

IMPERIAL

Student Perspectives on the Use of Generative AI in Assessment

Presenters: Dr Richard Bale, Dr Tiffany Chiu, Dr Monika Pazio Rossiter

Centre for Higher Education Research and Scholarship (CHERS)

Imperial College London

15 January 2025

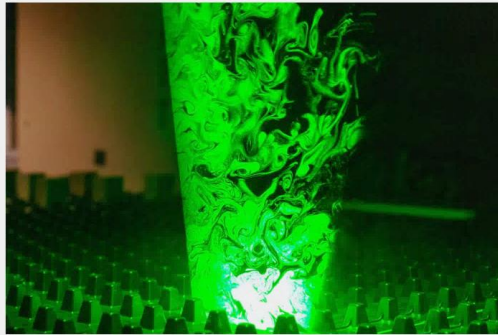
QAA-funded Project – Supporting the ethical and equitable use of GenAI for
formative and summative assessment: a transnational perspective

IMPERIAL

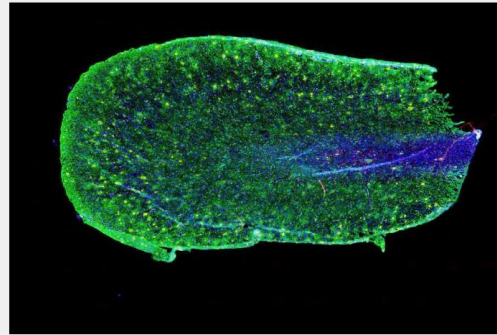
Institutional context

Imperial College London – *the institution*

- STEMMB focus
- Four faculties
- Russell Group; research-intensive



Faculty of Engineering



Faculty of Medicine



Faculty of Natural Sciences



Imperial College Business School

Imperial College London – *the students*

- ca. 23,000 students
- Accustomed to high achievement; assessment-focused
- Competitive
- High expectations; tendency to over-work

IMPERIAL

Participant and data overview

7

Focus
groups

27

Students

5

Departments

Physics, Computing, Life Sciences,
Medicine, Bioengineering

IMPERIAL

Themes

1 Influence of departmental and institutional beliefs of AI use on students

2 Students' wider societal perceptions of AI influence perceptions of AI in academia

3 AI literacy and criticality

4 Trust hierarchy - AI limitations affect trust of AI in academia

5

AI use has wider societal and ethical implications

6

AI use as a collaboration

7

Relationship between AI use and type of assessment

8

Perceptions of AI use are affected by beliefs of the role of staff, pedagogy and institution

9

AI as a beneficial study tool

IMPERIAL

Themes

1 Influence of departmental and institutional beliefs of AI use on students

2 Students' wider societal perceptions of AI influence perceptions of AI in academia

3 AI literacy and criticality

4 Trust hierarchy - AI limitations affect trust of AI in academia

5

AI use has wider societal and ethical implications

6

AI use as a collaboration

7

Relationship between AI use and type of assessment

8

Perceptions of AI use are affected by beliefs of the role of staff, pedagogy and institution

9

AI as a beneficial study tool

IMPERIAL

Theme: AI use as a collaboration - Student experiences with GenAI

Starting point for learning & Springboard for idea generation

- “I do use it for kind of **background reading** at the start of a project or kind of just to understand like the main points. But then I would kind of **weight it less heavy** as I continue **going into more detail.**” (Anika, FG5)
- “I use it as like to start my thinking process, especially when I'm quite like I don't know where to start. So for me it's a **good starting point in the springboard.**” (Leena, FG6)

Tool for simplifying complex concepts

- “I think there's something about a lot of the language models is that they're really **good at explaining things...** in cases where kind of students will ask, oh, can **you explain that in a different way?** And often lecturers just can't that maybe AI would be good at stepping in to help that process?” (Fred, FG2)
- “[ChatGPT is] very good at sort of **breaking down** each line of code and **explaining what it does.**” (Hana, FG7)


IMPERIAL

Theme: AI use as a collaboration - Student experiences with GenAI, cont'd

Supplementary, not primary, resource

- “I just started using it in my own work and it's just becoming like a little, I guess, **a support tool**. It's not the only thing I use, but it's part of my, my **toolkit of resources available**.” (Leena, FG6)
- “I definitely use it in part of my study. For example, like doing like a projects, but I **wouldn't use it as a central source of information** rather than just to play around with it and see what kind of suggestions they give. **Some kind of inspirations**.” (Theo, FG2)

FG question: Would you consider it ethical to copy in an assessment the output from an AI prompt? Please give examples relating to both summative and formative assessments.



- **The focus groups highlight diverse student perspectives on the ethical use of AI in assessments**
 - “I actually think it is...you can take credit for it because prompt engineering is such a big thing like you have”
 - “I'm not too sure if this is a good idea... I think you personally should also play the majority play a larger role in actually creating this assessment and the written work”
 - “AI has like an influence, but you are the majority of the work. It's like in your tone, it's in your voice. It's in like your perspective because someone says you have to argue your perspective so.” (FG2)

IMPERIAL

Implications for educators: Transparency in ethical AI use in education

- **“Human-centric:** viewing AI as a tool to augment, not replace, human insight, judgment and creativity.
- **Skill-centric:** focusing on the development of transferable skills (i.e., problem formulation, exploration, experimentation, critical thinking, and reflection) over mastering specific tools.
- **Responsibility-centric:** promoting effective and responsible use of generative AI tools.”
(Acar, 2023)

Acar, Oguz A., 2023. “Are Your Students Ready for AI? A 4-Step Framework to Prepare Learners for a ChatGPT World”, *Harvard Business Publishing Education*. Accessed at: <https://hbsp.harvard.edu/inspiring-minds/are-your-students-ready-for-ai?>

To support ethical AI use, educators should:

- Clearly define acceptable and unacceptable AI practices in their assessment contexts.
- Highlight the importance of critical engagement and originality in assessments.
- Foster discussions about AI as a tool to support learning rather than bypass it.

IMPERIAL

Integrating Generative AI (GenAI) in Higher Education: Supporting Concept Development and Technical Design in Engineering Group Projects

IMPERIAL

Integrating Generative AI (GenAI) in Higher Education: Supporting Concept Development and Technical Design in Engineering Group Projects



Introduction

This case study outlines how Generative AI tools were introduced to Bioengineering students to support their group assessment task. This practice helped students explore creative ideas, improve technical design processes, and reflect on the strengths and limitations of using GenAI. This case study also highlights key insights to help educators apply similar approaches in their teaching.



Assessment Overview



The first-year undergraduate module 'Design & Professional Practice' involved a 40% weighted group assignment, the 'Design Challenge', with 180 students working in teams of 5-7. Students were tasked with designing a desktop plant pot inspired by bioengineering, submitting a report detailing their design process, final prototype, and the bioengineering concepts behind their design.

Introducing and Using LLMs for the Assessment

Students were introduced to large language models like ChatGPT in a kick-off session that covered the assignment briefing, referencing guidelines, and the potential of GenAI to support design processes. The session included demonstrations of GenAI's capabilities through real-time queries and discussions on its applications and limitations in engineering design. Students were given the option to use ChatGPT or other LLMs, provided that proper referencing was applied.



Defining Boundaries: Ensuring Ethical and Transparent Use of GenAI



Students found GenAI tools are useful for design generation prompts for suitable materials, checking grammar and sentence structure. Some of the challenges students faced when using AI are listed below:

- Lack of citations from ChatGPT output
- Incorrect responses by LLMs
- Lack of Creativity in LLM outputs

Given the wide range of purposes and contexts in which LLMs can be used, staff and students co-developed clear guidelines regarding appropriate LLM usage and corresponding referencing requirements. Specifically, a rubric framework was created to clarify expectations for students and provide a standardised approach to acknowledging LLM contributions.

Template rubric →

ISSUES	Point to the Original Reference	Always Acknowledge the Contribution	Reference to Report Content/Process	Acknowledge Prompt/Process/Chat History	Use Published
Learning/Understanding Link				X	
Using AI to generate content		X			
Using LLMs to check grammar			X		
Using LLMs to generate content			X		
Using LLMs to generate content					X

Key Factors to consider for Effective Implementation



- Open discussion: Demonstrate GenAI tools like ChatGPT during class, running real-time queries. Facilitate conversations with students about the potential and limitations of AI to support their learning and engagement.
- Clear guidelines: Define how LLMs can be used in assignments, such as for brainstorming, technical support, or language assistance and promote AI as a tool to enhance, not replace, student creativity and critical thinking.
- Referencing: Provide a rubric to guide proper citation of AI use, ranging from basic references to detailed context required.

Challenges with referencing LLM outputs

- LLMs can be used for a wide range of purposes and contexts. How should students reference properly or is referencing needed? Would students be allowed to use LLM generated code for their assignments, and how are they expected to reference it?
- Communication and transparency are important.

<u>USE CASE</u>	Free to Use without Reference	Simple Reference to Acknowledge Use	Reference + Report Primary Prompt	Reference + Report Primary Prompt + Report Chat History	Use Prohibited
Generating Code using LLMs				X	
Using LLMs for linguistic assistance		X			
Using LLMs to aid in research			X		
Using LLMs to generate design concepts			X		
Using LLMs to generate large portions of text					X

Reference: From a case study by Bioengineering UROP students who evaluated the use of AI for this assignment.

IMPERIAL

Theme: Relationship between GenAI use and type of assessment

Mixed picture – use of GenAI to maximise summative grades vs. transparent, ethical use of GenAI

If it's summative, obviously I want good grades, everyone wants good grades, so I will **try my best to then maximise my grades** and that would mean obviously using ChatGPT. (Paul, FG 1).

I would say I would use it for both, but **for a graded task, obviously you don't want to be flagged for AI**, right? So for graded tasks instead of getting it to write me an essay for example, I'll get it to write me an essay plan and then I'll write my own essay, whereas **if it is non-graded, I would just submit the written piece by GenAI because there's no consequence to it.** (Malik, FG 1)

In terms of formative though, I think it sets the tone for summative, so it's important in that sense. But I guess **because formative doesn't actually count, it's like it is unethical to still use it, but it has less significance.** (Anika, FG 5)

IMPERIAL

Theme: Relationship between AI use and type of assessment

Formative assessment for learning – perception that GenAI does not facilitate learning

The funny thing I've noticed is whenever it's formative, I tend to not use it actually because **I care about my learning** [...] there's no pressure on me to get it right. (Paul, FG 1)

The whole point of them [formative assessments] is to kind of make you learn, right? And if you're just putting it into ChatGPT, it's probably not going to be the best for you as an individual [...] **the nice thing with formative assessments is you kind of get more room to be creative** sometimes [...] I think if people just ChatGPT'ed everything – I think you can be a bit less strict in the formative assessments in use of AI – but again, I would say it's not really worth just copy and pasting things. (Helena, FG 4)

I think it also- it's a little **disrespectful of the marker's time** if you're just- the marker is being paid to help you learn and you're effectively just giving them busy work that does nothing for your learning. (Amelia, FG 5)

IMPERIAL

Theme: Relationship between GenAI use and type of assessment

The nature of the task and authentic application is more important than the type of assessment

It doesn't really depend on the formative or summative assessment, but **it depends on the the nature of the tasks.** (Moon, FG 1)

I think if you are stupid enough to set an assignment where it doesn't test anything and it just tests a bunch of random facts, then the uni deserves that people use AI against it, to be honest. But if you're gonna actually **set an assignment where it's an actual application and synthesis of knowledge, then obviously it's wrong to use that** [...] I don't believe in pointless assessments, testing random facts that we don't need to know that in the real life actually have no applications for. (Kabir, FG 6)

IMPERIAL

Using GenAI to prepare for an oral presentation assessment

Student Perspectives on Utilising Generative AI for Oral Presentation Preparation: Idea Generation & Question Prompts



IMPERIAL

Integrating Generative AI (GenAI) in Higher Education: Embedding AI into coding assessments

IMPERIAL Integrating Generative AI (GenAI) in Higher Education: Embedding AI into Coding Assessments



Introduction

This case study highlights the integration of AI, notably ChatGPT, into coding assessments to enhance students' critical thinking and problem-solving skills. By generating, analysing, and refining AI-produced code, students gain practical experience with AI tools while addressing their limitations. This approach prepares them for AI's role in modern workflows, emphasising responsible and effective usage.



Why is this exercise useful?



- Builds critical thinking and problem-solving skills.
- Teaches ethical and effective AI use.
- Prepares students for AI-driven workflows.
- Enhances coding and prompt-crafting abilities.
- Encourages recognising and improving AI limitations.

Method

- **Goal:** to help students develop skills in Python programming, code testing, profiling and optimisation while fostering algorithmic thinking.
- Students are provided with a framework containing a skeleton Python model and tasked with completing the following: Code Development, Code Profiling, Code Optimisation, and Reflection and Discussion. AI is integral part of the task as they are asked to use AI to refine the code and reflect on the output.
- ChatGPT 3.5 was chosen for its widespread familiarity, accessibility and consistency.
- Students are assessed on Code Functionality (12 Marks) and Critical Discussion and Analysis (18 Marks) - for critical approach to analysis, refining and optimising the code.
- Additional marks are provided for identifying and correcting errors in ChatGPT's output, incentivising students to actively engage with AI.

Key Factors for an Effective Implementation



- Ensure a Clear Method for Submitting AI Interactions
- Define Consequences for Non-Compliance with Submission Guidelines
- Ensure Equal Access to AI Tools for All Students
- Test AI Tools and Assessments Before Launch
- Monitor and Limit the Impact of Your AI Interactions
- Encourage Critical Engagement with AI Output
- Provide Guidelines for Ethical AI Use
- Clarify the Role of AI in the Assessment Process

Quotes from Students

An example of how a student used AI for debugging:

- 'I was trying to debug codes and found that Chat GPT has become good at being able to do that, especially when it's just small things like human error tends to miss like oh, missing a comma, or missing hyphen' (Hana, FG 7)

An example of how clear rules are required when AI is to be used in assessment, to avoid any confusion:

- 'Just at least from a college perspective, when it's your degree, then we need to have more clear rules. I understand that it keeps changing, but I just feel that even when it does change, they just can't answer half our questions and that's really tricky from the perspective, especially [of] summative assessments' (Anika, FG 5)

An example of how educators can positively incorporate Gen AI in teaching:

- 'If an educator was to use GenAI in a positive way, I do think it's something that would be positive because I think we're taught especially in a kind of old fashioned [or] traditional lectures, it's very technology averse. It's nice to have someone who's kind of embracing that side of things as well' (Kamal, FG 7)

IMPERIAL

Theme: Trust hierarchy – AI limitations affect trust in AI in academia

Awareness of issues related to AI outputs and the need to approach its use critically

"If it's a little bit **nuanced** in what the answer should be, **it gives it something slightly different, which is concerning** sometimes."

(Mei, FG 1)

"If I'm getting it to kind of do anything that involves something which I think is **more niche** in terms of the content information, I'd probably instinctively **wouldn't trust it.**" (Kamal, FG 7)

"I still don't find it entirely trustworthy when you're into the detail, I think it is **fine in terms of summaries** and things like that." (Anika, FG 5)

'I find that **the more I use [AI], the less I trust it.**' (Fred, FG 2)

IMPERIAL

Theme: Trust hierarchy – AI limitations affect trust in AI in academia

Sources of knowledge - Who do you trust? AI vs own knowledge vs peers

"But sometimes I can't tell it's clearly wrong 'cause I **keep thinking, oh, it can't be wrong**. It's AI. And sometimes I **question my own knowledge**, and so I'd spend like an hour trying to understand chat, GPT's answers and then realise at the end that actually it's wrong." (Ling, FG 2)

"I got I started using GTP lot more than all my friends around me. So whenever like I sometimes get an answer but then like **my friends ask like is it like your answer or did you just get this from chat GPT** and it's starting to get into more and more like a trend?" (Leena, FG 6)

"I wouldn't say I would trust it more than my friends. And it might be because **I'm surrounded by people who are smarter than me**. But like majority of the people that I like, all my friends, they're quite good. So if I'm stuck on a question, **I would tend to actually usually ask them first**, as opposed to, then go on to ChatGPT." (Alaia, FG 4)

"So I do trust it. I wouldn't say I would trust it more than people, but **I do trust my more than myself**." (Theo, FG 2)

IMPERIAL

Theme: Perceptions of AI use are affected by beliefs of the role of staff, pedagogy and institution

AI Undermining Lecturer's Expertise and Competence

"The lecturers probably know way more than what ChatGPT or any AI tool will ever know."
(Dylan, FG 3)

"If lecturers use ChatGPT to plan lectures, **it feels like they're a bit incompetent.**" (Aarush, FG 4)

'I would think that as lectures like I'd expect them [to not use AI], **they're smart people, they wouldn't need AI.**' (Mei, FG 1)

IMPERIAL

Theme: Perceptions of AI use are affected by beliefs of the role of staff, pedagogy and institution

Perceived Devaluation of a University Degree

"It feels like **AI usage in teaching devalues the purpose of coming to college** when we're paying so much." (Hana, FG 7)

"**We're paying** the university, we're our students. **We're here for the lecturers' knowledge.**" (Aisha, FG 3)

"The reason why professors are in their job is because of their experience and their knowledge. Using an AI, which doesn't have that experience, feels like, **what's the point of having my professor?**" (Jia, FG 2)

A more balanced view...

"I think it would be **hypocritical to criticize teachers for using AI when most students use it** regularly. **It's just about how it's being used.**" (Moon, FG 1)

IMPERIAL

Takeaway
messages...

