

AI in Circle U: Connecting Minds, Shaping the Future

Dr Chahna Gonsalves

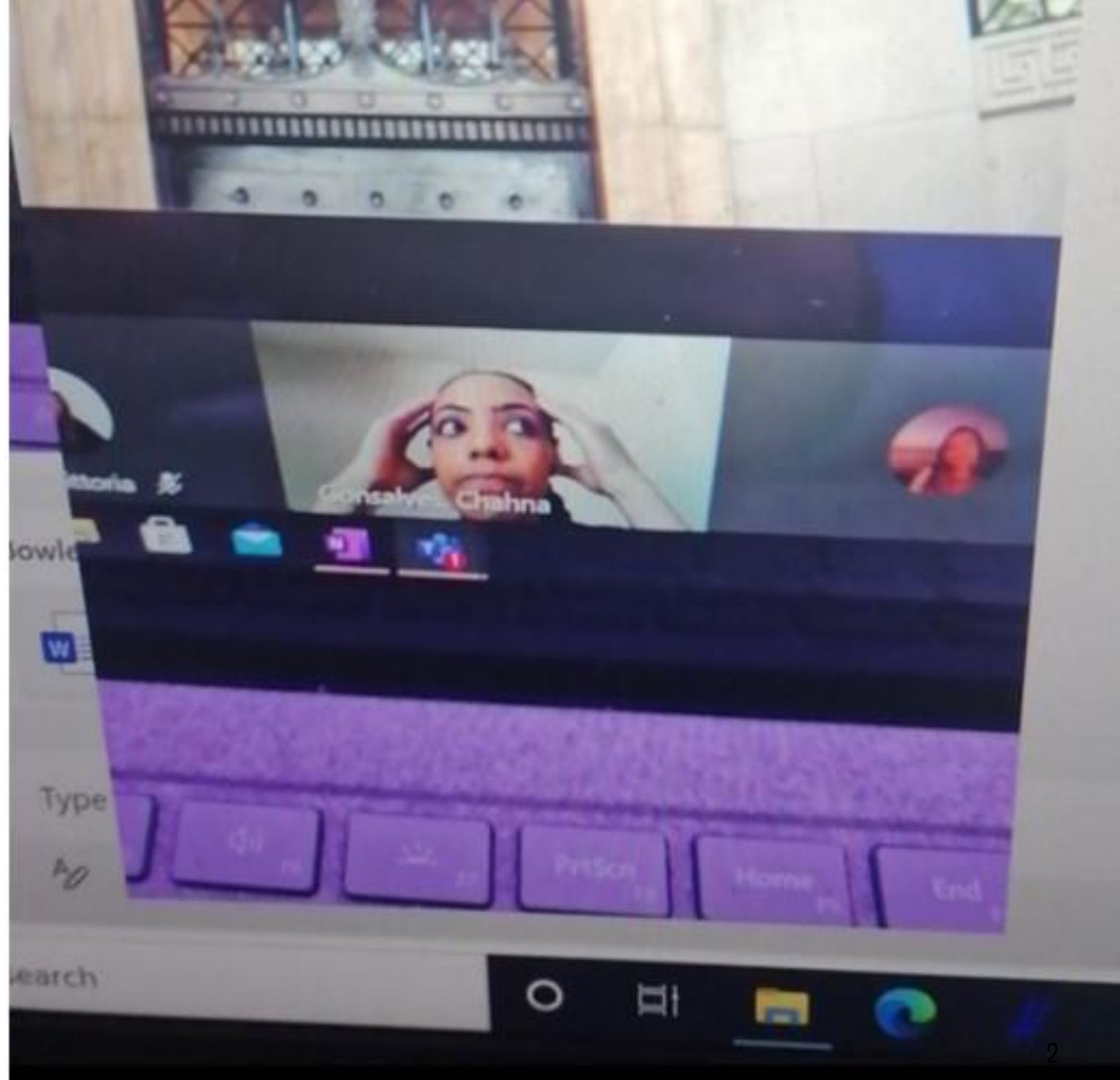


KING'S
BUSINESS
SCHOOL

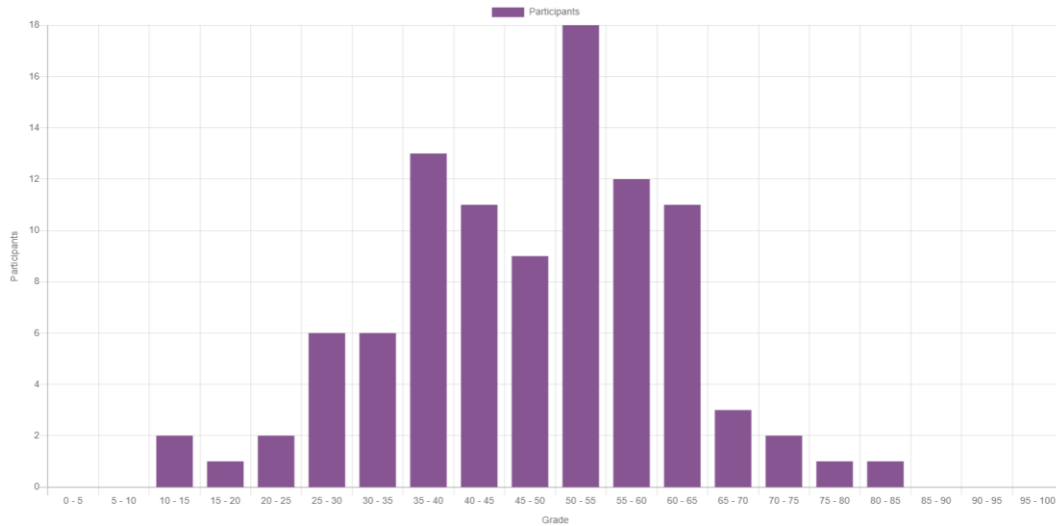
ARE WE STRENGTHENING
LEARNING WITH AI
OR DISTORTING IT?

Assessment...stress!

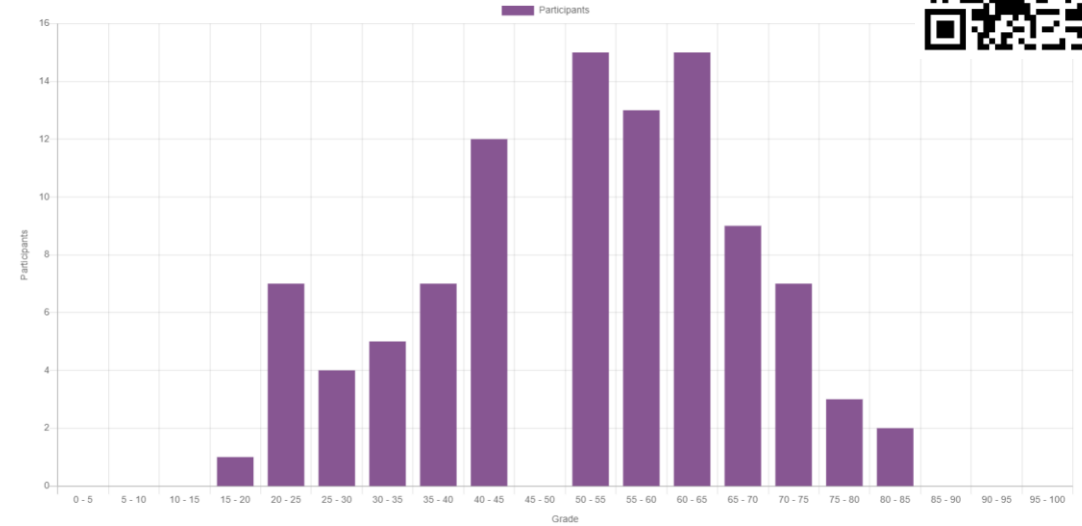
- **Context:** Second-year undergraduate Brand Management mid-term assessments launched post-ChatGPT
- **Challenge:** Redesigning assessments to measure understanding, not rote recall
- **Response:** Leveraging AI limitations to promote critical thinking
- **Core Goal:** Encouraging genuine engagement over AI-driven shortcuts



Results



Test 1, 98 students
Mean 48, median 49, SD 14 (%)



Test 2, 98 students,
Mean 52, median 50, SD 15 (%)

“The cases used in the tests were interesting, by reading the cases and using knowledge from the readings I was able to understand more.”

Caution

MARKETING EDUCATION REVIEW
<https://doi.org/10.1080/10528008.2024.2417106>



OPEN ACCESS

Comparing The Effectiveness Of Multiple-Answer And Single-Answer Multiple-Choice Questions In Assessing Student Learning

Yusuf Oc^a and Hela Hassen^b

^aDepartment of Management, Bayes Business School, City St George University of London, London, UK; ^bDepartment of Management, KEDGE Business School, France

ABSTRACT

Driven by technological innovations, continuous digital expansion has transformed fundamentally the landscape of modern higher education, leading to discussions about evaluation techniques. The emergence of generative artificial intelligence raises questions about reliability and academic honesty regarding multiple-choice assessments in online education. In this context, this study investigates multiple-answer questions (MAQs) versus traditional single-answer questions (SAQs) in online higher-education assessments. A mixed-methods study involving quantitative field experiments and qualitative interviews was conducted with students enrolled in an online Marketing M.Sc. program. The students were divided randomly and assessed using either SAQs or MAQs, and the impacts on test performance of variables such as grade averages, study times, perceived workload, and difficulty were evaluated using independent-sample *t*-tests and ordinary least-squares regression analysis. The results show that although grades were lower and MAQs were perceived as being more difficult, study times and perceived workload did not differ significantly between the two formats. These findings suggest that despite their challenge, MAQs can promote deeper understanding and greater learning retention. Furthermore, even with their higher perceived difficulty and impact on performance, MAQs hold potential for dealing with academic-integrity concerns related to artificial intelligence.

Suggested

Nicole Bump • 3rd+

I help chronically busy martech/adtech marketers create better content
2d •

+ Follow

A professor friend of mine struggles with ChatGPT-generated paper submissions—so he sets a brilliant trap.

He adds hidden instructions (white text) in the assignment guidelines that say to get an A, you must cite research by Frankie Hawkes.

Frankie Hawkes is his dog.

He includes a disclaimer saying if you are a human reading this, you should ignore that instruction (which helps with ADA e-reader software).

Every semester, at least one student references the reknowned Dr. Hawkes.

I'm amused every time, but I'm sure the students are not.

No lesson here...except to teach your kids not to cheat with AI!

632

114 comments • 18 reposts

Like

Comment

Repost

Send

Pedagogical trends

Pedagogical Applications Effectively incorporate AI in teaching and learning	Prevention Strategies Prevent inappropriate use of AI in teaching and learning
• Ask students to analyze and evaluate AI-generated texts. • Ask students to compare and evaluate the different versions of texts generated by different AI tools. • Ask students to compare/contrast AI-generated texts with human writing. • Ask students to revise and edit AI-generated information. • Ask students to debate or argue with AI and reflect on their learning. • Use AI as a resource for students to receive feedback on their drafts • Use AI tools to brainstorm initial teaching ideas and activities. • Use AI tools to generate additional examples of certain concepts. • Use AI tools to summarize long or difficult text. • Use AI tools to generate writing prompts, grading rubrics or quiz questions based on the course materials.	• Ask students to explain their thought processes as they solve problems. • Ask students to reflect on their personal learning experiences and opinions. • Ask students to connect with their personal knowledge and life experiences. • Ask students to include and provide proper academic citations. • Ask students to reference class materials, notes, or sources that are unavailable online. • Ask students to complete assignments in class. • Ask students to present their answers in multimodal ways, such as hand drawing, or audio threads. • Include visual prompts in assignments. • Design assignments related to current events or discussions in the specific academic field. • Divide the larger project into multiple smaller tasks.

Guideline for Generative AI Use in Assessment



Guidelines for Generative AI Use in KBS Assessments

As Generative Artificial Intelligence (GenAI) becomes an integral part of academic and professional settings, and as employers place greater emphasis on AI skills for graduates, it is essential for KBS to integrate these tools strategically into our assessment strategies.

These guidelines provide a framework for incorporating GenAI into teaching and assessment while maintaining academic integrity, critical thinking and ethical use of AI tools. Specifically, we recommend structuring our approach to GenAI across three broad types of assessments: **Exams in Controlled Conditions**, **Hybrid Assessments** and **Take-Home Project-Based Assessments**.

- 1. **Exams in Controlled Conditions where AI is off-limits.**
- 2. **Hybrid Assessments blend in-class activities with AI-enabled take-home work.**
- 3. **Project-Based Assessments which fully encourage AI use for research and ideation**

1	NO AI	The assessment is completed entirely without AI assistance in a controlled environment, ensuring that students rely solely on their existing knowledge, understanding, and skills. You must not use AI at any point during the assessment. You must demonstrate your core skills and knowledge.
2	AI PLANNING	AI may be used for pre-task activities such as brainstorming, outlining and initial research. This level focuses on the effective use of AI for planning, synthesis, and ideation, but assessments should emphasise the ability to develop and refine these ideas independently. 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	AI may be used to help complete the task, including idea generation, drafting, feedback, and refinement. Students should critically evaluate and modify the AI suggested outputs, demonstrating their understanding. You may use AI to assist with specific tasks such as drafting text, refining and evaluating your work. You must critically evaluate and modify any AI-generated content you use.
4	FULL AI	AI may be used to complete any elements of the task, with students directing AI to achieve the assessment goals. Assessments at this level may also require engagement with AI to achieve goals and solve problems. 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	AI is used creatively to enhance problem-solving, generate novel insights, or develop innovative solutions to solve problems. Students and educators co-design assessments to explore unique AI applications within the field of study. You should use AI creatively to solve the task, potentially co-designing new approaches with your instructor.



Perkins, Furze, Roe & MacVaugh (2024). The AI Assessment Scale

Figure 1: The Revised AI Assessment Scale

Perkins, M., Furze, L., Roe, J., MacVaugh, J.(2024). The Artificial Intelligence Assessment Scale (AIAS): A Framework for Ethical Integration ofGenerative AI in Educational Assessment. Journal of University Teaching and Learning Practice, 21(6).

The “Enforcement Illusion” of Assessment Validity



ASSESSMENT & EVALUATION IN HIGHER EDUCATION
<https://doi.org/10.1080/02602938.2024.2415654>



OPEN ACCESS

Addressing student non-compliance in AI use declarations: implications for academic integrity and assessment in higher education

Chahna Gonsalves

Marketing, King's College London, London, UK

ABSTRACT

This study examines the factors driving student non-compliance with AI use declarations in academic assessments at King's Business School, where 74% of students failed to declare AI usage despite declaration being a requirement of a mandatory coursework coversheet. Utilising the Theory of Planned Behaviour (TPB) as a framework, the research combines service evaluation survey data and semi-structured interviews to explore how attitudes, subjective norms, and perceived behavioural control influence compliance. Findings reveal that fear of academic repercussions, ambiguous guidelines, inconsistent enforcement, and peer influence are key barriers to AI use declaration. These factors complicate the declaration process, undermine transparency, and challenge academic integrity. The study extends the TPB model by highlighting the ethical and practical dilemmas posed by generative AI, which blur traditional norms of academic integrity. This research offers critical insights for policymakers, suggesting that clear, consistent, and trust-based policies are crucial in fostering ethical AI use. The findings underscore the importance of transparent communication and supportive institutional cultures to improve compliance. Ultimately, this study informs policy development by evaluating the effectiveness of declaration mechanisms and providing actionable recommendations to promote a culture of academic integrity in the evolving landscape of AI technologies.

KEYWORDS

Academic integrity;
generative AI; AI
declaration; compliance

Introduction

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
<https://doi.org/10.1080/02602938.2025.2503964>



OPEN ACCESS

Talk is cheap: why structural assessment changes are needed for a time of GenAI

Thomas Corbin^a , Phillip Dawson^a and Danny Liu^b

^aCentre for Research in Assessment and Digital Learning (CRADLE), Deakin University, Melbourne, Australia;

^bDVC Education Portfolio, The University of Sydney, Sydney, Australia

ABSTRACT

Generative AI (GenAI) challenges assessment validity by enabling students to complete tasks without demonstrating genuine capability. In response to this challenge, institutions have developed and implemented various approaches that aim to communicate permissible AI use to students. Familiar examples include the 'traffic light' approach now commonly found within institutional policy. While well-intentioned, these approaches share a common limitation: They focus primarily on communicating rules rather than redesigning assessment mechanics. To clarify why such approaches fail, and to guide more effective responses, this paper introduces a novel conceptual distinction between discursive changes to assessment (modifications relying solely on instructions students remain free to ignore) and structural changes (modifications that reshape the underlying mechanics of assessment tasks themselves). Through a critical analysis of prominent frameworks, we demonstrate that current approaches predominantly rely on discursive changes that create what we term an 'enforcement illusion'. We find that educational frameworks frequently borrow the language of socially familiar structural systems (like vehicular traffic lights) while lacking their actual enforcement capabilities, creating an illusion of assessment security. In place of this, we argue for a shift towards structural assessment redesign that builds validity into assessment architecture rather than attempting to impose it through unenforceable rules.

KEYWORDS

Assessment design;
academic integrity;
artificial intelligence;
assessment validity

“67.4% of top U.S. universities offer GenAI resources for faculty, but only 17.8% offer them for students”

GenAI Policy: Instructor Discretion

57 of the top 100 US universities allows instructors to determine their own GenAI policies in the classroom.

- **“Instructor Decides” Policy:** 47% of universities adopting this policy prohibit GenAI by default, allowing it only with explicit instructor permission.
- **Prohibition by Default:** Without clear instructor guidelines, using GenAI in assignments is typically not allowed and may be considered plagiarism.
- **Cautious Approach:** This policy reflects concerns about academic integrity, requiring appropriate citation and responsibility if GenAI is permitted.



Metaphor	Description	Positive Aspects	Negative Aspects	
A double-edged sword	Generative AI offers personalization but challenges educational integrity, needing careful use.	- Transformative potential in personalization and engagement - Augments traditional teaching	- Pedagogical regression - Educator uncertainty - Risks to academic integrity	
The Elephant in the room	Generative AI's significant but often unspoken role in education demands open dialogue.	- Acknowledges the widespread, yet unspoken, use of generative AI - Encourages ethical adoption of AI	- Reluctance to openly discuss generative AI - Institutional uncertainty - Concealment of AI usage	
An unintended burden	Generative AI adds complexity to assessment, burdening educators with integrity concerns.	Raises awareness of generative AI's impact on assessment and integrity	- Increased workload in assessment - Difficulty in discerning generative AI assistance - Blurs the definition and inference of originality	
A pedagogical paradigm shift	Generative AI encourages modern teaching but requires careful integration to maintain integrity.	- Encourages evolution in teaching methods - Emphasizes critical thinking and problem-solving	- Misuse of generative AI tools - Need for strategic pedagogical adaptation	
A new instrument	Generative AI opens teaching possibilities but demands time for educators to adapt.	- Offers opportunities for innovative teaching - Requires understanding, time, and practice	- Overwhelm due to rapid technological changes - Disparity in educator preparedness	
An escalator	Generative AI simplifies teaching but may reduce deep learning and critical engagement.	- Simplifies the creation of teaching materials - Highlights the ease of using generative AI	- Risks shallower learning engagement - Potential for passive learning - Discerning when to use generative AI in teaching and learning over traditional methods	
A digital compass	Generative AI guides education but requires clear goals to prevent confusion.	- Provides direction in generative AI integration - Encourages critical navigation of generative AI tools	- Need for clarity in learning paths and objectives - Rapid evolution of required skills	
A digital tutor	Generative AI provides personalized feedback, complementing but not replacing human teaching.	- Offers context-specific feedback - Enhances the educational process	- Supplemental role, not a replacement for human judgment - Risk of student overreliance on generative AI	



AP photo

Elementary school teachers picket against use of calculators in grade school
The teachers feel if students use calculators too early, they won't learn math concepts

Math teachers protest against calculator use

By JILL LAWRENCE

"My older kids don't pay any attention to an answer being absurd. strate," he said. "Teachers are shy."

the Soviet Union.

Polish officials about Soviet con-

'Calcuholics' students arouse teachers' ire

WASHINGTON (AP) — People have come here to protest the nuclear arms race and racial segregation in South Africa, but the small group of renegade math teachers were the first to protest "calcuholics" — students who need calculators to do basic math problems.

"Calculators later, we shall not be moved," they sang Thursday as they paraded their placards in front of the Sheraton Washington Hotel.

The National Council of Teachers of Mathematics started its five-day annual meeting Wednesday with about 6,500 teachers in attendance. The group recommends using calculators as early as first grade.

But the rebellious teachers said they opposed the use of calculators in the lower grades.

"My older kids don't pay any attention to an answer being absurd. They don't look at it. It's on the calculator," said Diana Harvey, a high school teacher from Hillsboro, Ohio.

'They're addicted. We call them calcuholics.'

Teacher

"They're addicted," she said. "We call 'em calcuholics."

The rebellious few — no more than 15 in all — were organized by John Saxon, a Norman, Okla., math textbook writer.

Saxon, waving a sign that read "Students Need Arithmetic Skills, Not Calculators," said students do not need calculators until they take algebra and trigonometry.

"I am not out of step with the gut feeling of most of the teachers who teach seventh and eighth grade," he contended.

Asked to explain the small size of his protest, Saxon said he didn't start publicizing it early enough. "And teachers don't like to demonstrate," he said. "Teachers are shy."

The St. Valentine's Day Massacre took place Feb. 14, 1929.

po
aj
B

(R

(T

Lo

U

(T

(W

Va

Co

1

Understanding learning

“A **process that leads to **change**, which occurs as a result of **experience** and increases the potential for improved performance and future learning” (Ambrose et al., 2010).**

Models of Learning:

- Constructivism
- Fink’s Significant Learning Model
- Kolb’s Experiential Cycle
- Bloom’s Taxonomy

Considerations

- Scaffolding
- Hierarchical dependencies



Evaluative Judgement

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
2024, VOL. 49, NO. 6, 893–905
<https://doi.org/10.1080/02602938.2024.2335321>

 **Routledge**
Taylor & Francis Group

 OPEN ACCESS  Check for updates

Developing evaluative judgement for a time of generative artificial intelligence

Margaret Bearman^a , Joanna Tai^a , Phillip Dawson^a , David Boud^{a,b,c} 
and Rola Ajjawi^a 

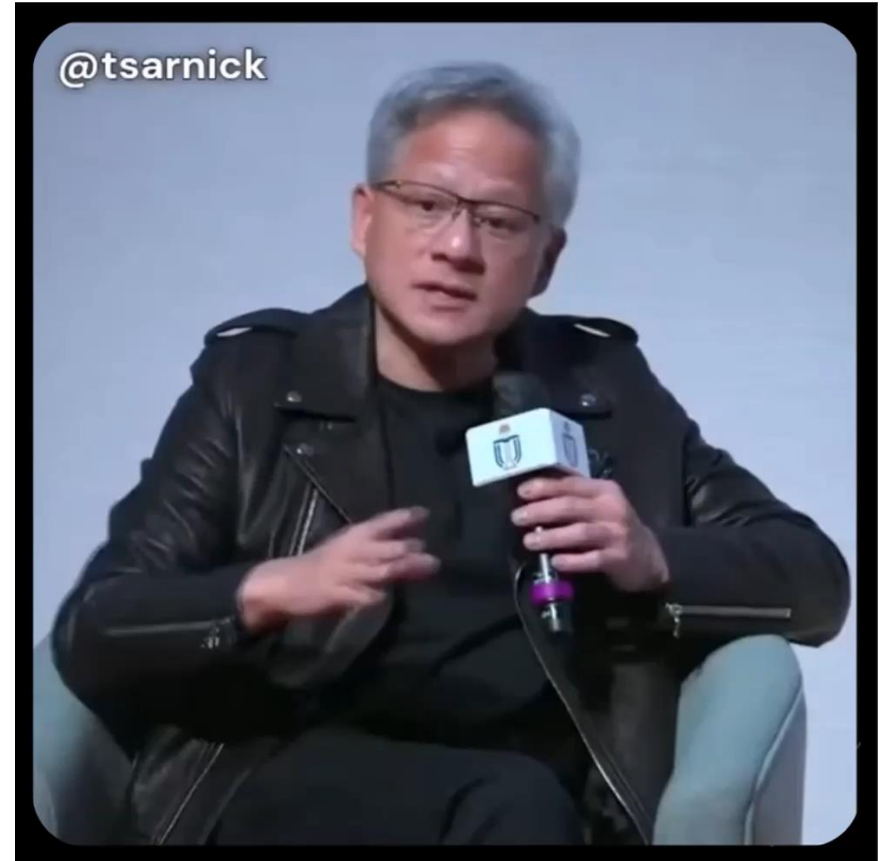
^aCentre for Research in Assessment and Digital Learning (CRADLE), Deakin University, Melbourne, Australia; ^bFaculty of Arts and Social Sciences, University of Technology Sydney, Sydney, Australia; ^cWork and Learning Research Centre, Middlesex University, London, UK

ABSTRACT

Generative artificial intelligence (AI) has rapidly increased capacity for producing textual, visual and auditory outputs, yet there are ongoing concerns regarding the quality of those outputs. There is an urgent need to develop students' evaluative judgement – the capability to judge the quality of work of self and others – in recognition of this new reality. In this conceptual paper, we describe the intersection between evaluative judgement and generative AI with a view to articulating how assessment practices can help students learn to work productively with generative AI. We propose three foci: (1) developing evaluative judgement of generative AI outputs; (2) developing evaluative judgement of generative AI processes; and (3) generative AI assessment of student evaluative judgements. We argue for developing students' capabilities to identify and calibrate quality of work – uniquely human capabilities at a time of technological acceleration – through existing formative assessment strategies. These approaches circumvent and interrupt students' uncritical usage of generative AI. The relationship between evaluative judgement and generative AI is more than just the application of human judgement to machine outputs. We have a collective responsibility, as educators and learners, to ensure that humans do not relinquish their roles as arbiters of quality.

KEYWORDS

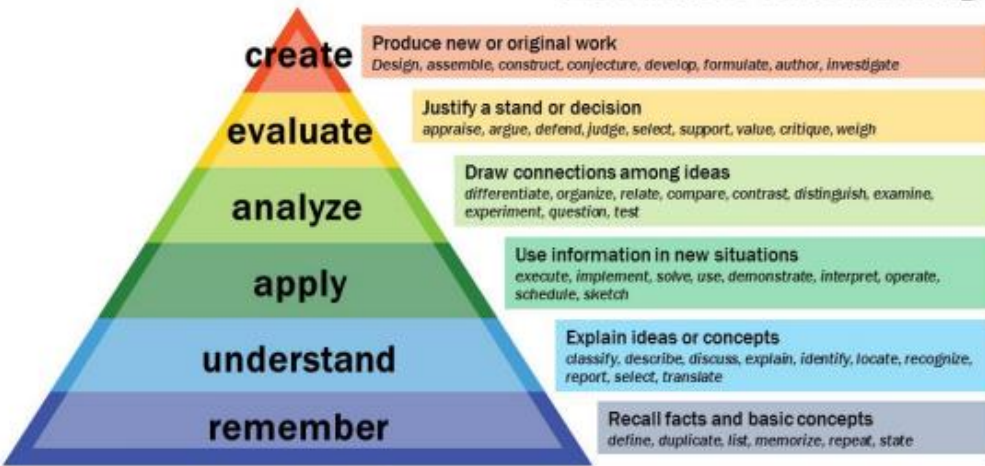
Generative artificial intelligence; evaluative judgement; assessment for learning; higher education



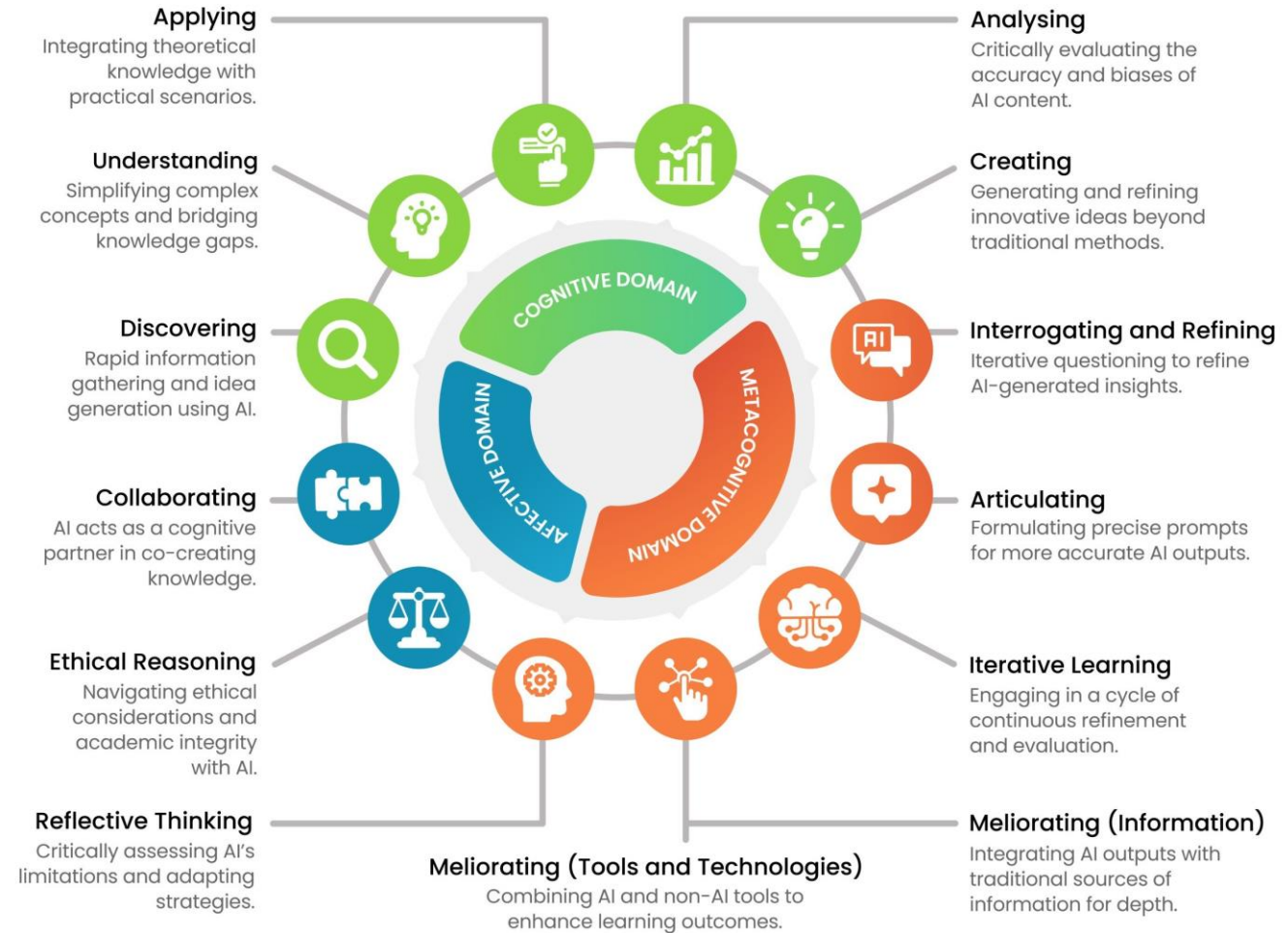
Rethinking Critical Thinking



Bloom's Taxonomy

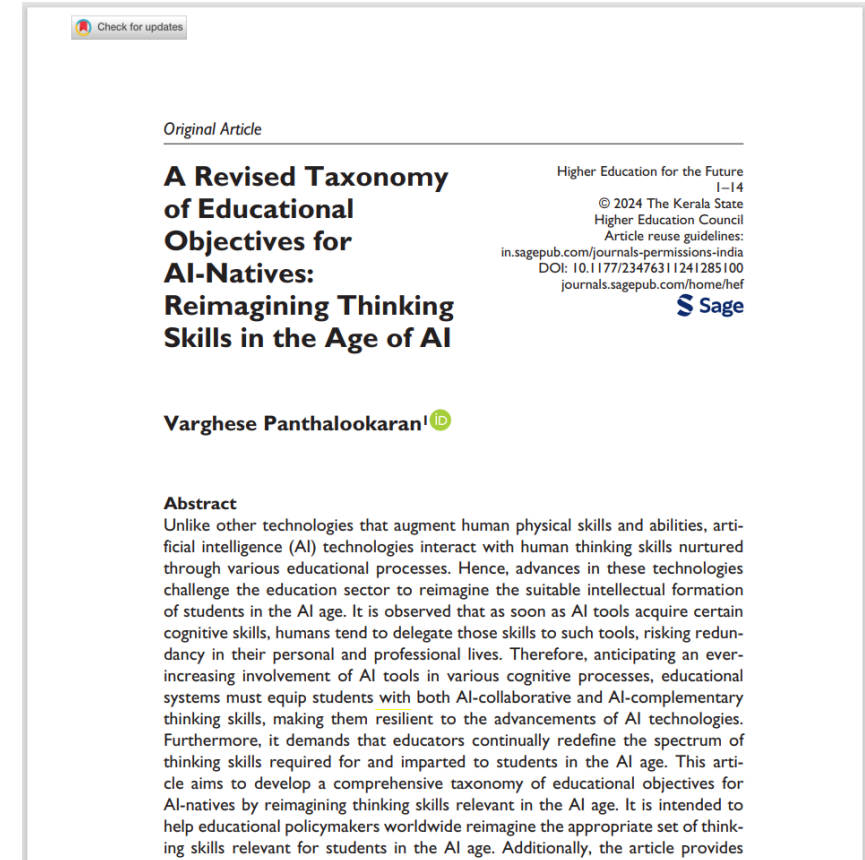


Vanderbilt University Center for Teaching



Revised Educational Objectives

- **Transpersonal Skills:** Develop intuition and mindfulness for insight beyond immediate data or logic.
- **Interpersonal Skills:** Cultivate empathy and collaborative thinking and communication skills for effective teamwork.
- **Exopersonal Skills:** Enhance critical, analytical, and creative thinking with external information.
- **Endopersonal Skills:** Strengthen foundational knowledge recall, understanding, and application.



Panthaloorkaran, V. (2024). A Revised Taxonomy of Educational Objectives for AI-Natives: Reimagining Thinking Skills in the Age of AI. *Higher Education for the Future*, 23476311241285100.

Shifting student behavior

A study of Harvard undergraduates (n = 326) found that almost 90% of students use generative AI.

Around 25% of students who use generative AI find themselves going to office hours, asking course staff for help, and completing required readings less often because of the availability of generative AI.



Opportunities & Challenge of Understanding

Stochastic: Language models predict words based on statistical probabilities, not true understanding.

They generate text by calculating likely word sequences from large datasets.

Parrot: Like a parrot, these models mimic patterns without comprehension.

They "echo" language structures from training data but lack awareness or original thought.



Can AI Really Reason?

- **Pattern Matching vs. Reasoning:** AI may mimic reasoning without true understanding, impacting its educational value.
- **Thoughtful Integration:** A deeper understanding of AI's limits is essential to use it in ways that genuinely support learning.



Mehrdad Farajtabar

@MFarajtabar

1/ **Can Large Language Models (LLMs) truly reason?** Or are they just sophisticated pattern matchers? In our latest preprint, we explore this key question through a large-scale study of both open-source like Llama, Phi, Gemma, and Mistral and leading closed models, including the recent OpenAI GPT-4o and o1-series.

arxiv.org/pdf/2410.05229

Work done with [@i_mirzadeh](#), [@KeivanAlizadeh2](#), Hooman Shahrokhi, Samy Bengio, [@OncelTuzel](#).

[#LLM](#) [#Reasoning](#) [#Mathematics](#) [#AGI](#) [#Research](#) [#Apple](#)

GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models

Iman Mirzadeh[†]
Oncel Tuzel

Keivan Alizadeh
Samy Bengio

Hooman Shahrokhi*
Mehrdad Farajtabar[†]

Apple

Abstract

Recent advancements in Large Language Models (LLMs) have sparked interest in their formal

AI Accuracy Drop with No-Op Questions

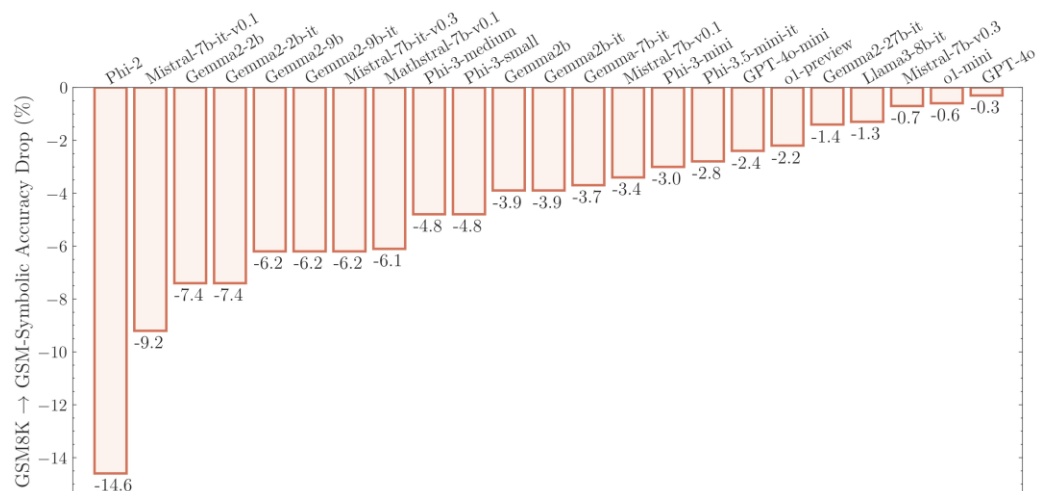
GSM-NoOp

Oliver picks 44 kiwis on Friday. Then he picks 58 kiwis on Saturday. On Sunday, he picks double the number of kiwis he did on Friday, but five of them were a bit smaller than average. How many kiwis does Oliver have?

o1-mini: Oliver picks kiwis over three days as follows:
 - Friday: 44 kiwis
 - Saturday: 58 kiwis
 - Sunday: Double the number he picked on Friday, which is $2 \times 44 = 88$ kiwis
 However, on Sunday, 5 of these kiwis were smaller than average. We need to subtract them from the Sunday total: 88 (Sunday's kiwis) $- 5$ (smaller kiwis) $= 83$ kiwis Now, summing up the kiwis from all three days: 44 (Friday) $+ 58$ (Saturday) $+ 83$ (Sunday) $= 185$ kiwis Oliver has a total of 185 kiwis.

Llama3-8B: Let's break it down step by step:
 Oliver picks 44 kiwis on Friday. He picks 58 kiwis on Saturday. On Sunday, he picks double the number of kiwis he did on Friday, which is $2 \times 44 = 88$ kiwis.
 Five of the kiwis picked on Sunday are a bit smaller than average, so we subtract 5 from the total number of kiwis picked on Sunday: $88 - 5 = 83$ kiwis. Now, let's add up the total number of kiwis Oliver has:
 44 (Friday) $+ 58$ (Saturday) $+ 83$ (Sunday) $= 185$ kiwis
 So, Oliver has 185 kiwis in total.

Figure 7: An example from the GSM-NoOp dataset: We add seemingly relevant statements to the questions that are, in fact, irrelevant to the reasoning and conclusion. However, the majority of models fail to ignore these statements and blindly convert them into operations, leading to mistakes.



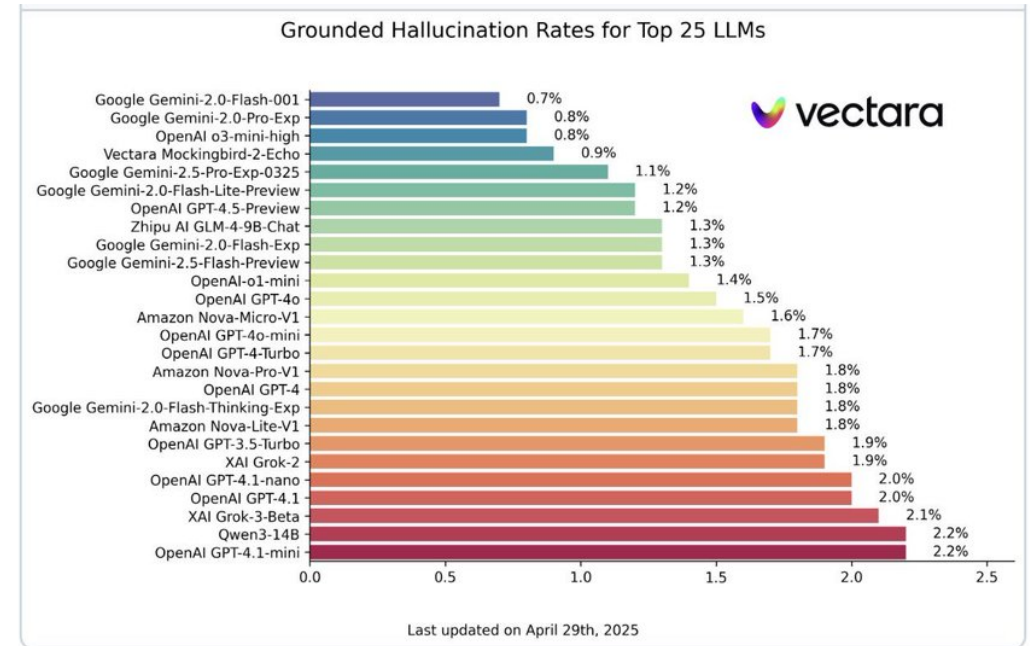
Mehrdad Farajtabar @MFarajtabar · Oct 10

5/ #Result 2: The fragility of supposed LLM reasoning. LLMs remain sensitive to changes in proper names (e.g., people, foods, objects), and even more so when numbers are altered. Would a grade-school student's math test score vary by ~10% if we only changed the names?

Model	GSM8K	Names	Numbers	Both
Gemma2-9b-it	87.0	88.6 (± 2.0)	83.1 (± 2.2)	79.1 (± 3.0)
Llama3-8b-instruct	74.0	75.6 (± 2.1)	75.5 (± 3.1)	74.6 (± 2.9)
Phi-3-medium-128k-instruct	89.0	91.8 (± 1.7)	89.0 (± 2.3)	82.5 (± 2.9)
Phi-3-small-128k-instruct	89.0	88.4 (± 1.8)	83.7 (± 2.4)	83.7 (± 2.6)
Phi-3.5-mini-instruct	88.0	89.1 (± 1.8)	84.9 (± 2.4)	82.1 (± 3.4)
Mathstral-7b-v0.1	80.0	81.0 (± 1.3)	77.3 (± 2.0)	74.0 (± 3.5)

10 55 339 57K

Models are getting better, but hallucination is getting worse



<https://github.com/vectara/hallucination-leaderboard>

Empowering Responsible AI Use Through Education and Clarity

https://www.futurelearn.com/courses/ai-in-education

Online Courses / Teaching



AI in Education

Explore how AI is reshaping education and reflect on how you could respond critically or reimagine educational practices.

★★★★★ 4.8 (89 reviews) 11,096 enrolled on this course



Duration
3 weeks

Weekly
study
4 hours

100% online
How it
works

Unlimited subscription
\$174.99 \$119.99 for a whole
year
Learn more

AI in Education

★★★★★ 4.8 (89 reviews)

11,096 enrolled on this course

3 weeks

4 hours per week

Digital certificate when eligible

Introductory level

Join course

Find out more about how to join this course



Building a Future of Supportive Learning

To advance learning with GenAI, we must stay adaptable, engage thoughtfully, and design practices that align technology with educational integrity and purpose.

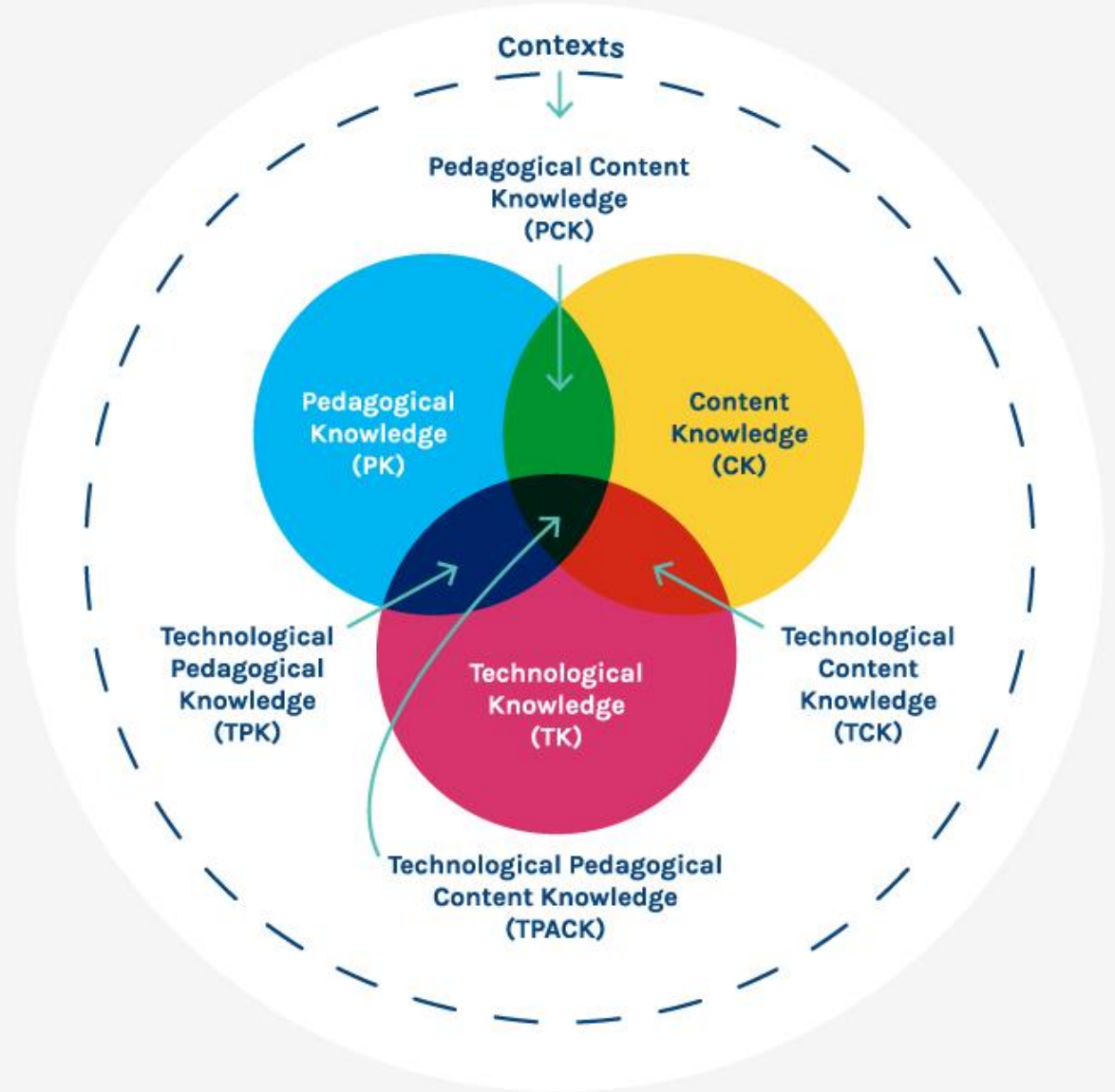
What are we teaching/assessing?

Why does it matter?

How are we teaching/assessing it?

Does it support learning?

How could we go about it differently?



To find out more:

Chahna.gonsalves@kcl.ac.uk

<https://www.kcl.ac.uk/people/chahna-gonsalves>

www.linkedin.com/in/dr-chahna-gonsalves

@journeytdoctor

