Reconstructing Formative Assessment in the Era of Generative Artificial Intelligence and the Upstream Governance of Academic Integrity: A Case Study of Student Voices in an Introduction to New Media Course

Hefei Li

Guangdong Baiyun University, Guangzhou, China
lihefei@gdbyxy.wecom.work

Abstract. As generative artificial intelligence enters university courses, how assessment can identify genuine learning has become a critical question. Taking the course Introduction to New Media as a case, this paper analyzes 210 questionnaire responses and 259 open-ended statements, and employs thematic analysis to examine students' use of AI and their expectations of assessment. The findings show that students mostly used AI to generate ideas, build structure, and refine expression; tasks marked by high difficulty, broad scope, and strong template conformity, together with the absence of process checks, were perceived as conditions that invite direct AI use. Students expected rules to be established in advance, processes to be visible, contributions to be attributable, and tasks to be authentic. On this basis, the paper proposes four mechanisms—personalized tasks, process visibility, original thinking assessment, and AI-use transparency—which move the governance of academic integrity upstream into course design.

Keywords: Firgenerative artificial intelligence, formative assessment, academic integrity, journalism and communication education.

1 Introduction

1.1 A Subsection Sample

Generative artificial intelligence has entered the stages of information retrieval, ideation, and expression in university coursework. An outright ban proves difficult to enforce, while unregulated use blurs both the learning process and the attribution of competence; judged on the final product alone, it is also hard to tell what thinking a student has actually performed. The problem is particularly acute in journalism and communication programs, where content production lies at the core of the discipline and the intervention of AI touches the very foundations of professional training. The three formative assignments in the author's Introduction to New Media course face the same difficulty. In the end-of-semester questionnaire, 90% of students considered that the course had improved their abilities, whereas only 4%–5% self-reported that their

© The Author(s) 2026
M. B. Badzis et al. (eds.), *Proceedings of the 2026 2nd International Conference on Education Research and Training Technologies (ERTT 2026)*, Advances in Social Science, Education and Humanities Research 1042,
https://doi.org/10.2991/978-2-38476-628-4_24

work was “entirely or largely AI-generated”. Against this background, this paper examines how students use AI, which task conditions are perceived as risky, and how assessment can retain evidence of genuine learning.

2 Literature Review and Theoretical Foundations

2.1 Generative AI and the Reconstruction of Assessment in Higher Education

Recent research has discussed the impact of generative AI on higher education along two main lines: one concerns learners’ patterns of use and their ethical awareness [1]; the other examines the knock-on effects on faculty capability, curriculum provision, usage policies, and ethics training in journalism and communication education [2]. In broader higher education settings, cross-national surveys of teachers and students indicate that generative AI is prompting instructors to reconsider assessment design and to place the learning process and higher-order abilities in a more prominent position [3]. Research on students’ use and perceptions of chatbot assistance in completing assessments likewise reflects a continuing need for courses to specify the boundaries of generative AI use and to adjust assessment arrangements accordingly [4]. Analyses of institutional generative AI assessment policies further argue that the definition of “originality” in student work needs to be re-examined under conditions of digital knowledge production [5]. Course assessment therefore needs to bring learning evidence, process-based performance, and usage rules into a single design frame.

2.2 Formative Assessment and Learner Agency

Formative assessment provides feedback during the course of instruction and uses that feedback to adjust teaching and learning; its central purpose is to foster students’ self-regulated learning, that is, the active monitoring and adjustment of goals, processes, and strategies [6]. In situations where AI can participate in producing the output, the focus of assessment should shift from “whether the product was written by the student” to whether the student took part in judging, revising, and reflecting; related experiments also suggest that the learning effects of AI depend on whether the student’s cognitive engagement is preserved [7].**Sample Heading (Third Level).** Only two levels of headings should be numbered. Lower level headings remain unnumbered; they are formatted as run-in headings.

2.3 Cognitive Scaffolding: A Working Definition

This paper adopts “cognitive scaffolding” as an analytic concept: AI may offer temporary support for initiation, organization, and expression, but the selection of sources, the exercise of judgement, and the work of revision should remain with the student. The definition draws on the notion of “support without replacement” in scaffolding theory

[8]. It is used to interpret students' self-reported experiences and the risks embedded in course design, rather than to establish what students actually did.

2.4 Research Gap and Analytical Framework

Existing research seldom presents student experience within concrete course tasks, and course-level material connecting "patterns of use – risk conditions – assessment expectations" remains scarce. Using course questionnaires and open-ended items as data, this paper brings student voice into the discussion of assessment in the AI era. The analysis does not infer actual AI use directly from student opinion; instead, it treats such opinion as a clue for understanding task experience and assessment expectations. It first describes students' self-reported patterns of AI use, then identifies the task and assessment features that raise the likelihood of direct AI use, and finally distils students' demands concerning rules, process, and authenticity, from which adjustments to formative assessment are derived.

3 Research Design

3.1 Research Questions and Data Sources

This paper adopts a single-case research design, taking the Introduction to New Media course taught by the author as the object of study. The course is offered to first-year students majoring in Network and New Media and sets three formative assignments: a personal digital profile report, a brand communication planning proposal, and a new media platform advertising analysis report. All three assignments involve information retrieval, proposal design, written expression, and case analysis, and thus provide a concrete context for examining how students draw on AI across different tasks and how they understand assessment rules.

Centred on students' experience of using AI and their assessment expectations, the paper poses three research questions. RQ1: How do students self-report their use of generative AI in course assignments, and what characterizes their patterns of use? RQ2: Which task and assessment features do students believe increase the likelihood of direct AI use? RQ3: What demands do students make of course assessment, and what problems in assessment mechanisms do these demands expose?

The data come from an anonymous end-of-semester questionnaire administered on Wenjuanxing, a Chinese online survey platform. The valid sample comprises 210 responses from students of the 2025 cohort in the Network and New Media program at Guangdong Baiyun University. The questionnaire contained closed-ended items and three open-ended questions asking for suggestions on improving the formative assignments, the final examination, and course instruction respectively. After removing answers with no substantive content, such as "none" or "nothing", the three open-ended questions retained 86, 80, and 93 statements, giving a total of 259 statements, which constitute the primary material for thematic analysis. The closed-ended items are used only to describe patterns of use, risk perception, and assessment preferences within the sample; no inferential statistics are reported. It should be noted that the open-ended

questions were not framed around AI, and they are therefore used here to analyze course assessment experience and task expectations rather than read directly as attitudes toward AI. The closed-ended items in Section 4 concerning task risk, assessment preferences, and task authenticity permitted multiple responses; percentages for each option were accordingly calculated against the number of valid responses and do not sum to 100%.

3.2 Thematic Analysis Procedure

The 259 open-ended statements serve as the primary analytic material, with the closed-ended results used as a descriptive backdrop for understanding students' patterns of use, risk perception, and assessment preferences. Following the thematic analysis procedure of Braun and Clarke [9], the texts were first subjected to initial coding, then consolidated into candidate themes, which were checked by repeated comparison between coded extracts and the full dataset. After definition and naming, four themes were finalized. To reduce the risk of selective presentation, coding memos were maintained and negative cases were examined; nonetheless, single-researcher coding and the dual teacher–researcher role may still have shaped the interpretation.

3.3 Research Ethics and Limitations

The questionnaire followed the principles of anonymity and voluntary participation, and no personally identifiable information was collected. The sample comes from a single course in a single semester (N = 210), the data rely on self-report, and the open-ended items were not designed specifically around AI. The results are therefore positioned as interpretive clues for course design, and are not intended to support causal inference or broad generalization.

4 Findings: AI Use, Task Perception, and Assessment Expectations

4.1 From "Not Knowing Where to Start" to "AI as Scaffolding": Self-Reported Patterns of Use

Descriptive statistics show that the proportions of students self-reporting that their work was "entirely or largely AI-generated" were 4.2%, 4.2%, and 5.2% for the three assignments. Correspondingly, most students reported that AI was used mainly to produce a framework or parts of the content followed by substantial additions and revisions, or else for retrieval, organization, and language polishing. Among the main reasons for using AI, "having no ideas and not knowing how to start" (165 students, 78.5%) and "inadequate writing ability and needing to improve the text" (144 students, 68.5%) were the most common, whereas "saving time and not wanting to invest much effort" was cited by only 9 students (4.2%). These distributions point to a tendency to use AI as a way of lowering initiation costs and supporting expression.

The open-ended responses reflect the same logic. One student wrote:

"Could you teach us how to analyze data without using AI, and help us strengthen our own ability to analyze data." — Q32-04

This suggests that students sought help from AI while at the same time worrying that their analytical ability would be displaced. Judged by the standard set out above, most self-reported experiences approximate a scaffolding pattern of "task framing – inspiration – expression support"; whether key judgements and the evaluation of results remained with the student would need to be verified through process materials or interviews. A low self-reported rate of substitutive use therefore does not mean that risk is absent; it is better read as a starting point that course design can work with.

4.2 From "Task Difficulty" to "Perceived Risk of Direct AI Use"

Although self-reported substitutive use was uncommon, students were remarkably consistent in judging "which assignment features would make people inclined to generate work directly with AI": 138 students (65.7%) selected "the assignment is too difficult", 105 (50.0%) selected "the prompt is too broad and generic", 54 (25.7%) selected "the assignment is heavily templated", and 39 (18.5%) selected "only the final product is assessed, with no process checks". These figures reflect students' subjective judgements about risk conditions, not a test of causal relations among variables.

Several responses concerning group work illustrate the point concretely:

"In a group, everyone has different goals; some members are keen to participate while others are not. Those with little willingness simply hand the original prompt straight to AI, so that later analysis gets produced before the earlier data have been integrated — putting the cart before the horse." — Q32-36

"I hope the part each person is responsible for could be explained in more detail before the assignment is set. Sometimes with collaborative work, or when you draw a section you are not good at or unfamiliar with, there is simply no way to start if you do not use tools." — Q34-129

These data reflect students' subjective judgements about risk conditions and cannot support causal inference. Stated more cautiously, students regard task difficulty, overly broad prompts, templating, and insufficient process constraints as conditions likely to increase direct AI use. This suggests that the issue cannot be understood solely in terms of individual integrity; task and assessment design must also be brought into view. Course design influences both how students seek support and the extent to which they take responsibility for their own learning.

4.3 From "Final Product" to "Process Assessment": Students' Demand for Rules

When asked "which adjustments would increase independent thinking and reduce direct AI use", students most often selected "adding personalized, task-specific requirements" (132 students, 62.8%). "Refining the grading criteria and raising the weight of original thinking and personal reflection" and "providing more detailed

assignment guidance and reference examples before class" were each selected by 60 students (28.5%), while a further 54 (25.7%) selected staged process-based assessment.

In the open-ended responses, students likewise called actively for rules to be set in advance and made transparent:

"I suggest releasing detailed grading rubrics for each assignment in advance, along with an analysis of the key and difficult points, so that the scoring criteria are clear and students do not have to revise again and again because the criteria are vague." — Q32-202

In the assessment of group work, students were especially concerned with whether individual contribution could be seen. Some 56% of students preferred forming their own teams, 25% reported "uneven division of labour and low participation", and 18% explicitly proposed "grading by individual contribution". Their statements were direct:

"Grade by individual contribution." — Q32-01

"Set individually assessed assignments; the friction of communicating in group work uses up all your energy." — Q32-08

Taken together, students are not simply asking for AI to be "banned"; they want rules and support that are clear, set in advance, and predictable. Staged feedback, criteria released in advance, and identifiable individual contribution correspond respectively to criteria clarification, process monitoring, and evidence attribution in formative assessment; they also indicate that current course assessment still leaves room for improvement in the supply of process information.

4.4 From ""Generic Prompts" to "Authentic Practice": Students' Expectation of Task Authenticity

Students' expectations of course content also point toward task authenticity. Among the "hands-on content they most wanted added", account operation (136 students, 64.7%), employment and internship guidance (102 students, 48.5%), commercial monetization methods (90 students, 42.8%), practical exercises in public opinion management (88 students, 41.9%), and short video production (86 students, 40.9%) ranked highest. In evaluating what they gained from different teaching methods, real industry case analysis (84 students, 40.0%) ranked above hands-on classroom exercises (75 students, 35.7%) and lectures (53 students, 25.2%).

"Some assignments need industry data and research reports to support them; I hope the assignments come with guidance on data retrieval, or that the class tells us which free, authoritative, and efficient methods and platforms there are for finding data." — Q32-20

"I hope the questions could draw more on current popular cases of new media commercial monetization and AI-driven operation." — Q33-87

Such expectations carry two implications for assessment design. Students want support for the compliant use of AI, for data retrieval, and for industry judgement within authentic tasks; at the same time, tasks anchored in specific contexts, personal experience, and process materials are harder to complete by applying generic AI answers directly. Authenticity is thus both a learning need and a means of reducing substitution risk.

In sum, the thematic analysis condensed the 259 open-ended statements into four themes, read against the descriptive distributions of the closed-ended items:

T1 Orientation of use: scaffolding rather than full substitution — students mostly used AI for task framing, inspiration, and expression support; this judgement rests on self-reported material and needs further verification with process data.

T2 Risk perception: task and assessment structures — students considered high difficulty, broad prompts, templating, and the absence of process checks to increase the likelihood of direct AI use.

T3 Demand for rules: visible process and clear criteria — students expected grading criteria to be released in advance, staged feedback, and identifiable individual contribution.

T4 Expectation of authenticity: connecting with industry practice — students expected real cases, data retrieval skills, and platform operation to be built into tasks and assessment.

The four themes form a coherent logical sequence: T2 and T4 highlight the risks and authenticity requirements in task design; T1 reflects students' tendency to employ AI for problem-solving, inspiration, and support in expression; T3 underscores the necessity of reconstructing assessment evidence through a transparent process, clearly defined standards, and identifiable contributions.

5 Discussion and Mechanism Construction

Taken together, the four themes indicate that task structure determines how AI enters learning, while assessment structure determines whether such use can become credible evidence of learning. When task goals, pathways, and support are unclear, AI more readily becomes the first option for lowering initiation costs; when assessment attends only to the final product, the risk of substitutive use rises accordingly. The problem should therefore not be reduced to individual integrity; it must also be traced back to task and assessment design.

The key to formative assessment is to make students' judgement, revision, and collaboration visible again. The staged feedback, criteria released in advance, and contribution records raised under T3 provide the evidence base for this, while T4 requires tasks to constrain the direct application of generic answers through authentic contexts and specific materials.

On this basis, the mechanisms can be linked along the chain of"task–process–evaluation–norm"(Fig.1): personalized tasks pose questions that students themselves must judge; drafts, search trails, and version records render the process visible; problem definition, evidence selection, and revision notes become the focus of grading; and a brief statement of AI use records where the tool intervened and where responsibility lies. In this way, AI-use transparency ceases to be a matter of retrospective accountability alone and becomes part of front-end assessment design.

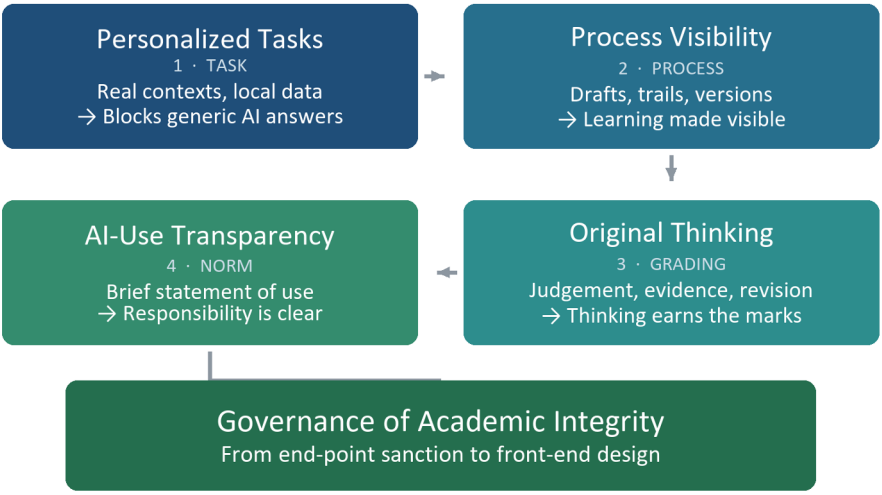


Fig. 1. A four-mechanism model of human–AI collaborative formative assessment.

6 Conclusion and Outlook

The paper answers the three research questions as follows: students used AI mainly for task framing, structuring, and expression support (RQ1); tasks characterized by high difficulty, broad scope, templating, and the absence of process checks were perceived as risk conditions for direct AI use (RQ2); and students expected rules set in advance, visible processes, attributable contribution, and task authenticity (RQ3). On this basis, the paper proposes four mechanisms: personalized tasks, process visibility, original thinking assessment, and AI-use transparency.

Limited by a single course, a single semester, a small sample, self-reported data, and single-researcher coding, the conclusions should not be generalized broadly. Once genuinely new samples become available, future work could undertake cross-course and cross-semester tracking, and combine process materials, interviews, and multiple-coder analysis to test the applicability of these mechanisms in different contexts.

References

1. Kankevičienė, L., Gedvilienė, G.: The interplay between generative artificial intelligence use and ethical awareness: Learners' perspectives. Pedagogika 161(1), 149–169 (2026)Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999).
2. Wenger, D., Hossain, M. S., Senseman, J. R.: AI and the impact on journalism education. Journalism & Mass Communication Educator 80(1), 97–114 (2025)LNCS Homepage, http://www.springer.com/lncs, last accessed 2023/10/25.

3. Kizilcec, R. F., Huber, E., Papanastasiou, E. C., et al.: Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. Computers and Education: Artificial Intelligence 7, 100269 (2024).
4. Gruenhagen, J. H., Sinclair, P. M., Carroll, J. A., et al.: The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments. Computers and Education: Artificial Intelligence 7, 100273 (2024).
5. Luo, J.: A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. Assessment & Evaluation in Higher Education 49(5), 651–664 (2024).
6. Nicol, D. J., Macfarlane-Dick, D.: Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education 31(2), 199–218 (2006).
7. Bastani, H., Bastani, O., Sungu, A., et al.: Generative AI without guardrails can harm learning: Evidence from high school mathematics. Proceedings of the National Academy of Sciences 122(26), e2422633122 (2025).
8. Wood, D., Bruner, J. S., Ross, G.: The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry 17(2), 89–100 (1976).
9. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative Research in Psychology 3(2), 77–101 (2006).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

