Higher Ed Educator Guide

Redefining Formative Assessment in a Generative AI Era

Practical Strategies for Future-Ready Assessment

Designed for HE 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 cheating, there is a real opportunity to enhance **learning integrity** and build AI literacy in our institutions. 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 educators are already thoughtfully engaging with GenAI. 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 the experiences of educators in Higher Education institutions. 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 focused 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. We intend for this guide to give educators practical steps for reflecting on and evolving their assessments.

A snapshot of educator perceptions

Given how quickly GenAI is evolving, educators, understandably, have different perspectives on the technology, 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.

HE Educator Profiles

Protective Skeptics — **37%, n=194**

"Their learning declines because instead of thinking through a problem, they're just copying what AI tells them. There's no consolidation of information in their brain."

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 — **44%, n=232**

"I have deliberately included GenAI so that my students recognize it as a useful tool that can enhance their own efforts. During workshops, I have them use their AI product in specific ways that preclude them from using AI to draft entire papers or portions of papers. I treat GenAI as another research, drafting and editing tool."

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 — **6%, n=33**

"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 — **12%, n=65**

"The assessments gauge students' knowledge, retention, and application of theory. Using GenAI as a research tool would not drastically change that goal, however using GenAI to the extreme (copy and paste completely) would destroy that goal."

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.

- **Neutral** 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

"Lots of Higher Ed students are battling with using AI for short-term gains, knowing that they are at risk of losing fundamental skills because of it."

Pat Yongpradit
Chief Academic Officer of Code.org and Lead of TeachAI

"Are students becoming uncritical users of AI? Uncritical use doesn't necessarily have to connect to learning outcomes. It may be that they don't think about it much. Are students being irresponsible users of it, which is a bit different to uncritical. Irresponsible means, are they throwing in private data or are they throwing in things that they shouldn't into AI."

Danny Liu
Professor of Educational Technologies
University of Sydney

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 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 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 GenAImisuse. 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 all share a few key points:

- 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 with all the 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 Assessment Evolved report, we also recognize that substantial change cannot come without considerable support from both administrators 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 course, 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 the assessment unsupervised. Here, GenAI usage may be allowed, invited, or even required, depending on the course and/or particular assessment targets.

> "I think the answer comes down to what is the ultimate pedagogical goal of the class? What's the thing that you want the student to acquire?"
>
> **Ryan Baker**
> Professor of Artificial Intelligence and Education, University of South Australia and Director of the Penn Center for Learning Analytics

#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?

Most course syllabi outline goals like, "Acquire and synthesize scientific information from a variety of sources." To reflect GenAI, models like Kennedy HQ's AI Literacy Framework[5] (adapted from UNESCO' Digital Competence Framework) can help reframe assessment targets with a human-centered approach. This framework highlights domains like content creation ("use AI to create and personalize new content"), problem solving ("use AI to create insights and improve processes"), and career competencies ("combine industry expertise with AI analysis"). This lens updates the original goal to: "Critically use GenAI to acquire and synthesize scientific information from a variety of sources."

Reframing assessment targets to include AI proficiency requires explicit instruction on responsible AI use, including structured formative assessment activities that integrate AI use. For example, if students use GenAI for drafting, require a reflective commentary on what the AI contributed, what was revised, and where their own thinking is evident. This helps students think critically about AI outputs alongside their own work.

> "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

> "In the Computer Science course I teach, it's even more relevant. I think it's probably one of the areas where AI is the most disruptive. I would say AI is very good at writing code. And there we do really talk about what do you actually learn and we actually do a little experiment in class - some people use AI to do an assignment, some people don't use AI to do the assignment and we have a discussion around what they learn and their process of doing that. So, they really start to understand where AI is shortcutting learning and where it is not shortcutting learning."
>
> **Eric Klopfer**
> Professor, Director of the Scheller Teacher Education Program and Director, The Education Arcade, Massachusetts Institute of Technology

#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.

When assessment targets emphasize **how** students demonstrate what they know, learning becomes more authentic and better aligned with the real situations in which they'll need these abilities.

Authentic assessment doesn't mean recreating the workplace in your classroom. It simply asks students to apply their knowledge in new, meaningful situations. GenAI can support this process. When used to generate ideas, refine drafts, or seek feedback, it helps students develop both disciplinary skills and AI competence.

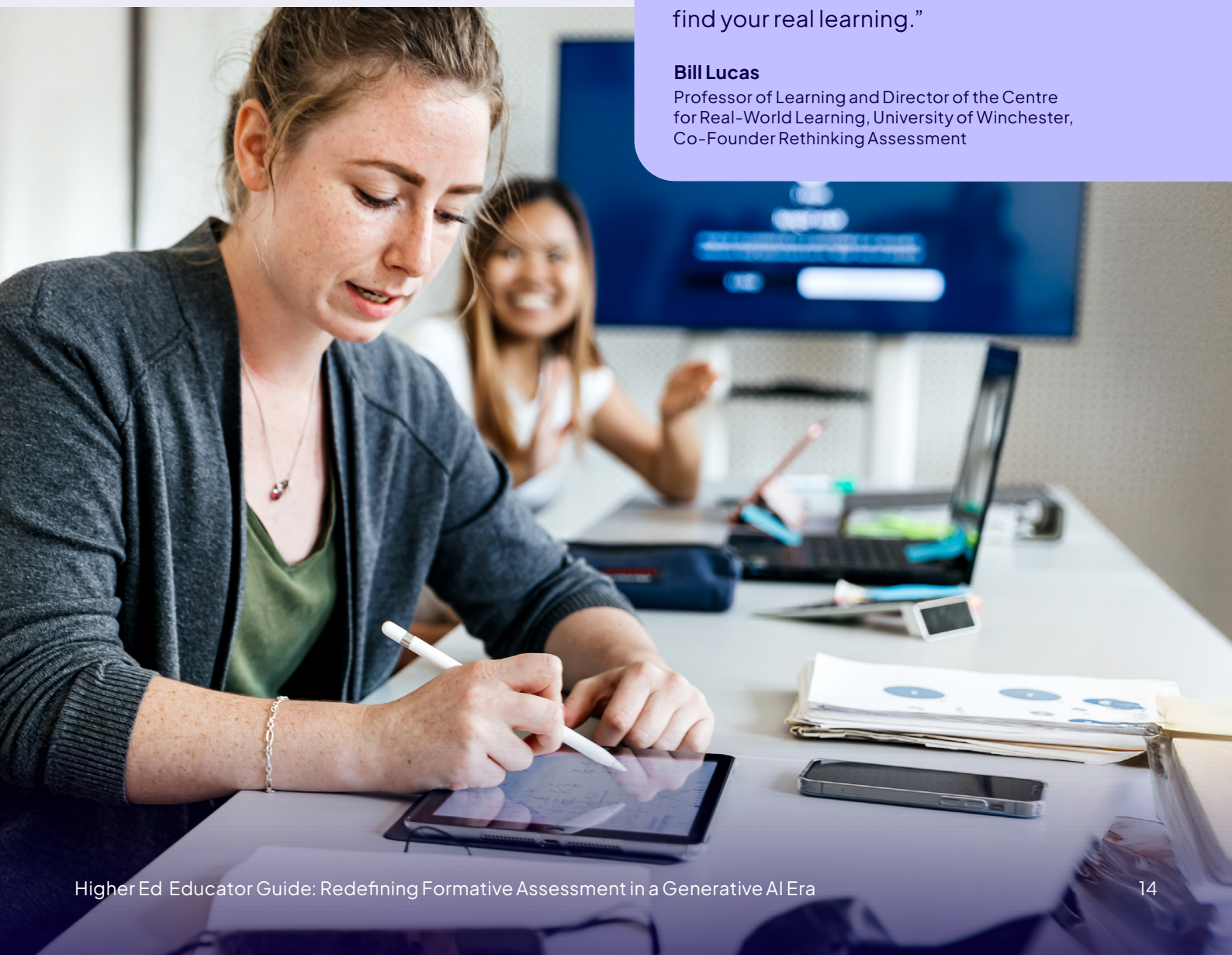
A process-focused approach often requires more structure: breaking an assignment into stages, adding checkpoints, or providing scaffolded steps. Part of the evaluation should reflect this process. Students may need to submit logs of their LLM interactions or brief notes showing how prompts, revisions, and feedback shaped their thinking. These checkpoints can occur in class or accompany the final submission.

> "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 futurein 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 instructor 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.

> 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

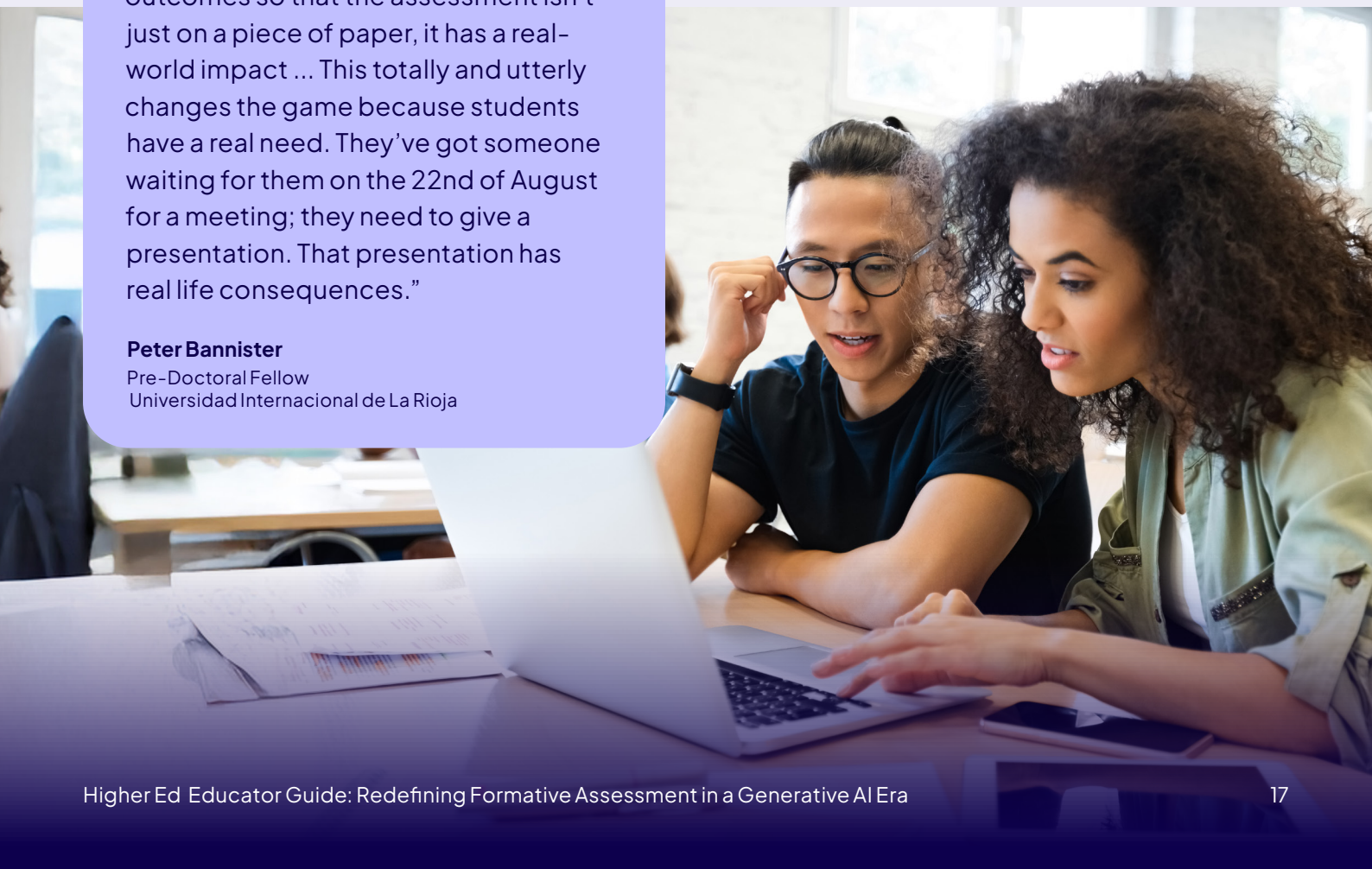
> Let me talk about the self-regulation of learning. Now they talk about the co-regulation of learning, acknowledging that others - teachers, peers, the book, the computer and the questions, the textbook - all of those things also help you regulate your learning. But you can't regulate learning until 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 of a theory or concept, prompt students to critique, compare, or apply concepts to a novel, real-world situation.

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.

> “I’m talking about collaborating with local stakeholders within the community, i.e., charities, schools, immigration centers, whoever we may like, and designing challenges, specific outcomes so that the assessment isn’t just on a piece of paper, it has a real-world impact ... This totally and utterly changes the game because students have a real need. They’ve got someone waiting for them on the 22nd of August for a meeting; they need to give a presentation. That presentation has real life consequences.”
>
> **Peter Bannister**
> Pre-Doctoral Fellow
> Universidad Internacional de La Rioja



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 for a nursing student might consist of an assignment where students must produce a care plan based on ethical nursing principles and evidence based practices. 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 educator 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 drafting process, a skill that will inevitably serve them well in future.

Figure 1. Traditional higher education nursing care plan task versus updated task ideas integrating GenAI

Approach	Task description
Target learning objective: Create care plans that are informed by ethical nursing principles and evidence-based practices	
Original task	Choosing one of **five** case studies shared by your instructor, identify two priority nursing diagnoses. Create a care plan for one of the **two** diagnoses drawing on your knowledge of ethical nursing principles and evidence-based interventions.
AI as a tool for critique or debate	Choose one of **five** case studies shared by your instructor. Identify two priority nursing diagnoses. Create a care plan for one of the two diagnoses, drawing on your knowledge of ethical nursing principles and evidence-based interventions. Then, ask a Large Language Model (LLM) to provide a care plan for said diagnoses. Submit a brief reflection on how your care plan was similar to or different from the suggestions from the LLM. Identify where the LLM supported your process and where it was incorrect or missing context. Provide a statement at the end of your care plan to indicate how and where you used GenAI and the role it (and you) played in developing the final product.

Approach	Task description
Reminder of target learning objective: Create care plans that are informed by ethical nursing principles and evidence-based practices	
Process over product	Submit all three parts of this assignment which follows the first three steps of the nursing process: 1 **Assessment:** Choose one of **five** incomplete case studies shared by your instructor. Identify what information is missing from the case studies that should be gathered from the patient, family members, and/or medical records to enable diagnosis. For any subjective sources of data, describe at least one method for validating the information. 2 **Diagnosis:** Review the updated case study provided by your instructor and compare it to your Step #1 list of missing information. Did you request all the relevant pieces of information? Did you request anything that you didn't need to? Based on the new information, identify two priority nursing diagnoses and provide a rationale for them by applying the prioritization framework you think is most relevant to this case. 3 **Planning:** Create a care plan for the diagnoses selected in Step #2 that includes at least two treatment goals and 3–4 nursing interventions. Submit the document from Steps 1–3 with version history enabled for the assessor. **Bonus:** Submit a short reflection about this process. Include what you enjoyed, what was difficult, etc.). In that reflection, indicate how and where you used GenAI, what it did well, and what it did not do well.
Multiple modalities	Choosing one of **five** case studies presented by your instructor, identify two priority nursing diagnoses. Create a written care plan for one of the **two** diagnoses, drawing on your knowledge of the ethical nursing principles and evidence-based interventions. In class, you will be asked to present your case, diagnoses, and care plan for questions in a 10–minute presentation.

Approach	Task description
Reminder of target learning objective: Create care plans that are informed by ethical nursing principles and evidence-based practices	
Requiring AI justification	Choosing one of **five** case studies shared by your instructor, identify two priority nursing diagnoses. Create a care plan for one of the **two** diagnoses, drawing on your knowledge of the ethical nursing principles and evidence-based interventions. If you would like to use GenAI, indicate why, where, and how you plan to use it. Requests to use AI are due at least one week before the assignment due date. Submit a reflection on your use of GenAI, including what GenAI did well and what it did not do well.
Real-world assessment	In class, you will be presented with a mock patient (a classmate who will roleplay one of **five** case studies, which you will have access to 48 hours in advance). You can ask up to five follow-up questions to the patient/classmate. In real time, identify two priority nursing diagnoses for the case. Using the next 45 minutes, create a care plan for one of the two diagnoses drawing on your knowledge of ethical nursing principles and evidence-based interventions.

Remember: While this example showcases some of the strategies we've outlined, there will be adjustments of varying degrees that you can make across the course. We encourage you to experiment, to consider how different approaches might work across a course sequence, 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 higher education.

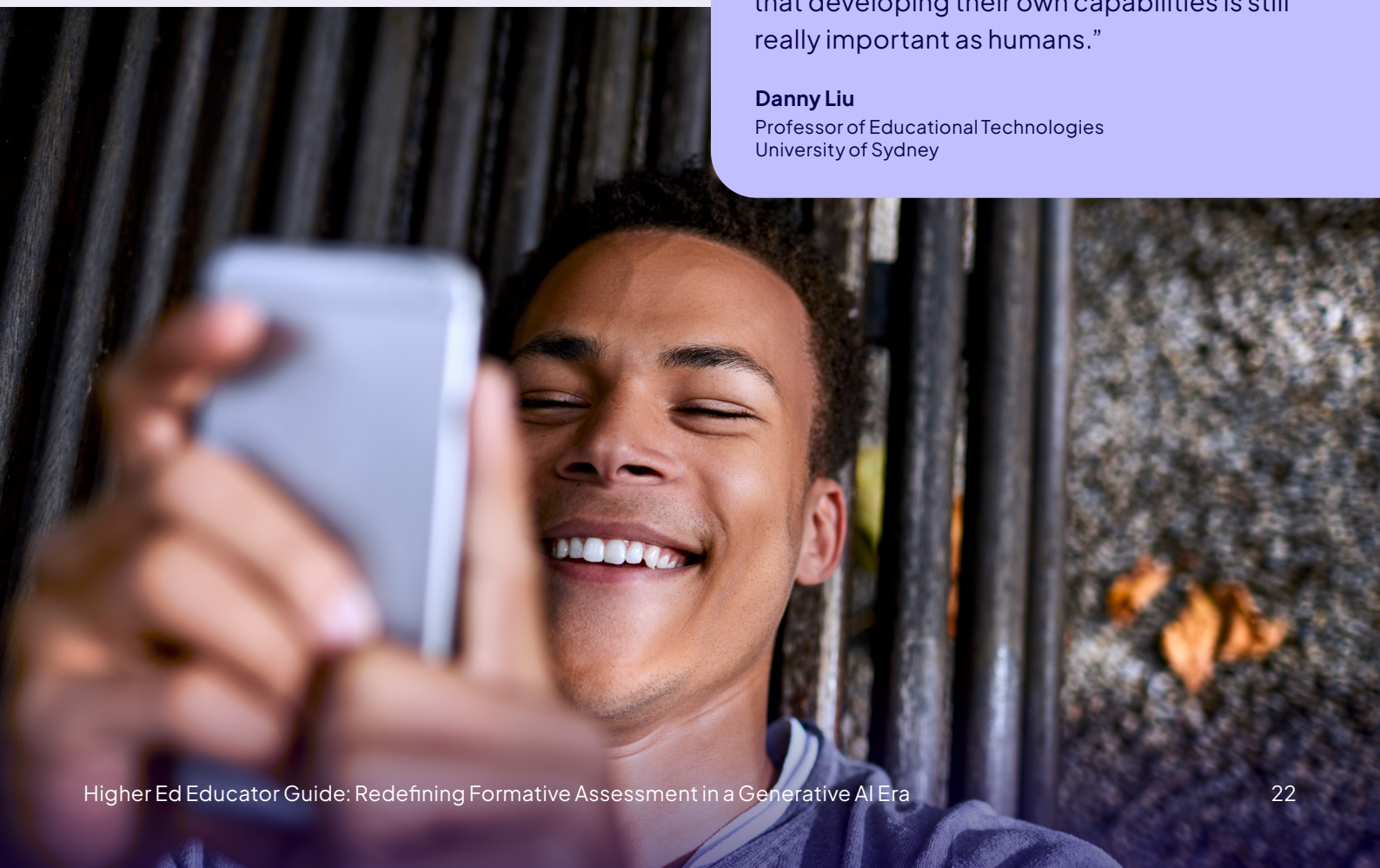
Students now have a host of AI tools at their fingertips, and while it is tempting to try to block their access to this technology in the name of learning, these students will eventually need to navigate careers and everyday lives in an AI-driven world. This places an upfront responsibility on educators 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 instructors 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.

Ryan Baker, Professor of Artificial Intelligence and Education, University of South Australia and Director of the Penn Center for Learning Analytics.

Peter Bannister, Pre-doctoral Fellow, Universidad Internacional de La Rioja.

Amanda Bickerstaff, Co-Founder and CEO of AI for Education.

Susan Brookhart, Professor Emerita, School of Education, Duquesne University.

Mutlu Cukurova, Professor of Learning and AI, University College London.

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

Rachel Evans, Director of Digital Transformation, Girls' Day School Trust.

Charles Fadel, Founder and Chairman, Center for Curriculum Redesign.

Steve Fitzpatrick, K12 Educator and AI Practitioner, Hackley School.

Eric Klopfer, Professor, Director of the Scheller Teacher Education Program and Director, The Education Arcade, Massachusetts Institute of Technology .

Danny Liu, Professor of Educational Technologies, University of Sydney.

Bill Lucas, Professor of Learning and Director for the Centre of Real World Learning at the University of Winchester and Co-Founder of Rethinking Assessment.

Kane Murdoch, Head of Complaints, Appeals and Misconduct, Macquarie University.

Mike Perkins, Head of the Centre for Research & Innovation, British University Vietnam.

Pat Yongpradit, Chief Academic Officer of Code.org and Lead of Teach AI.

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 learning objective: Create care plans that are informed by ethical nursing principles and evidence-based practices		
Instructions	Choosing one of **five** case studies shared by your instructor, identify two priority nursing diagnoses. Create a care plan for one of the **two** diagnoses drawing on your knowledge of the ethical nursing principles and evidence-based interventions.	Choose one of **five** case studies shared by your instructor. Follow the first three steps of the nursing process: Assess, Diagnose, and Plan, documenting your data and steps along the way.
Structure		**Part 1: Assess** **On your own**, review the case file and outline the most pertinent case information, any missing information, and any information that needs to be verified. **During class,** work in teams to interview your mock patient to verify patient information and gather any additional data needed to make a diagnosis. **Part 2: Diagnose** **On your own,** enter the pertinent information from the case file into an LLM and ask it to provide two priority diagnoses, explaining its reasoning step by step. Carefully evaluate these diagnoses. Submit your prompts, the output, and a written critique of the diagnoses. **During class**, work with your team to compare prompts and outputs from the LLM. As a group, document what worked well, and what didn't. **Part 3: Plan** **On your own,** create a care plan for one of the two diagnoses that includes at least two treatment goals and 3–4 nursing interventions. **During class,** present your case, diagnoses, and care plan for questions in a 10–minute presentation.

	Pre-GenAI task	Updated task
Reminder of target learning objective: Create care plans that are informed by ethical nursing principles and evidence-based practices		
Evaluation/Artifacts		Timestamped outline documenting case file analysis and missing information. Updates to the outline as more information is gathered. Annotated logs of LLM prompts and output, and a written critique of the diagnoses. Group analysis of pros and cons of using an LLM to help with diagnosis. Presentation on care plan, including responses to questions.

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:

- EDUCAUSE: Generative AI
- Stanford: Defining AI and Chatbots
- Harvard: Teach with Generative AI
- Vanderbilt: Faculty Guidance

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 Kennedy, K. (2023, December 21). AI literacy framework. Kennedy HQ. https://kennedyhq.com/wp/2023/12/21/ai-literacy-framework/

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 institution. Share your feedback with us at **assessmentevolved@pearson.com**

Learn more at
plc.pearson.com/assessmentevolved