Prompt-as-Lesson: Authoring AI-Delivered Lessons for Fidelity and Observability

David Lefevre
Imperial College Business School, Imperial College London, UK

david.lefevre@imperial.ac.uk

Abstract: This paper reports the development of the prompt-as-lesson: a lesson encoded as a single prompt and run in the student's own LLM tool, with the model cast throughout as a tutor delivering the teacher's lesson, and the session closing with a report of the dialogue returned to the teacher. The purpose of the prompt-as-lesson is to enable students to pursue teacher-specified learning outcomes through tutorial dialogue with an LLM. The interactive, conversational nature of an LLM makes Laurillard's Conversational Framework the natural interpretive lens. Under this lens, the paper argues, the LLM is the first single medium capable of enacting the iterative teacher–learner dialogue the framework holds essential to formal teaching, yet this dialogue is frequently lacking in higher education. Introducing LLM activities into the formal learning environment risks inconsistency as, left to its defaults, the model teaches from its own conception and constructed learning environment rather than the teacher's. The prompt-as-lesson is a response. Utilising the prompt-as-lesson within the teacher's formal environment makes three demands: consistency, so that what is taught is drawn from the teacher's conception and pedagogy; adherence, so that the enacted lesson remains the authored one; and reporting, so that the dialogue is not opaque to the teacher. Design-based iteration on two contrasting modules to address these demands yielded two primary findings: contemporary LLMs struggled to adopt the learner's perspective and would not reliably hold to an unstructured lesson. The format that emerged seeks to address both issues and comprises three artefacts: a control surface of fourteen dimensions; an eight-section structure; and a four-phase authoring method. The paper claims no effect on learning; its contribution is a control surface, a structure and an authoring method. A formal evaluation of the format's fidelity of implementation begins in September 2026.

Keywords: Prompt-as-Lesson; Generative AI in Education; Conversational Framework; Fidelity of Implementation; Learning Design; Design-Based Research

1. Introduction

1.1 The Second Learning Environment

Students increasingly use generative AI in their studies. In a recent survey, 95% of UK undergraduates reported using AI and 94% reported using generative AI to help with assessed work (Stephenson and Armstrong, 2026). The AI investigated here is the large language model (LLM). It is an interactive medium, accessed via conversation, and, this paper argues, capable of enacting a formal teaching environment.

This interactive nature makes Laurillard's (2002) Conversational Framework an appropriate interpretive lens. This framework models formal education not as linear transfer but as an iterative dialogue across four components: the teacher's conception, the teacher's constructed learning environment, the learner's conception and the learner's actions. Here, the teacher's conception refers to the expert knowledge, theories and mental models that the teacher aims to communicate. The teacher's constructed learning environment refers to the pedagogical tasks, tools and settings designed for students to interact with and test those concepts.

Students have always utilised learning media outside the formal environment convened by the teacher, for example websites and MOOCs. However, Laurillard's pre-generative-AI analysis of learning media reached a precise verdict: no single extant medium enabled the iterative model 'in the way that a teacher in a practical session could'. The extant media covered the essential activities only in combination (Laurillard, 2002: 174). The combining required the work of a teacher.

The LLM is the first single medium capable of enacting the entirety of the formal model. It holds the four properties the framework demands, being discursive, adaptive, interactive and reflective. It converses at the level of descriptions, sets and adapts tasks, gives feedback on actions and prompts reflection. It requires no assembly, and therefore no assembler.

The LLM may be a means by which iterative conversation is expanded within the teacher's formal learning environment. The Conversational Framework holds iterative conversation to be essential to learning, yet iteration is frequently lacking in formal higher education. As an example, coursework typically comprises a single feedback cycle rather than a dialogue.

This challenge has a long lineage in tutoring research, from Bloom's (1984) tutorial advantage to intelligent tutoring systems (VanLehn, 2011), and more recently educators have begun to exploit the capability of generative AI. Mollick and Mollick (2023; 2024) deploy single-activity prompts for use as class activities and persona-based assistants. Prompts embedded in custom agents have been shown to be valuable in the field of physics education (Kestin et al., 2025). Recent systematic reviews of prompt engineering in education (Lee and Palmer, 2025; Qian, 2025) document a literature in which the model is cast as the teacher's assistant or as a tutoring persona, but not as the enactor of a teacher-authored lesson.

A student using an LLM now has a second environment fulfilling the Conversational Framework's description of teaching. However, this raises the issue of inconsistency for students: left to its defaults, the model teaches from its own conception and constructed learning environment rather than the teacher's.

1.2 The Prompt-as-Lesson Format

This paper investigates a response to this potential inconsistency, the prompt-as-lesson. This is a lesson plan authored by the teacher to be consistent with the teacher's conception and constructed learning environment (Laurillard, 2002). It is delivered via the student's own LLM tool with the model cast throughout as a tutor with a brief to deliver the lesson as a one-to-one tutorial. A report of the resulting dialogue is returned to the teacher, maintaining the iterative dialogue the Conversational Framework holds essential to formal teaching.

To this end, the lesson is bounded by a beginning and an end and proceeds as a sequence of timed teaching phases, described in Section 4 as 'acts'. Rather than being left to the model's defaults, it is specified against an explicit control surface, the dimensions of which are defined in Section 5.

The purpose of the prompt-as-lesson is to enable students to pursue teacher-specified learning outcomes through tutorial dialogue with an LLM, conducted within the teacher's formal learning environment rather than outside it. The following sections describe how the approach was chosen and how the subsequent format developed. They then present the resulting structure, a generalisable authoring method, and a planned evaluation.

2. Requirements and the Challenge of Fidelity

Bringing the LLM into the teacher's formal learning environment makes two demands. The first is consistency. The knowledge base, pedagogy and tone need to be aligned with the teacher's conception and pedagogical approach. Left to its defaults the model will draw upon its own knowledge base, which may differ from the teacher's conception and the curriculum curated for a class. Pedagogy may also differ. For example, LLMs are currently optimised for helpfulness, whereas desirable difficulties (Bjork, 1994) and productive failure (Kapur, 2008) hold that instruction frequently requires assistance to be withheld. The second demand is reporting. When a student uses an LLM the conversation and its outcomes are typically opaque to the teacher; maintaining the teacher–student dialogue requires some report of them.

The prompt-as-lesson format introduced in Section 1.2 adds a third demand: adherence. The lesson must be enacted as specified, adapt to the individual learner only within bounds the teacher sets, and bring the session to its planned completion. This is demanding because the model performs two roles at once: within an activity it is the instrument the student works with; across the lesson it is the tutor directing that work and holding the lesson to its course. Whether a programme is delivered as designed is the subject of an established literature termed fidelity of implementation (Stains and Vickrey, 2017).

These demands must be met under technical, resource and user-experience constraints. Technically, the platform must be capable of ingesting a prompt, delivering the lesson as intended and providing a means for activity to be reported to the teacher, and it must not conflict with institutional IT policy. On resource constraints, the format must be affordable to every party and represent value for money. Finally, the user experience must be sufficiently intuitive and easy to use such that it does not derail the lesson.

3. Arriving at the Technical Format

Initial investigations addressed the technical question of how students should access an LLM. Three options merited serious inquiry.

The first was to develop a bespoke platform sitting between the learner and the LLM. This method offers the greatest control as the interface can be specific to this context; transcripts resulting from the lesson would be

easily accessible; a student's progress through the lesson could be continually monitored and interventions could be made to keep a lesson on track. It is, however, the most complex and expensive option.

The second option considered was to embed a lesson configuration within the LLM platform provided by the institution. This would have the advantage of sitting within existing IT infrastructure and, as per option one, offer significant control over the lesson delivery. However, given the sensitivities surrounding so new a technology, engagement with institutional ICT would introduce significant organisational complexity.

The third, the prompt-as-lesson, sends the lesson to students as a prompt they run in their own LLM. This approach is relatively inexpensive and simple to implement. In addition, students use tools they already hold and do not need to learn a new interface. Reporting, however, is manual (the student must submit a report of the conversation) and control ends at the specification: unlike the first two options, the lesson cannot be monitored at run time. Provider-hosted configurations such as custom GPTs are considered here to be a delivery variant of the prompt-as-lesson.

The prompt-as-lesson was selected. It offers the least run-time control, but simplicity of implementation and reach into students' existing tools were, on balance, considered decisive.

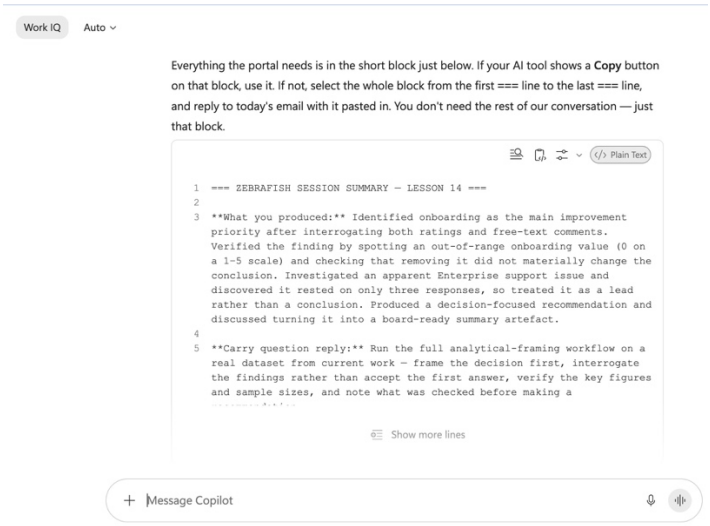


Figure 1: The close of an author test session in Microsoft Copilot: the summary block returned for submission

4. Formative Development

Once the technical approach was determined, the focus moved to the form the prompt-as-lesson should take. A format emerged from the development of prompt-as-lessons on two modules differing in audience, content and setting. The first, 'Sparks', is a set of self-contained micro-lessons on the topic of AI fluency for use in executive education. The second, the 'AI Companion', was developed for a ten-week strand of a postgraduate entrepreneurship module at a UK business school, in which nine weekly prompt-as-lessons each guide the student through a 30–45-minute tutorial applying that week's teaching to the student's venture.

The resulting format resembles a lesson plan stated in natural language but written knowing the teacher executing the plan will be an LLM. The specification was developed by iteration: scripts were drafted, run in repeated informal tests across the LLMs students typically use (ChatGPT, Gemini, Claude and Copilot), and revised wherever the model's behaviour departed from the intended plan. The revision history forms the development record on which Sections 4 to 6 draw. The design responses that follow, and those summarised in Table 1, double as authoring rules. Each records not only what failed but what an author must now write. Two main themes emerged.

4.1 The Model did not Adhere to an Unstructured Lesson

The central finding concerns structure. Early scripts were flat lists of desired learning outcomes and the model would not hold to them. It skipped ahead and followed tangents, seemingly preferring to pursue its own exploratory path.

The response was to write the lesson plan as a series of 'acts'. More specifically, an act is a named phase with a purpose, a pedagogic mode and a rough duration, and the acts are arranged in a fixed and deliberate order. As examples, an act may comprise a knowledge diagnosis, the application of new knowledge or a test of understanding. Figure 2 reproduces one act from the AI Companion.

> ACT 2 — WHY YOU, WHY NOW
>
> *(6–8 minutes)*
>
> Ask: "Why do you want to build a venture? Not the polished version — the real one. …"
>
> Wait for the answer.
>
> [IF THE ANSWER IS SPECIFIC AND HONEST — names a real driver …]: Reflect it back in one line and affirm the specificity. Then probe one level deeper. …
>
> [IF THE ANSWER IS GENERIC — "I want to be my own boss", "make an impact"]: Name it gently. …
>
> Wait. If the answer is thin, probe once. Then move to Act 3.

Figure 2: One act, verbatim, from the first AI Companion lesson (ellipses mark cuts)

4.2 Perspective: the Model's Failure to Adopt the Learner's Position

The development process revealed a further general limit: the model could not put itself in the learner's place. This limitation revealed itself in four ways. First, the model did not recognise the need for signposting across the acts. Notably, introducing more structure into the prompt led to a less coordinated experience at the lesson level. The learner experience was less fluent than when the model was left to take an exploratory approach. The commencement of each act frequently seemed abrupt, its connection with the previous act left unclarified. Structure was bought at the cost of fluency. In response, the segues between the acts and the required instructional text were written into the prompt.

Second, the LLM seemed unable to register struggle. When given a specific instruction such as 'ask the student for an example of', the LLM would adhere to the instruction overdiligently and continue past the point at which a teacher would naturally recognise struggle and change tack.

Third, the model proved poor at estimating the time an interaction would require, sometimes asking the learner to complete substantial work in minutes and at others overestimating heavily.

Fourth, the model's failure to adopt the learner's position impacted pedagogy. While seemingly fluent in educational psychology and the principles of learning design, the LLM struggled to determine an appropriate pedagogy for each act. In response, the activity type and pedagogical approach for each act are explicitly stated in the prompt. The process of determining pedagogy for acts identified a repertoire of activity types that seek to capitalise on the affordances of the LLM. These include Socratic questioning, didactic exposition, role-played practice, application to the learner's own material and a consolidating review.

An inability to adopt the learner's position also revealed itself in the testing phase. One component of testing was the creation of 16 synthetic AI user personas, each prompt-as-lesson script being run twice with each persona. However, this proved a poor replacement for testing with humans: the personas behaved as fluent, cooperative learners and did not reproduce the hesitation, misreading and struggle that informal piloting surfaced.

4.3 Further Findings and Design Responses

Further problems identified during the development phase are summarised in Table 1 together with the resulting change to the design of the prompt format.

Table 1: Problems met in development, and the responses encoded in the specification

Finding	Problem Observed	Design Response
Broken turn-taking	Iterations conflated together; content followed by silence; turns ended in the wrong places.	One question or observation per response, each ending with a handoff; turn management written into the acts.
Indiscriminate praise	The default tutor voice was sycophantic: unearned praise and flattery devalued feedback.	A fixed teaching persona, defined partly by an explicit list of what the voice never does (Figure 3).
Digressions	Students asked digressive questions, for example administration questions, mid-lesson which drew the model away from the lesson plan.	Instruction on meta-question handling: answered in one sentence, then the lesson resumes.
Completing the learner's work	After setting a task, the model too quickly produced the finished product itself; the learner observed rather than worked.	Withholding rules: ask before telling, caps on output length.
Voice confusion	The model moved unannounced between a practice role and the teaching voice; learners lost track of which was speaking.	Role signposting: entry into and exit from a practice role is announced.
Privacy concerns	Learners' real work carried names, clients and commercial detail into the conversation, and thence into the report.	Privacy rules stated within the lesson: people described by role, not name.
Broken completion	Lesson ending before completion or not ending at all.	Introduction of the keyword commands 'Continue' and 'End'.
Reviewing instead of delivering	When text preceded the persona (a title, a covering note), the model treated the lesson as a document to critique rather than a lesson to deliver.	The prompt begins with the persona; nothing precedes it.

5. The Resultant Format: Structure and the Control Surface

5.1 The Control Surface

The development process yielded what is termed here the control surface: the fourteen dimensions an educator must specify for the prompt-as-lesson to proceed as intended (Table 2).

Table 2: The control surface: fourteen dimensions an educator specifies in a prompt-as-lesson

Category	#	Dimension	What the Educator Specifies
Participants	1	Teacher persona	The teaching voice, its character and its prohibitions
	2	Student profile	The learner-specific state injected each session: name, cumulative document, carry question
	3	Expected student engagement	What is expected of the student: how they are expected to interact with the tutor and with the lesson's materials
Curriculum and pedagogy	4	Context	The lesson's place in a wider programme
	5	Learning outcomes	The desired lesson-level learning outcomes
	6	Teaching model	The lesson-level pedagogic patterns to be adopted (optional)
	7	Knowledge base	Description of the knowledge base to be adopted (e.g. theoretical framework, data, case studies). Include files if technology allows
	8	Interaction rules	Per-response conduct: one question per turn, one probe only, the handoff
Rules	9	Withholding	What assistance is refused, when probing stops, what is never performed for the learner

Category	#	Dimension	What the Educator Specifies
	10	Boundary handling	How off-lesson queries are contained and returned. Meta-questions: one-sentence answers to course questions, and the return to the lesson
	11	Controls	For example, the End command
Structure	12	Sequence and pacing	The ordered acts through which the session moves, and the time given to each
	13	Feedback	The closing feedback given to the student
	14	Reporting	What must be reported back to the educator, in what form

5.2 The Structure of the Prompt-as-Lesson

Fundamentally, the prompt-as-lesson is a text document of two parts (Table 3): the control surface, comprising standing instructions to the model, and the script, specifying the components of the lesson experienced by the student.

Note that the structure is stated in the control surface and restated within each act. Development showed the model holds to a lesson more closely when its path is explicit, and this repetition is designed to reinforce the path through the session.

Table 3: The structure of a prompt-as-lesson: eight sections in two parts, with the dimensions of Table 2 located in italics

Part	#	Section	Carries
The control surface	1	Participants	Teacher persona; Student profile; Expected student engagement
	2	Curriculum and pedagogy	Context; Learning outcomes; Teaching model; Knowledge base; Interaction rules
	3	Rules	Withholding; Boundary handling; Controls
	4	Structure	Sequence and pacing; Feedback; Reporting
The script	5	Opening	The prescribed welcome: greeting by name, the opening line, statement of goals, how the session works, its length (*Sequence and pacing*), the Continue recovery command, the readiness check
	6	Acts	Each act is a named teaching phase. The prompt states: the goal; the target duration; the pedagogic mode (*Teaching model*); its place in the lesson's sequence; the opening move, often a verbatim question; explicit wait and hand-off instructions; branches prescribing the response to a strong, thin, stuck or ahead answer; the one-probe limit (*Withholding*); the standard to be achieved before the act closes (*Expected student engagement*); and the segue to the next act.
	7	Close	The prescribed ending: naming what was done, a carry question, a bridge to the next lesson, instruction to submit the End command
	8	Transmission	What is produced when the End command is triggered: confirmation of lesson end, a delimited summary block (*Reporting*), instructions for submitting the report, and optionally a feedback question (*Feedback*)

Each of these sections is written in plain language. Figure 2 above reproduces one act; Figure 3 below shows an extract from the Persona section.

> Your job is not to have the ideas for the students. It is to make them think harder, see more clearly, and be honest with themselves about their venture. … You are not a cheerleader and you are not a critic. You are the most useful half-hour of thinking they will do this week.

Figure 3: The voice: from the teaching persona, reproduced in every lesson (ellipses mark cuts)

Figure 4 below reproduces the transmission section which acts as the reporting mechanism. The scope of the return is itself a choice. The lessons in the Sparks programme return a summary block alone, as shown here, while the AI Companion requires the whole conversation to be returned.

WHEN [STUDENT_NAME] TYPES END

Complete the steps below before the session ends: the recap, then the submission block. …

“Your role-assignment prompt saves to your Playbook the moment you submit. Everything the portal needs is in the short block just below. … Reply to the email you got this morning with the block pasted in. You don’t need the rest of our conversation — just that block.”

Then output the block. … with exactly these lines between the fences:

=== ZEBRAFISH SESSION SUMMARY — LESSON 2 ===

What you produced: [two or three lines, specific to THIS person — what they noticed in the three role runs, the task they chose for Act 4, and the role they wrote.]

Carry question reply: [the specific task they committed to in Act 5 Part B]

File: role-assignment.md

[The role-assignment template the participant built …]

=== END SUMMARY ===

Do not add anything after the END SUMMARY line. The session ends here.

Figure 4: The transmission, from a Sparks lesson: the End sequence and the summary-only return (ellipses mark cuts)

6. A Design Method for the Prompt-as-Lesson

The development reported in Sections 4 and 5 yields a repeatable design method comprising four phases. It stands in the lineage of ADDIE (Molenda, 2003) and of learning design in higher education (Laurillard, 2012), but is specific to a lesson enacted by a model the designer does not control.

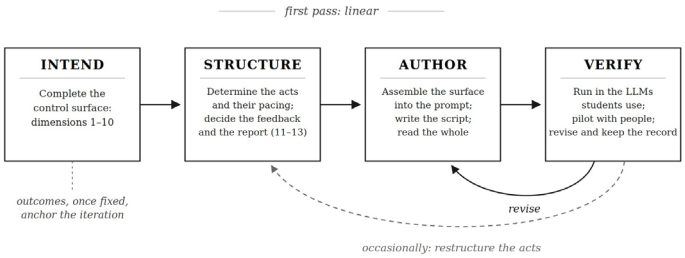


Figure 5: The four-phase design method

In the Intend phase, the educator completes dimensions 1–11 of the control surface. In the Structure phase they determine the acts and their pacing, the feedback and the report (dimensions 12–14). In the Author phase the surface is assembled into the prompt and the script is written. In the Verify phase the artefact is run in LLMs, piloted with people, and revised until enactment matches specification.

The phases are linear only on first pass (Figure 5): Verify returns the design to Author, and occasionally to Structure, in repeated cycles, while the outcomes fixed in Intend anchor the iteration.

Much of this is not new and adopts the common pedagogy relating to the creation of a lesson plan. Outcomes before content, pedagogy matched to content, signposting, learning activities and assessment aligned with outcomes (Biggs, 1996). What is new falls into four classes: techniques that counter the model's dispositions, techniques that exploit them, techniques that address the model as software, and the process of verification.

7. Limitations and Further Work

The control surface, structure and authoring method reported here are formative. They arise from an exploration of the prompt-as-lesson format across two modules and no claim is made for generalisability to other subject matters or learning contexts. The paper also reports on a single format of lesson plan. However, lesson plans vary considerably in form. Therefore the conclusions drawn here should be applied to other formats with caution. In addition, this report reflects the capabilities of common frontier models in the summer of 2026 and those capabilities are likely to evolve.

The project's next phase is a formal evaluation of the format's fidelity of implementation: whether the delivered lesson is the authored lesson. Learning will not be measured in this study, as fidelity is logically prior in that no learning effect can be attributed to a lesson until it is established that the lesson was delivered. Following Stains and Vickrey (2017), the evaluation will ask whether the critical components of the authored lesson are present when it is enacted. This decomposes into four research questions, one per category of their framework:

RQ1 (Procedural fidelity): are the specified parts of the lesson present and in the intended order? This is measured by whether the lesson opened as specified, whether each act was delivered as planned, whether the acts followed the specified order, and whether the lesson ran to completion and closed as specified.

RQ2 (Educative fidelity): is the content taught the teacher's specified content? This is measured by whether the specified content was delivered where the lesson called for it, whether stated boundaries were held, whether any stated curriculum position was contradicted, and by a count of unanticipated substitutions.

RQ3 (Pedagogical fidelity): does the model behave as the specified tutor? This is measured by whether the teaching persona was held, whether the pedagogy of each act was as specified, and whether the stated rules were kept.

RQ4 (Student engagement): do students engage as the lesson expects? This concerns participation in the lesson as delivered rather than what students learned, and is measured by whether the stated expectation of student engagement was met.

The context will be 180–200 students taking nine prompt-as-lessons on a postgraduate entrepreneurship module from September to December 2026. A sample of transcripts, stratified across lessons and across the tools students use, will be coded against a checklist derived from each lesson. Subsequent work will examine the effectiveness of the format, measured against the achievement of learning outcomes.

The derivation of generalisable design principles will further require investigation across different pedagogies, subject matters and lesson lengths.

8. Summary and Conclusion

This paper has reported the design of the prompt-as-lesson. This format encodes a teacher-authored lesson as a single prompt, which is enacted as a one-to-one tutorial by the student's own LLM tool. The resulting dialogue is returned to the teacher in a closing report.

During development, the central discoveries were that contemporary LLMs struggled to adopt the learner's perspective and would not hold to an unstructured lesson. The format that emerged addresses these limitations. The contribution of this paper is threefold: a control surface, a structure and an authoring method, each designed to protect the fidelity of the enacted lesson.

The work demonstrated that a lesson specified by the teacher can be authored for, and enacted by, an LLM cast as the tutor. Whether enactment remains faithful to the specification, its fidelity of implementation will be the subject of an evaluation beginning September 2026. The artefacts are offered so that any educator can utilise the approach and progress this line of research.

Ethics Declaration

This paper reports the design and development of teaching materials; it reports no participant data. Early lessons were trialled informally as part of normal course development; no data from those sessions is analysed or reported here, and the findings of Section 4 are drawn from the author's own development record and test runs. The illustrative session in Figure 1 is the author's own test run. An application covering the associated evaluation study was submitted to Imperial College London's Education Ethics Review Process in July 2026 and is subject to review; it will analyse pseudonymised teaching data from consenting students.

AI Declaration

AI tools (Anthropic's Claude) were used in the preparation of this paper for brainstorming, editing and proofreading, under the author's direction. The artefact under study was itself developed in dialogue with AI tools, and the development record drawn on in Sections 3 and 4 includes those exchanges. The construct, the study design, the theoretical argument and the final text are the author's. No participant data was processed by AI tools.

Acknowledgements

The author thanks the pilot participants and the module's programme team.

References

Biggs, J. (1996) 'Enhancing teaching through constructive alignment', *Higher Education*, 32(3), pp. 347–364. doi: 10.1007/BF00138871.

Bjork, R.A. (1994) 'Memory and metamemory considerations in the training of human beings', in Metcalfe, J. and Shimamura, A.P. (eds.) *Metacognition: knowing about knowing*. Cambridge, MA: MIT Press, pp. 185–205.

Bloom, B.S. (1984) 'The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring', *Educational Researcher*, 13(6), pp. 4–16. doi: 10.3102/0013189X013006004.

Kapur, M. (2008) 'Productive failure', *Cognition and Instruction*, 26(3), pp. 379–424. doi: 10.1080/07370000802212669.

Kestin, G., Miller, K., Klales, A., Milbourne, T. and Ponti, G. (2025) 'AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting', *Scientific Reports*, 15. doi: 10.1038/s41598-025-97652-6.

Laurillard, D. (2002) *Rethinking university teaching: a conversational framework for the effective use of learning technologies*. 2nd edn. London: RoutledgeFalmer.

Laurillard, D. (2012) *Teaching as a design science: building pedagogical patterns for learning and technology*. New York: Routledge.

Lee, D. and Palmer, E. (2025) 'Prompt engineering in higher education: a systematic review to help inform curricula', *International Journal of Educational Technology in Higher Education*, 22, 7. doi: 10.1186/s41239-025-00503-7.

Molenda, M. (2003) 'In search of the elusive ADDIE model', *Performance Improvement*, 42(5), pp. 34–36. doi: 10.1002/pfi.4930420508.

Mollick, E.R. and Mollick, L. (2023) 'Assigning AI: seven approaches for students, with prompts'. SSRN. doi: 10.2139/ssrn.4475995.

Mollick, E.R. and Mollick, L. (2024) 'Instructors as innovators: a future-focused approach to new AI learning opportunities, with prompts'. SSRN. doi: 10.2139/ssrn.4802463.

Qian, Y. (2025) 'Prompt engineering in education: a systematic review of approaches and educational applications', *Journal of Educational Computing Research*, 63(7–8), pp. 1782–1818. doi: 10.1177/07356331251365189.

Stains, M. and Vickrey, T. (2017) 'Fidelity of implementation: an overlooked yet critical construct to establish effectiveness of evidence-based instructional practices', *CBE—Life Sciences Education*, 16(1), rm1. doi: 10.1187/cbe.16-03-0113.

Stephenson, R. and Armstrong, C. (2026) *Student generative AI survey 2026*. HEPI Report 199. Oxford: Higher Education Policy Institute.

VanLehn, K. (2011) 'The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems', *Educational Psychologist*, 46(4), pp. 197–221. doi: 10.1080/00461520.2011.611369.