Reconfiguring Assessment in the Age of AI: Making Learning Visible through Meta-Task Awareness (MTA)

Lung-Hsiang Wong [1*], Guat Poh Aw[1]
[1] National Institute of Education, Nanyang Technological University, Singapore
[*] lhwong.acad@gmail.com

Abstract: *The widespread use of large language models (LLMs) is reshaping assessment practices, making it increasingly difficult to determine what learners actually know and can do, as learning processes become less visible in AI-mediated work. Rather than treating AI use as a problem of detection, this study explores how assessment can be reconfigured to make learning visible through learners' interaction with AI. Focusing on a postgraduate teacher education course, the study introduces Meta-Task Awareness (MTA) as an analytic lens for examining how learners regulate task goals, pedagogical considerations, and epistemic responsibility in LLM-supported lesson design. Using an exploratory qualitative design with descriptive code-and-count summaries, interaction trace data from 11 students were analysed at the episode level using a theory-informed coding framework. Findings show that students' engagement was primarily oriented toward task clarification, pedagogical reasoning, and evaluative judgment, with varying patterns resulting in three distinct orientations toward AI-mediated design. The study shows how assessment can be reconfigured so that learning becomes more visible, more accountable, and increasingly unavoidable in the age of AI.*
Keywords: *Generative AI in education; Human-AI collaboration; Meta-Task Awareness (MTA); AI-mediated assessment*

1. Introduction

The rapid diffusion of generative artificial intelligence (GenAI) has fundamentally disrupted learning and evaluation practices across educational contexts. Tasks commonly used for assessment – such as essays, lesson plans, design artefacts, or computer programs – can now be produced at high quality by Large Language Models (LLMs) within seconds. As a result, traditional product-oriented assessment approaches are increasingly challenged, raising concerns about the validity of evaluation in AI-rich learning environments (Yan et al., 2023).

Recent reviews suggest that much of the early research have focused on detection, restriction, or alternative assessment formats (Wong et al., 2024). However, such approaches often address surface-level symptoms rather than the underlying issue. Changing the form of assessment does not necessarily make learners' understanding, judgment, or decision-making processes more visible (Kim et al., 2024). The core challenge is not whether students use AI, but whether learning and evaluation designs can meaningfully capture how learners engage with tasks while working alongside AI.

This tension reflects a growing mismatch between product-oriented evaluation and process-oriented learning. There is a need to rethink assessment design so that AI use becomes an explicit condition of learning and evaluation, rather than an external disturbance to be eliminated. The challenge addressed in this study is not how to teach students to use AI more effectively, but how task and assessment design can reposition their epistemic responsibility when AI is already embedded in the learning process. This shift calls for approaches that foreground learners' regulation of tasks, their decisions about when and how to involve AI, and their capacity to revise and justify those decisions over time.

In this study, we argue that such engagement depends critically on learners' awareness of the task itself during AI-supported activity. We refer to this capacity as Meta-Task Awareness (MTA) (Chen & Wong, 2026; Wong, 2025a, 2025b), defined here as learners' ability to monitor, regulate, and reframe tasks while interacting with GenAI. Rather than

proposing MTA as a fully developed theory, we introduce it as an analytical lens for examining student-AI interaction in complex tasks.

Focusing on a graduate-level teacher education context, this paper reports a proof-of-concept qualitative analysis with descriptive code-and-count summaries of students' interactions with GenAI during pedagogical design tasks. By examining how MTA is manifested across task-oriented activities, this study elucidates how students differ in their engagement with AI during lesson design, and contributes to ongoing discussions on how learning and evaluation can be reconfigured in AI-rich environments.

2. Beyond Prompting: Meta-Task Awareness (MTA) in AI-Supported Learning

Recent discussions on GenAI in education have increasingly focused on prompting practices. From prompt engineering strategies to prompt literacy frameworks, a growing body of work has examined how learners formulate inputs to elicit more useful or accurate responses from LLMs (e.g., Murray, 2025). While these discussions have provided valuable insights into human-AI interaction at the surface level, they often foreground the optimisation of prompts rather than learners' understanding and regulation of the tasks for which AI is employed.

This emphasis reflects an assumption that better prompts lead to better learning outcomes. However, prompt quality alone does not guarantee meaningful engagement with the underlying task. Students may produce sophisticated prompts while remaining unclear about task goals, pedagogical constraints, or the division of cognitive labour between themselves and AI. Thus, AI becomes a shortcut to task completion rather than a partner in learning, leaving students' reasoning processes opaque to instructors and evaluators. The key challenge is not how well students prompt AI, but how they understand the task itself and manage their engagement with AI across different stages of task completion.

Building on this line of thinking, we introduce MTA (Wong, 2025a, 2025b) as an analytic lens for examining learner–AI interaction. MTA refers to learners' capacity to remain aware of the task they are working on while interacting with AI, including their ability to **(a) recognise task goals and constraints, (b) decide which aspects of the task to delegate to AI and which to retain under human judgment, and (c) reframe or revise the task through iterative interaction**. It refers to learners' awareness of task ownership and epistemic responsibility in human–AI collaborative task environments, where traditional assumptions about agency and control no longer hold.

Conceptually, MTA is grounded in three strands of existing work. First, it builds on research on task and metacognitive regulation (Flavell, 1979), which emphasises learners' monitoring and control of goals, constraints, and strategies during task performance (Zimmerman & Schunk, 2001). Yet unlike metacognition, which concerns monitoring and regulation of one's own cognitive processes, MTA concerns learners' awareness of task framing, epistemic responsibility, and the distribution of agency between human and AI within a task. Second, it draws on perspectives on human-AI coordination, where effective performance depends on learners' ability to make informed decisions about task delegation and control when interacting with intelligent systems (Endsley, 2017). Third, it aligns with process-oriented views of learning and evaluation, which foreground learners' engagement and reasoning processes rather than final products (Nicol & Macfarlane-Dick, 2006). Together, these perspectives provide the theoretical underpinning for using MTA as an analytic lens in AI-supported learning contexts.

In the present study, MTA serves as a conceptual bridge between discussions on AI-supported learning and the practical problem of evaluation design. By analysing how MTA is manifested in students' interactions with GenAI during pedagogical design tasks, this study seeks to provide an empirical basis for understanding how learners move beyond surface-level prompting toward more reflective and regulated forms of human-AI collaboration.

3. Research Context, Assignment Design, and Research Design

3.1. Research Context and Course Design

This study was situated in a postgraduate-level course on technology-enhanced Chinese language teaching offered at teacher education institution in Singapore. The participants were 11 master's students aged between 23 and 30. The

majority came from humanities backgrounds such as linguistics or Chinese studies, and had no formal teaching experience at the time of the study. All participants were already frequent users of LLMs for academic or personal purposes. However, their use of AI was unevenly integrated into pedagogical thinking.

The 13-session course was grounded primarily in techno-pedagogical knowledge for language education, including language learning theories, technologies as cognitive tools, multimedia-assisted learning, mobile learning, computer-supported collaborative learning, and design for formative assessment. GenAI was introduced only in the later phase of the course, such that the assignment assessed students' ability to synthesise and enact these pedagogical foundations – rather than isolated prompt engineering skills – within AI-supported design contexts.

A key design principle of the course was the normalisation of AI use. Rather than restricting AI, students were required to use AI tools as part of their learning activities. This decision was grounded on the premise that contemporary teacher education cannot be meaningfully conducted by assuming the absence of AI. Instead, the pedagogical challenge lies in supporting future educators to engage with AI in ways that are reflective, purposeful, and pedagogically grounded.

3.2. Assignment and Evaluation Design: From Product to Process

The focal assignment was deliberately designed to operationalise MTA within a lesson design context, not through explicit instruction, but through structural features embedded in task requirements. Students were required to design and iteratively improve a lesson plan and associated learning resources, beginning with an initial draft generated using LLM-based tools, followed by three weeks of revision and refinement through iterative prompting and critical evaluation.

Crucially, assessment did not prioritise the quality of the final artefact alone. Instead, evaluation foregrounded the process of transformation from the initial LLM-generated version to the final student-submitted design. This delta-based evaluation approach – drawing on the notion of delta (Δ) as change over time – treated learning as evidenced through students' design decisions, revisions, and justifications across iterations, rather than through end products in isolation.

Within this design, LLM-generated first drafts functioned as epistemic starting points rather than authoritative solutions. Students were required to engage with these drafts at the lesson level by specifying lesson objectives, evaluating generated suggestions, and making pedagogically grounded decisions about whether to adopt, modify, or reject AI outputs. Importantly, they were required to conduct lesson planning and resource development within a continuous interactional thread, ensuring that design reasoning unfolded cumulatively rather than as fragmented, task-by-task exchanges. This continuity made their evolving task understanding and decision-making processes observable in their interaction traces.

Rather than teaching metacognitive skills explicitly, the assignment structurally activated meta-level engagement by making epistemic responsibility unavoidable. To demonstrate learning, students had to justify design changes and show coherent alignment between pedagogical goals, learner needs, and revised lesson components across iterations. In this way, MTA was enacted through students' awareness of task ownership and epistemic authority in lesson design, as reflected in how they negotiated control with LLMs and exercised judgment over AI-generated content. The assignment thus shifted evaluation from product-oriented performance to process-oriented evidence of learning, positioning MTA as an emergent outcome of task and assessment design rather than as a taught strategy.

3.3. Research Questions and Research Design

This study adopts an exploratory, proof-of-concept orientation to examine how students engaged with LLMs in lesson design tasks. The purpose of this study is qualitatively analytic rather than statistical generalisation. The analysis is underpinned by MTA, used primarily as an analytic lens rather than as a validated psychological construct. Rather than evaluating the quality of AI-generated outputs, the study focuses on how students demonstrated pedagogical reasoning, task ownership, and epistemic control through their interactions with LLMs. Methodologically, the study employs a qualitative, process-oriented analytical design (Chi, 1997), prioritising analytic depth and trace-level resolution over sample size. Students' interactions with LLMs were analysed at the level of lesson design episodes, defined as coherent segments of student-LLM interaction oriented toward a specific lesson design objective or sub-task (e.g., lesson structure,

learning activities, assessment rubrics, instructional materials, or simulated learner responses). Within this qualitative analytic framework, a theory-informed coding scheme was applied to identify observable manifestations of MTA, and descriptive code-and-count summaries were used to surface overall distributional patterns across participants. Accordingly, the study addresses the following research questions (RQs):

RQ1. How is MTA manifested in students' LLM-mediated lesson design tasks?

This RQ examines how MTA is enacted during students' interactions with LLMs in lesson design tasks. It focuses on observable process-level behaviours through which they exercised task ownership, pedagogical reasoning, evaluative judgment, and epistemic control when framing prompts, responding to model outputs, and revising design decisions. The analysis attends to how these actions emerged during iterative lesson design rather than to final artefacts.

RQ2. How do patterns of MTA vary across students with differing orientations toward AI-mediated lesson design?

Building on the manifestations identified in RQ1, this RQ examines how students differed in the patterning of MTA during AI-mediated lesson design. Adopting a level-based analytic approach, the analysis compares how task clarification, pedagogical justification, evaluative judgment, and epistemic authority were configured across students' LLM interactions. These are used to characterise distinct orientations toward engaging with AI in lesson design contexts. The analysis does not aim to model relationships among codes as independent variables, but to interpret recurring configurations of co-occurring practices within interaction traces.

Coding framework

To address RQ1, a theory-informed coding scheme was developed to operationalise MTA as observable practices in students' LLM-mediated lesson design tasks. Rather than treating MTA as an abstract disposition, the analysis focused on how students enacted task ownership, epistemic responsibility, and pedagogical decision-making through their interactions with LLMs. In addition to qualitative interpretation, a descriptive code-and-count summary was produced to surface overall distributional patterns. Frequencies were tallied at the episode level (i.e., each coded instance within a task/episode), and then aggregated across students for reporting in the Findings. Episodes were defined pragmatically based on shifts in lesson design objectives, rather than fixed interaction lengths. For analytic clarity, multi-valued codes were collapsed into binary presence/absence for frequency reporting.

The coding scheme was derived deductively from the conceptual framing of MTA and inductively refined through close examination of student–LLM chatlogs. The scheme comprised four core coding dimensions, each with clearly defined indicators and value categories, as summarised in Table 1. These dimensions captured (a) how lesson design tasks were specified and constrained, (b) how LLM-generated outputs were evaluated, (c) how pedagogical decisions were justified, and (d) how control over the lesson design process was distributed between the student and the LLM. Coding was carried out by a researcher and independently verified by the course instructor through a parallel pass over the data, yielding highly consistent code distributions with over 90% agreement across coded instances.

To address RQ2, coded tasks were examined across students using a qualitative aggregation approach. Rather than constructing formal typologies or validated scales, the analysis aimed to identify emergent levels of engagement with LLMs in lesson design. Students were compared in terms of the consistency, diversity, and depth of MTA-related coding patterns observed across their tasks. Given the nature of the task, the study assumes functional equivalence across mainstream LLMs and focuses on students' pedagogical reasoning as manifested through their prompts and revisions.

Through this analytical design, the study provides a concrete illustration of how MTA can be operationalised, observed, and compared in LLM-supported lesson design contexts, offering an initial proof of concept for MTA as a lens for learning and evaluation in generative AI–mediated education.

Table 1: *The coding scheme for qualitative analysis to address RQ1*

Code	Dimension	Description	Values
C1	Task specification	Whether the student explicitly specified or refined the lesson design task, goals, or constraints, rather than relying on default AI interpretations.	0 = Absent; 1 = Present
C2	Evaluation of LLM output	Whether the student critically evaluated, accepted, rejected, or questioned LLM-generated suggestions.	0 = Absent; 1 = Present
C3	Pedagogical justification	Whether pedagogical rationales were explicitly articulated to justify lesson design decisions, and whether such rationales were primarily initiated by the student, prompted by the LLM, or jointly constructed.	0 = Absent; 1 = LLM-led; 2 = Student-led; 3 = Co-constructed
C4	Epistemic authority	Whether the student exercised primary decision-making authority over lesson design choices, including overriding, restructuring, or redirecting LLM-generated outputs in accordance with pedagogical goals and constraints.	0 = Absent; 1 = Student-led

4. Findings

4.1. RQ1: Manifestations of Meta-Task Awareness in AI-Mediated Design Processes

To address RQ1, all student–LLM chatlogs were coded using six process-oriented codes representing different dimensions of MTA. Table 2 presents the aggregated frequencies of each code across the eleven participants.

Table 2: *Aggregated frequencies of MTA-related codes across students' LLM-mediated lesson design tasks*

Code	Description (concise)	Frequency (n)	Illustrative example (excerpt)
C1	Task specification	50	"The lesson is for Grade 8 Chinese Language students with mixed proficiency. Redesign the activity to focus on inferencing rather than vocabulary recall."
C2	Evaluation of LLM outputs	42	"This activity looks engaging, but it may be too demanding for weaker students. Can we simplify the output requirement?"
C3-1	Pedagogical justification (student-initiated)	19	"Insert a guided discussion here because students need support before independent writing."
C3-2	Pedagogical justification (co-constructed)	25	"If the goal is text appreciation, then the task should allow multiple interpretations rather than fixed answers."
C3-3	Pedagogical justification (LLM-initiated)	9	"You are right that peer discussion can lower affective filter. I will keep this step."
C4	Epistemic authority in lesson design	22	"Keep the structure, but remove the Kahoot! activity – it distracts from the main objective."

Overall, a total of 167 coded instances were identified. Among the six codes, C1 (task clarification and goal refinement) and C2 (pedagogical reasoning and justification) were the most frequently observed. These were followed by C4-1 (evaluation and critique of AI-generated outputs), and C3-2 (iterative refinement of prompts or outputs. In contrast, C3-3 (explicit meta-level reflection on AI use) occurred least frequently.

The distribution of codes suggests that students' engagement with AI was primarily oriented toward task-level and pedagogical decision-making, rather than abstract reflection on AI itself. High frequencies of C1 and C2 indicate that such actions were frequently observed across the dataset, with students articulating lesson goals, constraints, and pedagogical rationales when interacting with LLMs. Codes associated with evaluative judgment and selective adoption of AI outputs (C4) were also consistently observed across participants, indicating active scrutiny rather than uncritical acceptance of generated content.

Indeed, differences among students were not only reflected in the total frequency of coded behaviours, but also in the relative balance between generative, evaluative, and reflective actions. This variation provided an empirical basis for the level-based analysis reported in RQ2.

4.2. RQ2: Different MTA patterns across students with differing orientations toward AI-mediated lesson design

Building on the manifestations of MTA identified in RQ1, RQ2 examines how these process-level behaviours were configured differently across students, resulting in distinct orientations toward AI-mediated lesson design. Rather than treating MTA as a unidimensional construct, a level-based analytic grouping was conducted to capture how combinations of task clarification, pedagogical justification, evaluative judgment, and epistemic authority co-occurred within student-LLM interactions. Based on the relative density, balance, and functional roles of coded behaviours, students were

categorised into three MTA levels for analytic convenience. The three levels do not represent developmental stages or stable learner traits, but analytically derived groupings grounded in recurrent patterns of MTA enactment observed within the focal task. Table 3 summarises the mean per-student frequencies of MTA-related codes within each MTA level (not aggregated totals), with group sizes indicated by *n*.

Table 3: *Mean per-student frequencies of MTA-related codes by MTA level (n = number of students in each level)*

Code	C1	C2	C3-1	C3-2	C3-3	C4	Mean Total
Level 1 (*n* = 2)	6.5	5.5	3	2.5	1	2.5	21
Level 2 (*n* = 7)	4.7	4	1.7	2.6	1	2.1	16.1
Level 3 (*n* = 2)	2	1.5	0.5	1	0	1	6

Level 1: AI as a generative executor (Low MTA)

Students in Level 1 (*n*=2) primarily engaged with LLMs as content generators. Their interactions were dominated by task execution and material production, with limited evidence of student-initiated task refinement or evaluative intervention. Pedagogical justification, when present, was typically reactive rather than self-initiated, often emerging only in response to AI-generated suggestions. Across these cases, epistemic authority over lesson design decisions resided largely with the LLM. Prompts tended to specify outputs rather than constraints or pedagogical intentions, and AI-generated content was more often directly adopted. MTA at this level was characterised by minimal task ownership and limited reflective control over the design process.

Level 2: AI as a pedagogical assistant (Moderate MTA)

The majority of students (*n*=7) fell into Level 2, exhibiting a more balanced pattern of MTA manifestations. These students frequently articulated lesson goals and constraints, provided pedagogical rationales, and engaged in iterative refinement of prompts or outputs. Evaluative judgment was evident in selective adoption, modification, or rejection of AI-generated suggestions. While the LLM played an active role in idea generation, students at this level retained decision-making authority over pedagogical directions. AI was treated as a supportive assistant rather than a driver of design. Their MTA was thus characterised by sustained task ownership and functional integration of AI contributions within students' pedagogical reasoning processes.

Level 3: AI as a dialogic design partner (High MTA)

Students at Level 3 (*n*=2) demonstrated the most sophisticated configurations of MTA. Their interactions were marked by dense and diverse combinations of task clarification, pedagogical justification, iterative refinement, and explicit evaluation of AI outputs. These students actively shaped the trajectory of the design process, using AI responses as objects for reflection, comparison, and pedagogical decision-making. At this level, epistemic authority was clearly retained by the student. LLM-generated content functioned as a catalyst for design-level reasoning. Although explicit meta-level reflection on AI use was not frequent, reflective control was implicitly enacted through students' consistent regulation of task goals, constraints, and design quality across iterations.

5. Discussion and Implications

This study responds to recent calls for more process-oriented and methodologically grounded research on GenAI in education, particularly work that foregrounds human–AI collaboration and rethinks assessment. The findings from RQ1 and RQ2 suggest that MTA is neither a generic metacognitive trait nor a function of prompt sophistication. Rather, it is a structurally activated orientation toward epistemic responsibility in AI-mediated work.

5.1. MTA as structurally activated, not instructionally taught

A key insight from the findings is that higher-level MTA did not emerge because students were explicitly taught how to "think about AI use". Instead, meta-level engagement was activated by the design of the assignment itself. By anchoring assessment on design transformation (Δ) rather than final products, students were compelled to demonstrate judgment, revision, and justification across iterations. In this sense, MTA functioned as a necessary condition for demonstrating learning, not an optional add-on.

This helps clarify a conceptual gap between metacognition and MTA. While MTA involves metacognitive processes, it is not framed as an internal strategy to be explicitly trained or self-reported. Rather, it is observable through task-level decisions made under epistemic accountability, as evidenced in interaction traces. Meta-level engagement was not prompted instructionally ("reflect on your AI use"), but elicited structurally through task and assessment design.

5.2. AI as both amplifier and mirror of pedagogical thinking

Across levels, the data acted simultaneously as intelligent amplifiers and mirrors of students' thinking. Students who demonstrated stronger MTA were not those who produced longer or more complex prompts, but those who integrated pedagogical goals, learner considerations, and evaluative judgment when deciding how to engage with or override AI.

Students' ability to critique and revise AI-generated lesson drafts did not arise in a vacuum. Prior to the AI-based assignment, they had already engaged in repeated cycles of lesson critique and peer negotiation across the course. These earlier activities fostered pedagogical diagnostic criteria – such as sensitivity to tool overload, misalignment with learner needs, or weak assessment design – which were later mobilised in AI-mediated contexts. The assignment thus did not reward pre-existing "AI talent", but activated already-constructed pedagogical judgment through structural accountability.

5.3. Differentiated MTA orientations in AI-mediated lesson design

The level-based analysis in RQ2 further illustrates that students differed not only in the frequency of MTA-related behaviours, but in their overall orientation toward AI-mediated lesson design. Lower-level MTA was characterised by efficiency-oriented use of LLMs as generators or executors of tasks, with design authority largely ceded to the model. Higher-level MTA, in contrast, involved sustained task ownership, evaluative judgment, and epistemic control.

These levels should not be interpreted as fixed traits or ability rankings. Instead, they represent situated orientations shaped by task structure and evaluative demands. This reinforces the argument that MTA is design-sensitive: different assessment and interaction designs are likely to elicit different patterns of AI engagement, even from the same learners.

5.4. Design principles for AI-mediated, process-oriented assessment

Drawing from the findings, four design principles can be articulated for eliciting MTA in AI-mediated learning:

1. **Anchor evaluation on transformation (Δ), not end products:** Assess learning through changes across iterations, making epistemic engagement unavoidable.
2. **Treat AI-generated outputs as epistemic starting points, not authoritative solutions:** Require learners to evaluate, adapt, or reject AI suggestions in relation to learning goals.
3. **Require sustained work within a continuous interactional thread:** Design tasks that support iterative development within a single AI interaction, rendering evolving task ownership traceable.
4. **Activate meta-level engagement structurally, not instructionally:** Embed epistemic responsibility into task and assessment design rather than teaching metacognitive strategies explicitly.

Together, these principles reposition AI-mediated assessment from a problem of detection or control to an opportunity for reconfiguring how learning is evidenced.

5.5. Limitations

Several limitations should be noted. First, the study involved a small cohort within a specific teacher education context, and findings should be interpreted as a proof-of-concept rather than generalisable claims. The small sample size is appropriate for this qualitative, trace-based analysis, where analytic depth is prioritised over representativeness. Second, while rich interaction trace data were analysed, no direct measures of learning outcomes beyond task performance were included. Future work could examine how MTA relates to longer-term pedagogical development or transfer across tasks.

6. Conclusion

This study demonstrates that MTA can be meaningfully operationalised and empirically examined through interaction trace analysis in AI-mediated lesson design tasks. By shifting assessment from product- to process-oriented transformation, the study offers a viable pathway for addressing contemporary challenges in AI-era evaluation design.

More broadly, it suggests that the question is no longer whether AI should be allowed in learning, but how assessment becomes more visible, more accountable, and increasingly unavoidable in the presence of LLMs.

Acknowledgments

This project received no external funding. Consent from the research participants was obtained based on ethics approval by Nanyang Technological University (IRB ref: IRB-2025-658).

References

Chen, H.-C., & Wong, L.-H. (2026). The MTA-TPACK Dynamic Collaboration Spiral: Making Pedagogical Thinking Visible in Human–AI Scientific Visualization for Sustainable Teacher Innovation. *Sustainability*, *18*(6), 2718. https://doi.org/10.3390/su18062718

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, *6*(3), 271-315.

Endsley, M. R. (2017). From here to autonomy: lessons learned from human–automation research. *Human Factors*, *59*(1), 5-27. https://doi.org/10.1177/0018720816681350

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*(10), 906. https://doi.org/10.1037/0003-066X.34.10.906

Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on Generative AI-assisted academic writing. . *Education and Information Technologies*, *30*(1), 1265-1300. https://doi.org/10.1007/s10639-024-12878-7

Murray, M. (2025). Bridging the Generative AI Literacy Gap: A Guide to Introducing Prompt Engineering in University Courses. *Issues in Informing Science and Information Technology,*, *22*, Article 10. https://doi.org/10.28945/5536

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199-218. https://doi.org/10.1080/03075070600572090

Wong, L.-H. (2025a). The prompt is not enough: Rethinking AI-learner interaction in the post-prompting era, *Proceedings of International Conference on Computers in Education 2025 (Volume II)* (pp. 311-319), Chennai, India.

Wong, L.-H. (2025b). Transforming education with Generative AI: Designing for the Post-Prompting Era. *Information and Technology in Education and Learning*, *5*(1), Inv-p003-Inv-p003. https://doi.org/10.12937/itel.5.1.inv.p003

Wong, L.-H., Park, H., & Looi, C.-K. (2024). From hype to insight: Exploring ChatGPT's early footprint in education via altmetrics and bibliometrics. *Journal of Computer Assisted Learning*, *40*(4), 1428-1446. https://doi.org/10.1111/jcal.12962

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology,*, *55*(1), 90-112. https://doi.org/10.1111/bjet.13370

Zimmerman, B. J., & Schunk, D. H. (2001). Reflections on theories of self-regulated learning and academic achievement. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives (2nd ed.)*. Lawrence Erlbaum Associates Publishers.