

How to Assess the Benefits of GenAI in Higher Education

Author: Adam El Haj

University College London (UCL), UCL School of Management



Introduction

The Impact of Generative AI

- Global AI market valued at **USD 279.22 billion in 2024**, projected to reach **USD 1,811.75 billion by 2030** (CAGR 35.9%).
- Three in ten UK adults** perceive AI as a negative influence (GOV.UK, 2024).
- 86% of university students** actively use AI tools, viewing them as indispensable.

Current Issues in Assessing AI

- Assessing benefits is **not standardised**, limiting universities’ ability to adapt curricula.
- Institutions must balance **value creation** with risks such as over-reliance, bias, and misconduct

(Digital Education Council, 2024)

This report explores how institutions should assess the benefits of generative AI in higher education.

Literature Review

Benefits for Students

- Academic Writing Support: Improvements in grammar, structure, and fluency; supports brainstorming and formatting.
- Personalisation: Customised, immediate feedback; boosts confidence, reduces anxiety.
- Productivity & Time Efficiency: Enhances learning speed, simplifies complex concepts, fosters independent study.
- GenAI seen as a “cognitive partner”, facilitating deeper learning and creativity.

Benefits for Educators

- Productivity: Saves time on grading, data analysis, admin tasks.
- Assessment & Feedback: Generates prompts, automates grading, delivers adaptive feedback.
- Curriculum Reform: Supports inclusive learning and AI-informed pedagogy; raises importance of AI literacy for educators.

Research Question

How can we assess the benefits of generative AI in higher education using both qualitative and quantitative methods?

The aim of this research question is to evaluate how the benefits of GenAI can be assessed in a holistic and evidence-based manner, through evaluating both qualitative and quantitative approaches in published literature

Proposal

- Introduces a hybrid rubric to address gaps in evaluating AI’s benefits.
- 70% universal: assesses knowledge, scholarly skills, industry application, and professionalism.
- 30% discipline-specific: designed by staff to capture field-level benefits.
- Provides a structured, transparent, and transferable framework for higher education.
- Pilot study at UCL (Management Education) proposed to test rubric on written assessments, decision-making, and career readiness.
- Aim: empirically validate rubric and guide AI integration across disciplines.

UCL Management Course Assessment Rubric (70-30 Split)

HEQF Level 6 – Final Undergraduate Level

Criteria	0–19 (Fail F)	20–39 (Fail E)	40–49 (Third D)	50–59 (Lower Second C)	60–69 (Upper Second B)	70–79 (First A)	80–100 (High First A)
Knowledge & Understanding (20%)							
Intellectual Skills (20%)							
Research & Enquiry (15%)							
Scholarly Practices (10%)							
Professional & Life Skills (5%)							
Application to Practice & Industry Relevance (10%)							
Strategic & Ethical Awareness (10%)							
Problem-Solving & Innovation (10%)							

Key: grading criteria in dotted line box are course specific (management course)

Findings

- Mixed-methods approaches dominate (surveys + scenario/rubric/thematic coding).
- Surveys & interviews capture student and teacher attitudes.
- Rubric-based evaluations allow objective comparisons of GenAI vs student work.
- GenAI strong in grammar/structure, weaker in conceptual reasoning.
- Effectiveness varies across disciplines: better in analytical fields, weaker in project-based.
- Domain-specific rubrics highlight nuanced benefits (e.g. dentistry).
- Trend: push towards systematic and transparent evaluation frameworks.

Discussion

- No standardisation across studies; comparisons difficult.
- Heavy reliance on student perceptions, which may skew results.
- Lack of attention to long-term outcomes (e.g. career readiness).
- Rubrics useful but often field-specific, limiting broad applicability.
- Some studies lack depth or transparency in scoring methods.

Overall, the discussion highlights that while recent studies provide valuable insights into the perceived and demonstrated benefits of generative AI in higher education, the lack of standardisation across approaches makes it difficult to assess benefits effectively and holistically. The reliance on student perceptions, limited attention to long-term outcomes, and inconsistency in rubric design restrict the applicability of findings for institutions. Taken together, these issues underscore the need for a structured and transparent framework that can bridge short-term perceptions with objective, discipline-sensitive measures, enabling universities to evaluate and incorporate AI tools in a meaningful way.

