content while letting AI support their creative process.

Primary Education Scenario

Evaluate AI-generated images to create story settings based on learner ideas (e.g., "a jungle in space"), then write new stories inspired by unexpected results.

Secondary Education Scenario

Use AI to develop counterarguments for a class debate to anticipate and address opposing viewpoints.

Taken from OECD (2025). Empowering learners for the age of AI: An AI literacy framework for primary and secondary education (Review draft). OECD. Paris. https://ailiteracyframework.org

---

"

In an AI world, we are probably going to focus on different sorts of learning outcomes, so that the ways to get indicators of those are different [too]."

Philip Dawson
Co-Director of the Centre for Research in Assessment and Digital Learning Deakin University

#3: Focus on process over product

Shift the focus from the tangible end products (e.g., the final answer to a math problem, the completed essay) to the process of developing those products.

Reframing the assessment target to be about *how* students responsibly show what they know and can do emphasizes the learning process over the end product. This can have the added benefit of making assessment more authentic, and more reflective of the future scenarios where students will be expected to exercise these abilities in the real world.

But increasing the authenticity of an assessment doesn't mean replicating the workplace in your classroom. Instead, it means designing assessments that require students to meaningfully *apply* their knowledge and skills in new situations to solve problems. Using tools like GenAI thoughtfully in the process to generate, refine and elaborate on their ideas and seek feedback for improvement also helps to build AI competences.

Focusing on the **process** may require putting more structure around an assessment, breaking it up into multiple iterations and outputs or providing scaffolding along the way. Evaluation should be based, at least in part, on the process itself by including checkpoints or graded milestones to emphasize the importance of engagement. This may also require students to submit logs of their interactions with an LLM to understand how it was used and how it influenced their thinking. These additional checkpoints can be scheduled for in-class discussion or submitted with evidence of process, like revision notes.

> "
>
> Create a short paragraph yourself. Put this into the model and ask what could be improved in this particular draft, then work iteratively with the model to come up with a version of your own ideas that improve on the initial draft... To me, that is a productive way of interacting with an LLM. I don't really see any problem if my students are using AI in this way. They are bringing their own ideas; they are working with these models to help them improve their own ideas."
>
> **Mutlu Cukurova**
> Professor of Learning and AI,
> University College London

This is not a call to abandon the final work product completely, however. Where GenAI is integrated, clear boundaries for AI's role should be provided and students should still be responsible for ownership of the final work product. This helps to cultivate the literacy, judgement and critical thinking needed to prepare students for a future in which GenAI is not a shortcut but an essential professional tool.

This could mean incorporating an accountability exercise into the assessment – e.g., a short conversation with the teacher to explain their decision-making process around the final submission, including what they did and why they did it that way.

> "
>
> Just as in the early days of Coursework 1.0 you asked to see a student's drafts, in the AI world you'll need to see time-stamped logs. The learning will exist between Log One, Log Two, Log Three, and whatever passes for the final version. That's where you'll find your real learning."
>
> **Bill Lucas**
> Professor of Learning and Director of the Centre for Real-World Learning, University of Winchester, Co-Founder Rethinking Assessment

#4: Diversify formats and modes of assessment

Another suggestion, which harkens back to the Swiss Cheese model, is to think more broadly about what constitutes evidence of/for learning, particularly non-traditional evidence and diverse assessment formats or modes.

For example, ask students to submit portfolios that showcase the evolution of a piece of work over time, or require students to submit a range of different evidence types, as appropriate to the assessment target and subject. As well as written work, this could be voice memos, videos, visual timelines, chatbot logs, etc.

Importantly, multi-modality calls for triangulation of inferences about student learning across different types and modes of assessment. The benefit is that when you piece together these different sources of information, you have a more complete and holistic sense of what the student knows and can do compared to a single task.

> "
>
> ...the entire program [should be] designed in a way that emphasizes gathering evidence of learning and students providing evidence. But we can't just trust one piece of paper. We need to verify that. It's having multiple cross-verifications, which aren't just all big sit-down exams."
>
> **Kane Murdoch**
> Head of Complaints, Appeals and Misconduct Macquarie University

> "
>
> The phrase I would use is 'multi-modality'. We need to stop worrying about single-source indicators, because just as if you triangulate something in a piece of research, you're getting three or more perspectives."
>
> **Bill Lucas**
> Professor of Learning and Director of the Centre for Real-World Learning, University of Winchester, Co-Founder Rethinking Assessment

#5: Introduce social or collaborative elements

Outside of formal education, people must often solve problems in groups.

Bringing in teamwork is one way to simulate a real-world context, and asking students to work together in their use of GenAI can help them be more reflective about how, when, and why they are using it.

In pairs or small groups, ask students to come up with potential prompts that could be used as a jumping-off point for students' own ideas. You could then lead a whole-class discussion evaluating the quality of GenAI outputs from different LLMs or compared to students' brainstorming. These activities can be sequenced, so that students first engage in individual preparation before class time is used for idea refinement, collective analysis or to exchange feedback.

"

Community in learning is important. Those social interactions within the room are really vital. You don't want to lose those, and so, you need to find a way that personalised or individualised learning doesn't look like every child completely on their own track, in their online learning system - that's a bit dystopian. You want to be able to keep them together, but through individually generated text, they could have a much more individual approach in the same general subject area."

Rachel Evans
Director of Digital Transformation
Girls' Day School Trust

"

Let me talk about the self-regulation of learning. Now they talk about the co-regulation of learning, acknowledging that others - teachers, peers, the book, the computer and the questions, the textbook - all of those things also help you regulate your learning. But you can't regulate learning until you have a goal. You start with a goal and then a learner regulates their cognitive, affective and behavioral resources that they have (and if they don't have them, they get them from somewhere, their peers or a book or something) in pursuit of that goal, and they need feedback along the way."

Susan Brookhart
Professor Emerita in the School of Education
Duquesne University

#6: Incorporate higher-order skills and processing

Instead of asking for a summary, prompt students to critique, compare, or apply concepts to a novel, real-world situation.

> "
>
> "I created a mini simulation for students saying 'You are running a small chain of coffee shops. Based on these different scenarios, what are you going to do?'... Being able to do those things has been something that I could never have done before [GenAI]."
>
> **Mike Perkins**
> Head of the Centre for Research and Innovation
> British University Vietnam

To elevate an assessment, ask students to propose solutions to a community issue, using evidence from multiple sources, or to reflect on how course concepts relate to current events. Because modern LLMs are now capable of producing sophisticated responses to even higher-order thinking prompts, you should try to create tasks that require students to weave together contextualized and personalized knowledge or apply concepts to personal contexts that would be more difficult for an LLM to create without extensive background knowledge about the student.

Evolving an assessment: In practice

Let's apply these suggestions to demonstrate how a traditional assessment could evolve to better reflect the commonplace reality of GenAI.

A traditional assessment might consist of a take-home creative writing assignment in which students must produce an imagined first-person account of what it would have been like during the volcanic eruption of Pompeii. Figure 1 provides several alternative versions of a traditional assessment activity, using the approaches discussed in this guide (such as emphasizing process over product). See the Appendix for another example of how to update this activity by applying multiple recommendations within a single extended task.

These alternative versions of the task may initially take more teacher time and planning to implement but could produce more diverse evidence of learning while touching on a wider array of knowledge, skills and attitudes. It would also provide opportunities for students to actively explore (in a supervised setting) how to responsibly and ethically use GenAI tools as an aid in the writing process, a skill that will inevitably serve them well in future.

Figure 1. Traditional grade 5 creative writing task versus updated task ideas integrating GenAI

Approach	Task description
Target standard: Write narratives to develop real or imagined experiences or events using effective techniques, descriptive details, and clear event sequences	
Original task	Write an imagined first-person account of what it would have been like during the volcanic eruption of Pompeii.
Ai as a tool for critique or debate	Ask a large language model (LLM) to describe the sequence of events during the volcanic eruption of Pompeii.
	Use track changes to edit this document and use comments to identify what the LLM did/did not do well. Submit this document along with your final essay.
Process over product	Submit all three parts of this assignment in sequential order:

1. (Outside of class): Generate a timeline of the sequence of events during the volcanic eruption of Pompeii using your source material or a Large Language Model (LLM; if you use an LLM, note that you will need to review the output for accuracy-please include note of any corrections you had to make)

2. (In class): Using the sequence of events from Step #1, generate 1–2 descriptive sentences about each event, imagining you were there and experiencing it in person.

   For example, an event from Step #1 might be "the mountain began to smoke" and in this step, you might write "The mountain began to shimmer in the sunlight with silver smoke."

3. (In class): Picking your favorite descriptive sentences, create a narrative that sequences the events of Pompeii from a first-person account. Submit the document with version history enabled for your teacher.

Bonus: Submit a short reflection about this process. Include what you enjoyed, what was difficult, etc. In that reflection, make sure to name how and where you used GenAI, what it did well, and what it did not do well.

Approach	Task description
Reminder of target standard: Write narratives to develop real or imagined experiences or events using effective techniques, descriptive details, and clear event sequences	
Multiple modalities	Write an imagined first-person account of what it would have been like during the volcanic eruption of Pompeii. Note: You will use class time to complete the first draft without the use of GenAI, giving you the foundation to build on for your next task/s.
	Create a digital storyboard including photos and videos. Include quotes from your first-person account throughout the storyboard. In addition to submitting the essay and storyboard, you may also be asked to present the storyboard to a partner or to the class during discussion.
Requiring AI justification	Write an imagined first-person account of what it would have been like during the volcanic eruption of Pompeii.
	If you would like to use GenAI, share why, where, and how you plan to use the tool. Requests to use GenAI are due at least one week before the assignment due date. If you use GenAI, submit a reflection on how and where you used GenAI, what it did well, and what it did not do well.
Real-world assessment	Imagine you have been hired by the Pompeii Archaeological Museum to help visitors imagine what life was like during the eruption of Mount Vesuvius. Your task is to write a first-person account (in diary, letter, or spoken monologue form) that could appear in an exhibit or educational podcast.
	Use historical evidence about Pompeii's daily life and the eruption to make your account realistic.
	You will be asked to present your monologue to a section of our class who will follow up with at least 5 questions about your process for writing the account. If you use AI tools, indicate how and where they are used both in your writing and in your presentation.

Remember: While this example showcases some of the strategies we have outlined, there will be adjustments of varying degrees that you can make across your assessment activities. We encourage you to experiment, to consider how different approaches might work across the curriculum, and to collaborate with your colleagues to share your learning and to innovate together.

Conclusion

The rise of GenAI provides a critical opportunity to re-evaluate formative assessment in schools.

Students now have a host of AI tools at their fingertips, and while it is tempting to try to block their access to this technology in the name of learning, these students will eventually need to navigate careers and everyday lives in an AI-driven world. This places an upfront responsibility on educators to identify how students should (or should not) integrate GenAI into their learning.

Fortunately, there are experts and educators already leading the way, and our research shows several tangible takeaways for responsible and effective AI use in assessment. Whether it was an expert framework, an educator profile, or one of our six key suggestions for assessment redesign, we hope that the above content sparks ideas, conversations, and practical changes in classrooms.

We encourage ongoing dialogue and sharing of experience, both among teachers themselves, and with their students, about how GenAI can be thoughtfully used to enhance and enrich the learning experience.

The opportunity ahead lies not in resisting these technologies, but in leveraging them in ways that foster deeper, more meaningful understanding and engagement.

"

And so, the question for us as educators is to think about how we can help our students engage thoughtfully with this new reality that we find ourselves in. And help them to realize that developing their own capabilities is still really important as humans."

Danny Liu
Professor of Educational Technologies
University of Sydney
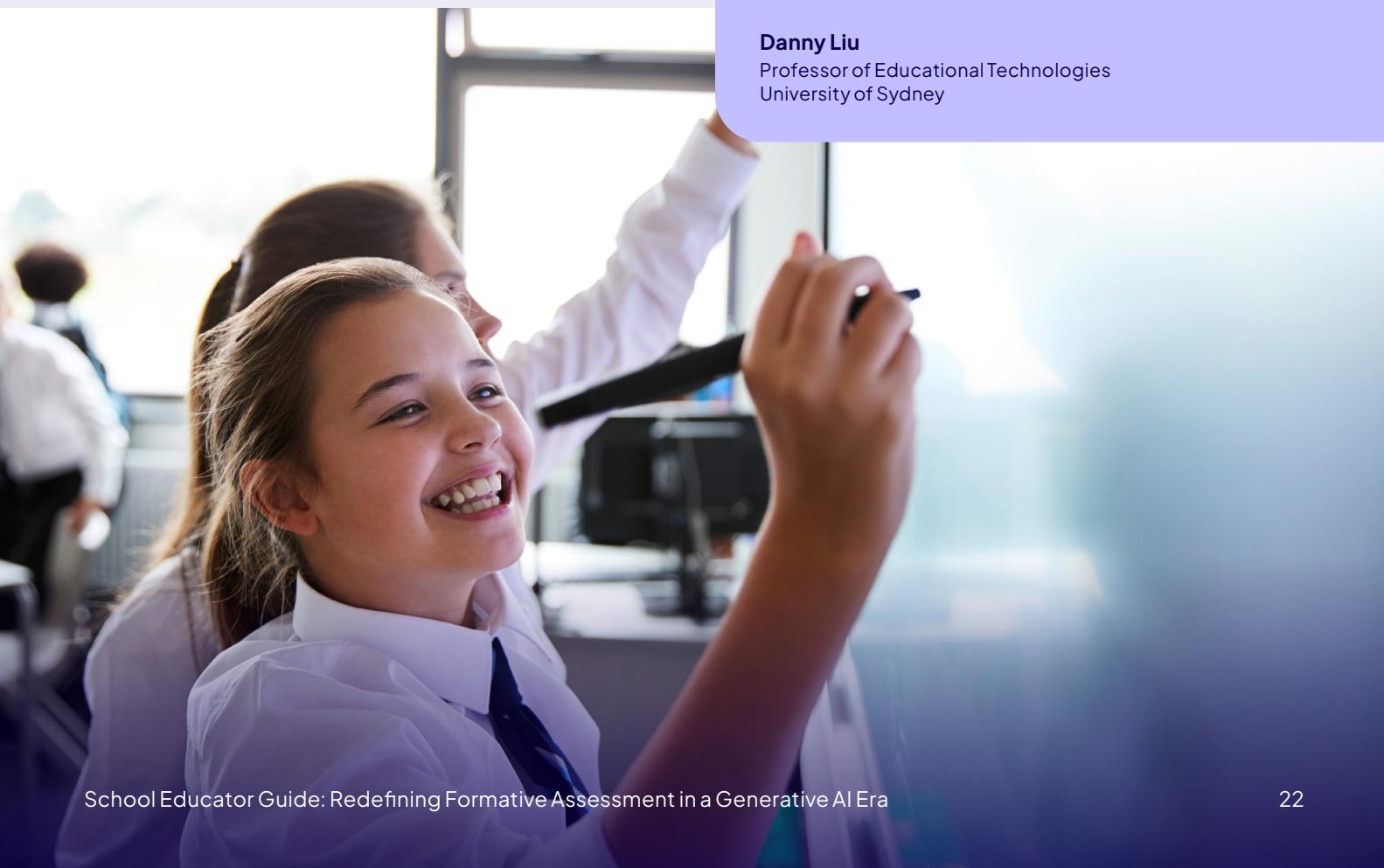
Acknowledgements

We would like to sincerely thank the educators who responded to our survey or generously gave their time to be interviewed for this project. Your thoughtful responses and perspectives have been invaluable in shaping this work.

Appendix

The following example provides ideas for updating the various components of an assessment task (e.g. instructions, structure, and evaluation) in order to intentionally integrate GenAI.

	Pre–GenAI task	Updated task
Target standard: Write narratives to develop real or imagined experiences or events using effective techniques, descriptive details, and clear event sequences		
Instructions	Write an imagined first-person account of what it would have been like during the volcanic eruption of Pompeii. Be sure to include effective techniques, descriptive details, and clearly-sequenced events.	Collaborate with an LLM to ethically and responsibly create an imagined first-person account of what it would have been like during the volcanic eruption of Pompeii.
Structure	Students create a first draft (out of class). Students share their draft and receive peer feedback (in class). Students revise their account based on peer feedback (out of class).	Teacher supervises students as they work in groups to use an LLM to generate imagined illustrations of the volcanic eruption at Pompeii to accompany their narrative. Students should test out different prompts and vote for the one that produces the best results (in class). Students select and sequence 3–5 of the AI-generated images to base their narrative on and record a voice memo explaining why they selected those (out of class). Students create a first draft of their narrative, weaving in details from the AI-generated images (out of class). Teacher supervises students as they use an LLM to generate feedback on their draft, in the guise of a friend hearing the first-person story and asking follow-up questions to clarify and collect more detail (in class.) Students revise their account based on their conversation with the chatbot (out of class). Students complete a 5–minute interview with the teacher about how using GenAI influenced their ideas and the final narrative (in class).

	Pre–GenAI task	Updated task
Reminder of target standard: Write narratives to develop real or imagined experiences or events using effective techniques, descriptive details, and clear event sequences		
Evaluation/ Artifacts	Final written account scored based on the level of detail included and clarity of the account.	"Exit tickets" or short reflections where students share their favorite prompt and one sentence explaining why. Students' sequencing of the images and recorded rationales for images selected. Students' first draft, especially quality of the details from the images included in the account. Students' chat logs from their conversation with an LLM showing their ability to answer questions about their account. Revised first-person account showing how AI-generated feedback was incorporated. Student reflections on the overall process, responsible and ethical attitudes toward GenAI use for creative writing.

Recommended resources

A Note on AI Basics

Note: This educator guide requires a basic understanding of Generative Artificial Intelligence (GenAI) and how it is being applied in education settings. For background on this topic, we recommend the following resources:

- Code.org: AI 101 for Teachers

- Teach AI: What is AI?

Endnotes

1   Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 327(1241), 475–484

2   Liu, D., & Bridgeman, A. (2023). What to do about assessments if we can't out-design or out-run AI. Teaching @ Sydney. https://educational-innovation.sydney.edu.au/teaching@sydney/what-to-do-about-assessments-if-we-cant-out-design-or-out-run-ai

3   Mormando, S. (2023). A stoplight model for guiding student AI usage. Edutopia. https://www.edutopia.org/article/creating-ai-usage-guidelines-students/

4   Perkins, M., Jasper, R., & Furze, L. (2025). Reimagining the Artificial Intelligence Assessment Scale: A refined framework for educational assessment. Journal of University Teaching and Learning Practice. https://doi.org/10.53761/rrm4y757

5   Organisation for Economic Co-operation and Development. (2025). Artificial intelligence literacy framework for primary and secondary education. https://www.oecd.org/en/about/projects/pisa-2029-media-and-artificial-intelligence-literacy.html

Pearson

This report marks the beginning of an important conversation on how assessment can evolve in the GenAI era. We'd love to hear your thoughts, both on this report and on how you're adapting formative assessments in your school. Share your feedback with us at **assessmentevolved@pearson.com**

Learn more at plc.pearson.com/assessmentevolved