

Article

AI as Scaffold or Shortcut? Metacognitive Regulation in Graduate Students' Educational App Design

Zaid Almajdoub ¹, Patricia Fidalgo ^{1,*}, Reem Hashem ² and Yahia Alramamneh ³

¹ Pedagogy, Technology and Teacher Education Division, Emirates College for Advanced Education, Khalifa City A SE-43, Abu Dhabi P.O. Box 295, United Arab Emirates; zaid.almajdoub@ecae.ac.ae

² Data Analytics, Policy and Leadership Division, Emirates College for Advanced Education, Khalifa City A SE-43, Abu Dhabi P.O. Box 295, United Arab Emirates; reem.hashem@ecae.ac.ae

³ Language, Cognition and Growth Division, Emirates College for Advanced Education, Khalifa City A SE-43, Abu Dhabi P.O. Box 295, United Arab Emirates; yahia.alramamneh@ecae.ac.ae

* Correspondence: pfidalgo@ecae.ac.ae

Abstract

The increasing use of artificial intelligence (AI) in education has raised important questions about how learners and teachers regulate their thinking when interacting with AI-supported tools. This study investigates metacognitive regulation during the design of AI-supported educational applications by graduate students ($N = 12$) in a UAE teacher education programme. The qualitative research adopted a multiple-case study method to analyse six student design teams involved in designing educational applications as part of a technology integration course. Data were collected from three main sources: the educational apps themselves, the corresponding Teacher Guides, and reflective interviews conducted using a semi-structured interview guide. Data analysis combined deductive qualitative coding based on the regulation-of-cognition dimensions of the Metacognitive Awareness Inventory (MAI) framework with inductive thematic analysis of participants' experiences and conceptualisations of AI. The findings show that AI was used differently across the six cases. It appeared to function more as a cognitive scaffold when participants questioned, checked, revised, or justified AI-generated outputs, and more as a shortcut when outputs were used mainly for convenience with limited further reflection. Planning was the most visible regulatory process, while stronger regulatory engagement was evident when AI-generated outputs were examined, questioned, and refined rather than accepted without further consideration.

Keywords: artificial intelligence in education; generative AI; educational technology design; metacognition; teacher education; technology integration



Academic Editors: José Fernández-Cerero and Marta Montenegro-Rueda

Received: 22 May 2026

Revised: 9 July 2026

Accepted: 22 July 2026

Published: 27 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Effective teaching and learning depend heavily on metacognition, which is generally defined as awareness and control of one's own thought processes (Stanton et al., 2021; Panadero, 2017). Metacognitive processes in teacher education enable pre-service teachers to intentionally plan lessons, monitor their decisions, manage dilemmas, and evaluate outcomes with respect to educational goals (Wass et al., 2023). These procedures are particularly crucial in areas like instructional design and technology integration that call for sophisticated judgement rather than adherence to protocol (Geitz et al., 2019).

It has been demonstrated that design-oriented tasks foster metacognitive engagement (Geitz et al., 2019). Teachers must explain their instructional intentions, defend their

pedagogical and technological decisions, and foresee the needs of their students when they are asked to create learning activities or educational resources (Chai & Koh, 2017). Both cognitive knowledge and cognitive regulation are activated by such tasks, which highlight how and why decisions are made. However, unless specifically scaffolded and examined, a large portion of this metacognitive activity remains implicit and is frequently underrepresented in traditional coursework (Wass et al., 2023; Geitz et al., 2019).

The context of technology use is itself a demanding one for metacognitive regulation. Although the domains of knowledge encompassed by such frameworks (e.g., Technological Pedagogical Content Knowledge, TPACK) highlight the need for knowledge to be integrated, they afford little understanding of the regulatory processes that teachers use to plan, monitor, and revise their decisions (Chai & Koh, 2017). Conversely, prescriptive planning models, for example, the Technology Integration Planning (TIP) model, are founded on the premise that goals, instructional strategies, theoretical underpinnings, and assessment methods are clearly articulated (Roblyer & Hughes, 2019). These stages align well with well-established metacognitive regulation models, allowing technology design activities to provide a useful context for investigating teachers' thinking-in-action (Dignath & Veenman, 2021; Panadero, 2017).

Above all, the growing involvement of Artificial Intelligence (AI) in education brings these metacognitive demands to a new level (Holmes et al., 2022; Zawacki-Richter et al., 2019). AI tools can facilitate thinking, organisation, and writing, but they raise new concerns about judgement, relevance, and ethical use (Akgun & Greenhow, 2022). For educators, AI is at once a design tool and a feature of the learning tools they develop. As a result, educators need to identify when AI can truly assist learning while watching for over-automation or a pedagogical mismatch (Ng et al., 2021; Holmes et al., 2022). Understanding how teachers regulate their thinking in these moments is therefore essential.

Although metacognition, educational technology, and AI are gaining more attention, they are usually studied individually. Little research has been conducted on the metacognitive regulation of graduate teacher-designers as they design AI-facilitated educational products, especially on the extent to which these metacognitive processes can be observed (or concealed) across different data sources (Chai & Koh, 2017; Wass et al., 2023). In addition, reports are often based on products or self-report measures, providing little information about monitoring, debugging, and evaluative activities during design (Zawacki-Richter et al., 2019).

Thus, this research focuses on the processes of metacognitive regulation among teacher education graduate students as they design educational applications involving AI. The following research questions guided the inquiry: (1) What metacognitive regulation processes were observable across artefacts, teacher guides, and interviews? (2) In what ways did the TIP framework make students' metacognitive decision-making visible or invisible in written design justifications? (3) How did students conceptualise the role of AI in supporting or constraining metacognitive regulation for themselves and their learners? (4) What design features and regulatory conditions were associated with stronger metacognitive profiles? The study adopts a qualitative methodology and is guided by a framework analysing metacognitive processes in design artefacts, structured planning reports, and reflective interviews. In response to the growing interest that teachers and researchers have in AI in education, this study aims to highlight the design process (not just the outcome) by which teachers make decisions related to technology and AI so that the metacognitive regulation that supports teachers' decision-making can be made visible (Holmes et al., 2022; Ng et al., 2021).

2. Literature Review

2.1. Technology Integration as a Metacognitive Practice

Metacognition has traditionally been understood as involving both knowledge of cognition and the regulation of cognitive activity (Flavell, 1979; Schraw & Dennison, 1994). Regulation includes processes such as planning how to approach a task, monitoring progress and understanding, changing strategies when difficulties arise, and evaluating outcomes. These processes are also central to models of self-regulated learning, which describe learning as a cyclical process in which goals, strategy use, monitoring, and reflection influence one another over time (Pintrich, 2002; Zimmerman, 2002). Importantly, metacognitive skills do not necessarily develop or become visible simply because learners are engaged in a challenging task. They often require explicit prompts, opportunities for reflection, and guidance that helps learners notice and examine their own thinking (Veenman et al., 2006). This is particularly relevant to technology design activities, where many decisions and adjustments may otherwise remain implicit.

Technology integration is a reflective and metacognitive practice and a technical issue. Though models such as the Technological Pedagogical Content Knowledge (TPACK) have provided broad direction, they tend to focus more on types of knowledge than on specific regulatory processes (Koehler & Mishra, 2012). Emerging research is starting to conceptualise technology integration as a metacognitive process, which entails intentional management of instructional decisions at the planning, implementation, and evaluation stages (Dignath & Veenman, 2021; Tondeur et al., 2020). Structured frameworks are necessary to make the invisible process of metacognitive decision-making visible and teachable. The TIP model provides a more metacognitively aligned scaffold (Roblyer & Hughes, 2019). The regulation of cognition is externalised in the TIP model, where the teacher is asked to identify a problem, map out strategies, ground choices in learning theory, anticipate outcomes, suggest assessments, and consider equity (Wang et al., 2025). Among the tasks that educators must carry out are planning (setting goals and strategies), information management (organising content and multimodal design), monitoring (making sure that the strategy and goal are in line), debugging (making the required adjustments when problems arise), and evaluation (judging results) (Ho & Lau, 2025).

There are also several points of connection between the TIP model and the models of self-regulated learning reviewed by Panadero (2017). In this study, however, Panadero's review was used to position the research within the broader literature on self-regulation rather than as the analytical framework. The empirical analysis was guided by the regulation-of-cognition dimensions of the MAI framework (Schraw & Dennison, 1994).

The value of the TIP model in this context lies in the way it encourages students to make parts of their design reasoning explicit. Asking students to identify the problem they intend to address, explain how they selected particular tools and teaching approaches, describe how they plan to implement the design, and consider how its effectiveness will be evaluated can make planning and evaluation more visible. Other processes, such as monitoring and debugging, may remain less explicit unless students are specifically asked to document the problems, adjustments, and revisions that occur during the design process.

These metacognitive processes can be developed through collaborative digital design activities (e.g., app design, simulation design, or multimedia lessons), as evidenced in recent studies of technology-enhanced teacher education. Recent research in teacher education has also highlighted the relationship between experiential learning, reflective practice, and metacognitive awareness (Ho & Lau, 2025). In addition, integrating reflective scaffolds such as report prompts, rubrics that focus on justification, or equity checklists results in more explicit reporting of decision-making (Novoa-Echaurren, 2025).

The planning and evaluation of teachers are usually well documented in reports, but monitoring and debugging are also less visible in written artefacts and are thought to become evident mainly through reflective interviews or think-aloud protocols (Wang et al., 2025). This suggests the value of a multi-source qualitative approach that combines artefact analysis with reflective interviews to examine different stages of metacognitive regulation.

2.2. Metacognition in Technology-Enhanced and AI-Mediated Learning

With the rise of AI in education, scholars are examining how AI systems both support and disrupt metacognitive processes. Through real-time guidance, progress visualisation, and adaptive recommendations, artificial intelligence-supported learning environments—such as adaptive quizzing environments and generative feedback tools—may be able to direct supervision and evaluation (Y. Yang & Xia, 2023). For example, AI tutors have been found to encourage students to reflect on their reasoning steps and compare alternative strategies, deepening both conditional knowledge and evaluation (Novoa-Echaurren, 2025).

Over-reliance on generative AI has recently been described as “metacognitive laziness,” referring to the possibility that learners may become less inclined to plan, monitor, and regulate their own thinking when too much of the cognitive work is shifted to AI (Y. Fan et al., 2024). The concept is still relatively new, and the evidence remains limited, so it should be interpreted with caution. Even so, it reflects a broader concern about how excessive reliance on AI may affect learners’ independent thinking and self-regulation. Similarly, although useful, personalisation features can impede productive struggle without explicit reflection prompts (Xu et al., 2025). There is widespread concern that learners will lose opportunities to practice strategic control due to the shifting of regulatory responsibilities away from learners because of automation.

There are two sides to this issue for teacher-designers. First, AI is used as a design tool (e.g., it supports thinking about design, generates design ideas and prototypes) and second, AI can be seen as a designed element of the products they produce. This dual level of interaction creates new challenges for designers from a metacognitive perspective, as they will not only have to manage their own use of AI but also their expectations of how learners will use AI elements in the final product.

The teachers’ perspectives also show that trust and agency are very important when using AI tools. In a study done across several countries, it was found that the extent to which teachers are willing to accept AI tools is based on their understanding of the technical aspects of the technology and, more importantly, their metacognitive processing of the AI responses—specifically, their levels of willingness to accept, modify or reject the information presented to them (Viberg et al., 2023; Wang et al., 2025). When teachers design with AI, their reflective choices about prompts, constraints, and feedback mechanisms constitute acts of metacognitive regulation that are crucial for equitable practice. Teachers must now possess the conditional knowledge to decide if AI is the appropriate solution to a learning problem. They have to keep an eye on the ethical ramifications, potential biases, and fairness of access regarding AI tools (Zawacki-Richter et al., 2019). They have to fix pedagogical misalignments brought on by the “black box” nature of many AI systems, in addition to technical ones. Importantly, they need to think about AI’s role in their own design process as well as in the metacognition of their future students. Will students’ critical thinking be replaced by AI tools, or can AI tools be built to support and improve it? Research on teachers’ AI readiness suggests that there is significant anxiety and a lack of conceptual understanding around these very questions (Biagini, 2025), making it a critical area for investigation.

2.3. *Designing as Thinking: Accessing Metacognitive Regulation Beyond Final Products*

Designing ICT products offers a unique window into teachers' metacognition because it distributes cognition across artefacts, collaborators, and reflective documents. The final product (e.g., an app) shows evidence of information management and procedural knowledge, while accompanying reports often highlight planning, theoretical alignment, and evaluation. However, monitoring and debugging, real-time adjustments, problem-solving, and iterative thinking are usually absent from final artefacts. Interviews provide the missing link, surfacing tacit decision-making and emotional regulation (e.g., frustration, persistence, or confidence).

For example, monitoring (identifying a mismatch) and debugging (putting a fix in place) are embodied when a teacher team changes an app feature after testing. Rejecting or improving AI recommendations demonstrates conditional knowledge and evaluative reasoning if AI was used to create the original content. These aspects risk going unnoticed if apps, reports, and interviews are not triangulated. For teacher-designers, AI is a dual-faced entity: it is both a tool for creation (e.g., using generative AI to brainstorm content or debug code) and a component of the product (e.g., designing an app with an AI-powered adaptive feedback system). This duality introduces new metacognitive demands.

This body of literature suggests that metacognitive regulation is an important part of technology integration, that AI creates both opportunities and challenges for such regulation, and that design tasks can make cognitive processes visible that might otherwise remain hidden. However, relatively little research has combined design artefacts, written justifications, and reflective interviews to examine how graduate teacher-designers regulate their thinking while developing AI-supported learning tools.

3. Materials and Methods

3.1. *Research Design*

This research applies a qualitative, multiple-case study design to explore how metacognitive regulation became visible as graduate students designed AI-supported educational applications. Since the study deals with processes rather than outcomes, such as planning, monitoring, debugging and evaluating, which are very often tacit and not easily captured through questionnaires and/or product-based analyses, a qualitative method of analysis was considered to be most appropriate.

The case study approach made it possible to examine metacognitive regulation within an authentic course-based design project, where cognition could be explored through artefacts, collaborative activity, and reflective accounts. Each student pair was considered a bounded case because both the educational application and Teacher Guide were developed collaboratively, and reflective interview evidence was collected from consenting participants. The dyad, therefore, represented the natural unit of the design process and analysis in this study. Treating each pair as a separate case made it possible to examine collaborative processes of planning, negotiation, problem-solving, and reflection as they unfolded during the design process. This approach is consistent with research showing that metacognitive regulation can be socially shared and co-constructed during collaborative activity (Iiskala et al., 2011; Järvelä et al., 2016). Cross-case analysis was then used to identify recurring patterns and differences in metacognitive regulation and the conceptualisation of AI across the six cases.

3.2. *Research Context*

The research is based on Technology Integration in Teaching and Learning, an Arabic-medium teacher education course at a higher education institution in the UAE. In this course, students are expected to create a technology-enhanced learning product and a

Teacher Guide that justify pedagogical integration, adherence to learning objectives, and evaluation methods.

This setting is especially relevant to the study, as it places participants at the confluence of theory, design, and reflective practice and explicitly obligates them to consider design rationales—fostering metacognitive externalisation.

3.3. Participants

Twelve graduate students, paired into six teams, took part in the study during the 2025–2026 academic year. Each pair was enrolled in a technology integration class in a teacher education programme in the UAE. The sample comprised six pairs of students (12 participants), which mirrors the collaborative nature of the course design task. The participants were drawn from one class cohort using convenience sampling. Students in the cohort who completed the design assignment and provided informed consent were eligible to participate in the study.

The sample size was considered appropriate for the exploratory multiple-case study because the main purpose was to examine the design process in depth rather than to identify statistically generalisable patterns. The six dyads provided six cases for comparison, and each case was analysed using three different data sources: the educational application, Teacher Guide, and reflective interview. While the six dyads constituted the primary cases for comparison, the interview evidence consisted of nine semi-structured reflective interviews with consenting participants from across these teams. At least one consenting participant from each of the six design teams took part in the interviews, ensuring that interview data were available for every case. This made it possible to compare evidence within each case and to identify recurring and contrasting patterns across cases.

Participants were considered appropriate for the study because they:

- Had prior exposure to pedagogy, assessment, and technology integration;
- Were engaged in a cognitively demanding design task;
- Were prospective teacher-leaders expected to engage in reflective technology use.

The pair-based format was deliberate, as teamwork fosters dialogue, co-reasoning, and co-articulation of decision-making, and makes metacognitive activities more observable. The use of dyads also reflected the nature of the course task, as the educational applications and Teacher Guides were developed collaboratively. The study, therefore, focused on shared regulatory processes within each pair rather than attempting to separate individual contributions to jointly produced work (Iiskala et al., 2011; Järvelä et al., 2016).

3.4. Participant Selection and Recruitment

A convenience sample was used, with participants recruited from a single cohort enrolled in the target course. The sampling approach, therefore, reflected access to a naturally occurring group whose members shared the same collaborative design experience. This approach was appropriate for the exploratory case study because the aim was to examine the design process in depth rather than to achieve statistical generalisation.

Recruitment occurred through:

- An in-class announcement;
- A follow-up invitation from the Research Office.

Participation was voluntary, and students were informed that:

- Participation would not affect their course grades;
- They could withdraw at any time.

Only coursework and interviews from students who had consented were used.

3.5. Data Collection Instruments

To capture metacognitive regulation in the process of AI-based educational design, multiple complementary data sources were used in the study. The aim was to identify what kinds of tacit process decision-making and explicit reasoning were involved. The focus of data collection was on artefacts created in the context of the course and on reflective dialogue prompted by interviews.

The first data source was the educational apps created by the student dyads. These ICT products offered a tangible demonstration of how participants, in practice, realised pedagogical goals, content, and AI within class contexts. However, the applications provided clear evidence of procedural decision-making and information management, as well as some indications of conditional knowledge through the appropriateness of tool selection and feature design.

The applications were developed using MIT App Inventor (Massachusetts Institute of Technology, Cambridge, MA, USA) as part of a university course in which students designed technology-supported learning experiences rather than commercial products. Although the projects differed in their content, target learners, and pedagogical approach, each one was intended to respond to a particular learning need and support specific learning outcomes. Across the six cases, the applications used different combinations of structured content, multimedia, interactive activities, quizzes, feedback, and learning pathways. AI also played different roles across the projects. Some participants reported using it (e.g., Chat GPT) during the design process to generate ideas, organise or simplify content, and develop learning activities. In other cases, AI-supported features were built into the applications themselves to provide guidance, feedback, adaptation, or content generation. The analysis, therefore, looked beyond the presence of technology or AI and considered how content and activities were organised, how technology choices supported pedagogical goals, and whether the designs created opportunities for feedback, monitoring, revision, and evaluation.

The second type of data came from the Teacher Guides that accompanied each educational application and formed part of the course assignment. Students used Microsoft Word or PowerPoint to create their guides. In the guides, students were required to describe the context and learning problem addressed by the application, identify the target learners and intended learning outcomes, explain the pedagogical rationale, describe the role of digital and AI technologies, and outline how the technology-supported learning activity would be implemented. They were also asked to consider inclusion and learner differences, assessment strategies, and the evaluation of the activity. The guides, therefore, provided evidence of participants' planning, design reasoning, and evaluative thinking. Monitoring and debugging, however, were less visible in the guides and emerged more clearly in the reflective interviews.

Third, semi-structured reflective interviews were conducted with consenting participants after completion of the design task. The interviews conducted and recorded using Microsoft Teams, lasted approximately 30–45 min, and were intended to explore aspects of the design process that were not clearly visible in the educational applications or Teacher Guides. The interview guide covered planning, information management, monitoring, debugging, and evaluation, based on the regulation-of-cognition dimensions of the MAI framework, and also explored participants' experiences using AI during the design process. Participants were asked to explain how they planned the project, organised content and multimedia resources, checked whether the design remained aligned with their goals, dealt with problems that arose, and evaluated the final product.

Sample questions included: "How did you plan the project at the beginning?", "During the development of the application, how did you check whether the design remained

aligned with your objectives?”, “Can you describe a challenge you faced during the design process and explain how you addressed it?”, and “What would you do differently if you could redo the project?” These questions were intended to encourage participants to reflect on planning, monitoring, debugging, and evaluation in relation to their actual design experience.

The teacher guide template and the interview protocols for the study were developed by the research team, and prepared and delivered in Arabic, the medium of instruction. To ensure conceptual equivalence and the authenticity of participants’ thoughts, the interviews were conducted in Arabic. The interview excerpts presented in the manuscript were transcribed and translated from Arabic into English manually by the bilingual research team, with care taken to preserve their intended meaning.

3.6. Data Analysis

Data analysis was conducted in two stages using a deductive–inductive qualitative approach. First, the regulation-of-cognition dimensions of the Metacognitive Awareness Inventory (MAI; [Schraw & Dennison, 1994](#)) were adapted into a qualitative coding scheme. In this study, the MAI was not administered to participants and was not used to score their metacognitive awareness. Five regulation categories—planning, information management, monitoring, debugging, and evaluation—were defined as codes for analysing evidence from the educational apps, Teacher Guides, and interviews. This approach was chosen because the study focused on how regulatory processes became visible in design activities and reflection rather than on measuring participants’ overall level of metacognitive awareness. Before coding independently, the two researchers went through the coding scheme together and discussed what each category meant and how it should be applied. They then coded a small sample of the data together to clarify how the scheme would be used across the three data sources. After that, each researcher independently coded 20% of the dataset using the initial codebook. Agreement for this subset was 85%. This figure was used as a practical check on coding consistency rather than as a formal measure of inter-rater reliability. The researchers then compared their coding, discussed and resolved any disagreements or ambiguities, refined the code definitions, and agreed on the final codebook before coding the remaining data.

The coding results from the three data sources were first organised by case and by source using a cross-source matrix to support systematic comparison ([Miles et al., 2020](#)). For each student pair, evidence from the educational application, Teacher Guide, and interview was then compared for each regulatory process. The comparison focused on three types of relationships across the data sources: convergence, where the sources supported a similar interpretation; complementarity, where one source added information or insight that was not visible in the others; and divergence, where the sources pointed to different interpretations. When the sources did not align, the researchers returned to the original data and discussed the case until a shared interpretation was reached. The resulting case-level interpretations were then compared across the six cases to identify recurring patterns and meaningful differences. Table 1 presents the operational definitions used for each regulation category and the types of evidence examined across the three data sources.

The second stage of data analysis employed an inductive thematic approach, informed by [Braun and Clarke \(2006\)](#), to explore how participants understood and experienced AI during the design process, particularly its affordances, limitations, and tensions between automation and pedagogical goals. First, the researchers became familiar with the interview data through repeated reading and identified initial codes based on recurring ideas and experiences described by the participants. Codes that were related to one another were then grouped into candidate themes, which were reviewed and refined through comparisons

within and across the six cases. Themes were retained when they appeared across more than one case and represented a distinct pattern in how participants understood or used AI during the design process. This process resulted in three main themes: AI as a Planning Scaffold, AI as Reactive Refinement, and AI as a Conditional Tool. Finally, these themes were compared with evidence from the educational apps and Teacher Guides.

Table 1. Joint Display of Triangulated Evidence (RQ1).

Regulation Category	Operational Definition in This Study	Evidence in Educational Apps	Evidence in Teacher Guides	Evidence in Interviews
Planning	Setting goals, selecting strategies, and anticipating how the design task will be approached before or during development.	Sequencing of activities, alignment between learning objectives and app structure, and planned use of tools or AI features.	Statements of learning objectives, design intentions, pedagogical strategies, and implementation plans.	Descriptions of initial planning, goal setting, task division, tool selection, and changes to the original plan.
Information management	Organising, structuring, and presenting information in ways that support clarity, usability, and learning.	Organisation of menus, content flow, multimedia integration, navigation, and information sequencing.	Explanations of how content, resources, multimedia, and learning activities were organised.	Accounts of decisions about organising content, simplifying information, sequencing ideas, and combining text, visuals, audio, or other media.
Monitoring	Checking whether the design process and emerging product remain aligned with intended goals, learner needs, and pedagogical purposes.	Feedback mechanisms, progress checks, conditional pathways, revision opportunities, or analytics features.	References to reviewing progress, checking alignment, or monitoring learner response and app performance.	Descriptions of noticing problems, recognising mismatches, checking progress, or questioning whether a design decision was working as intended.
Debugging	Identifying and correcting problems, errors, misalignments, or ineffective strategies during the design process.	Retry mechanisms, corrected technical features, revised navigation, modified activities, or changes made after testing.	Descriptions of technical or pedagogical problems and the changes made to address them.	Accounts of difficulties encountered, strategies used to solve them, changes in approach, and revisions to AI-generated or technology-based outputs.
Evaluation	Judging the quality, effectiveness, or outcomes of the design process and final product using criteria or evidence.	Assessment components, feedback structures, performance measures, reflection prompts, or evaluation features.	Evaluation plans, assessment criteria, success indicators, or discussion of expected outcomes.	Reflections on what worked, what did not work, why particular decisions were effective, and what participants would change in future.

Once deductive coding was completed, the researchers developed analytical intensity ratings to support within-case and cross-case comparison. These ratings were not MAI scores and did not indicate participants' level of metacognitive awareness. Instead, they summarised the strength and frequency of evidence for each regulatory process across the three data sources. A five-point scale was used: a rating of 1 indicated isolated or

very limited evidence; 2, limited but clear evidence; 3, moderate and recurring evidence; 4, strong evidence supported by more than one data source; and 5, very strong evidence across the data sources. The ratings were agreed upon through discussion among the researchers after all coded evidence had been reviewed. For example, a lower planning rating was assigned when there was only limited evidence of goal setting, strategy selection, or sequencing, such as a brief mention of objectives or a single design decision. A higher rating was assigned when planning was visible repeatedly across more than one data source, for example, through a clearly structured application, an explicit planning rationale in the Teacher Guide, and further explanation in the interview. Similarly, lower monitoring ratings reflected little evidence that participants checked progress or alignment during development. In contrast, higher ratings were used when monitoring was evident through repeated checking, use of feedback, revision, or reflection.

3.7. Ethical Considerations

Ethical permission was granted in accordance with the institution's IRB norms. Consent was obtained from all participants prior to data collection. To reduce the risk of coercion, the interviews were conducted by a researcher who was not involved in assessing the participants' coursework. Although the research used artefacts embedded in the course, only the data from participants who consented were analysed. All identifiable information was removed during transcription and analysis, and the data were kept in a secure place.

4. Results

The findings are organised according to the research questions and presented through joint displays that combine the researcher-generated analytical intensity ratings with analyses of the artefacts, Teacher Guides, and interview themes. These displays are used to examine convergence, complementarity, divergence, and explanatory patterns across the three data sources. The six design teams are treated as the main cases for cross-case comparison, while interview excerpts are attributed to individual consenting participants using anonymised codes (P1–P9).

4.1. RQ1: What Metacognitive Regulation Processes Were Observable Across Artefacts, Teacher Guides, and Interviews?

The triangulated analysis revealed a clear pattern across all six cases, with analytical intensity ratings and source-specific evidence showing different levels of visibility across the regulatory processes (Table 2). Evaluation followed a different pattern: it was more evident in the applications and Teacher Guides but less clearly articulated in the interviews. Information management was also visible across the cases, particularly in the way content was organised and sequenced.

Planning was evident in the alignment of learning objectives, pedagogical sequencing, tool selection, and the rationale for integration within the artefacts. Participants demonstrated explicit instructional intent and foresight in developing learning pathways. The convergence between artefact architecture and interview discourse indicates that planning was both structurally embedded and consciously articulated. One participant explained how the pair approached the task from the outset: "I started by defining the objectives, which were to transform reading into an interactive experience that develops students' understanding and critical thinking. Then I designed the structure of the application into different pages and sections." (P1, Interview).

Information management was evident across all six cases, particularly in the organisation and sequencing of content and the integration of different media within the applications. This dimension was more visible in the structure of the applications and in participants' explanations of how they organised learning content than in explicit state-

ments in the Teacher Guides. However, it was less useful for distinguishing between higher- and lower-regulation cases because evidence of content organisation and sequencing was common across most of the applications.

Table 2. Joint Display of Triangulated Evidence and Analytical Intensity Ratings (RQ1).

Regulation Process	Analytical Intensity Rating	Interviews	Teacher Guides	App Artefacts	Integrated Pattern
Planning	4 in all cases	High across accounts	Explicit objectives, design intentions, and implementation plans	Strong structural alignment	Strong convergence
Monitoring	Mean = 2.83	Mostly moderate; one high case	Limited and rarely explicit	Weak unless embedded in the design	Conditional strength
Debugging	Mean = 2.83	Mainly described as procedural correction	Rarely documented explicitly	Retry mechanisms and technical corrections	Functional, but not fully reflective
Evaluation	Mean = 3.67	Least elaborated; no high cases	Evaluation plans and criteria were generally present	Assessment and feedback structures were visible	Structural–reflective divergence
Information Management	Mean = 4.00	Clearly described across accounts of organising content, simplifying information, and sequencing ideas	Content organisation, learning activities, and multimedia use were generally described	Strong evidence in menu structure, navigation, content flow, sequencing, and multimedia integration	Broad visibility across cases, but limited differentiation

Note: Analytical intensity ratings were researcher-generated summaries based on evidence across the three data sources and were not MAI scores. Ratings ranged from 1 (isolated or very limited evidence) to 5 (very strong evidence across the data sources).

However, the execution of monitoring and debugging varied among the cases. This was evident in one pair’s description of revising the design after identifying a problem: “During testing, I noticed that some links within the app did not open or took a long time to open. I discovered that this was due to the large size of the videos, so I compressed the files and replaced the links with more stable sources.” (P8, Interview). Although some artefacts included feedback loops, revision pathways, or conditional logic, these features were unevenly distributed across cases and were more evident where the design process offered explicit opportunities for checking and revision. In higher-regulation cases, debugging was a process of refinement and adaptation, whereas in lower-regulation cases, iteration remained procedural rather than reflective.

Evaluation was more complex. Many artefacts contained assessment or feedback structures, but the interviews often showed little articulation of the evaluative reasoning behind them. In several instances, the respondents described assessment as a product demand and not as a reflective activity with criteria, validation or decision-making based on evidence. In one lower-regulation account, evaluation was described mainly as a general judgement of product success rather than as a systematic process based on explicit criteria: “I think the simplicity of the design and the clarity of the content were the most successful elements. The students interacted with the app easily because the steps were clear and

not complicated.” (P9, Interview). Evaluative structures can be procedurally implemented without corresponding metacognitive reflection.

Overall, the triangulated findings suggest that planning was the most consistently visible form of regulation across the six cases, while monitoring, debugging, and evaluation depended more on opportunities for checking, revision, and reflection. The data also showed that different regulatory processes became visible through different sources. Planning and information management were more evident in the applications and Teacher Guides, whereas monitoring and debugging became clearer in the interviews.

4.2. RQ2: In What Ways Did the TIP Framework Make Students’ Metacognitive Decision-Making Visible or Invisible in Written Design Justifications?

Analysis of the artefacts and design documentation showed that the TIP framework made planning decisions more visible, while monitoring, debugging, and iterative refinement were less consistently represented (Table 3).

Table 3. Joint Display of Triangulated Evidence for TIP Visibility (Teacher Guides and Interviews).

TIP Element	Evidence in Teacher Guides	Evidence in Interviews	Integrated Interpretation
Problem identification	Learner needs and goals clearly articulated	Participants start with classroom problems	TIP strongly externalises problem framing
Tool choice & learning pathway	Platforms and activity sequences justified	Tools selected based on affordances	TIP makes tool–pathway reasoning visible
Pedagogical/theoretical justification	Present, but variable depth	Pedagogical reasoning often implicit	TIP prompts explanation, not depth
Equity & learner differences	Mentioned in some guides	Occasionally discussed	Equity weakly surfaced
Monitoring & debugging	Rarely described	Changes made “as needed”	Regulation occurs but invisible
Data-driven adaptation rules	Largely absent	Rare evidence—revision links	TIP does not scaffold adaptation logic

The technology and pedagogy dimensions were consistently reflected in artefacts through tool integration, activity sequencing, and alignment with curriculum goals, indicating structured and intentional design principles.

Iterative practice differed across the cases. In stronger-regulation cases, the design process involved revision, feedback, and gradual refinement, with closer alignment between technological and pedagogical choices. In weaker-regulation cases, development followed a more linear path, with less evidence of systematic debugging or further revision after the initial implementation.

Participants also described how technology integration planning frameworks helped structure their design thinking, often using TPACK/TIP-related language to explain alignment among content, pedagogy, and technology. “Every decision was verified: Does it serve the content? Does it support learning? Does it employ technology in a meaningful way?” (P7, Interview).

The findings suggest that the TIP model helped make the planning process more visible, particularly regarding problem identification, tool selection, learning pathways, and pedagogical reasoning. Monitoring, debugging, and evaluation, however, were less clearly represented in the written documentation and became more visible in participants’ interview accounts of the problems they encountered and the changes they made during the design process.

4.3. RQ3: How Did Students Conceptualise the Role of AI in Supporting or Constraining Metacognitive Regulation for Themselves and Their Learners?

Across cases, AI appeared in two distinct but interconnected roles. First, participants used AI as a design tool during the development process, for example, to generate ideas, organise content, simplify material, or refine aspects of the design. Second, AI was incorporated into some applications as a feature intended to support learners through feedback, guidance, adaptation, or content generation. These two roles placed different demands on participants' metacognitive regulation. Using AI as a design tool required participants to evaluate, refine, or reject AI-generated outputs, while embedding AI as a feature required them to consider how learners would interact with the AI-supported function. Table 4 shows how the three main conceptualisations of AI identified in the data appeared across its two roles: as a design tool and as a feature embedded within the applications.

Table 4. Integrated Thematic Comparison of AI Conceptualisations.

AI Role	AI Conceptualisation	Evidence in Interviews	Evidence in App Artefacts	Integrated Interpretation
AI as a design tool	AI as Planning Scaffold	Idea generation, organisation, acceleration	Structured content and coherent sequencing	AI-supported planning and early-stage design activity
AI as a design tool	AI as Reactive Refinement	Content simplified or revised when misaligned	Limited but visible adaptation of content or features	AI-supported adjustment mainly after a mismatch was noticed
AI as a design tool and embedded feature	AI as Conditional Tool	Awareness of benefits, limits, and the need for human judgement	Human mediation or oversight is implied in some designs	Emerging evidence of conditional AI use and human judgement

AI as a Planning Scaffold was evident when participants described using AI to generate, organise, or develop initial ideas. One participant explained: "Artificial intelligence helped me organize ideas, divide pages, and clarify the final image before the actual application." (P2, Interview). AI as Reactive Refinement appeared when participants used AI to simplify, revise, or adapt an output that did not initially fit the learning context: "It helped generate new ideas and formulate content, but sometimes it produced longer texts than required, so summarization was done manually." (P7, Interview). AI as a Conditional Tool was evident when participants recognised that the value of AI depended on the task, context, and quality of its output: "Technology added great value when I used it to simplify abstract religious concepts. However, it was not effective in lessons that required in-depth discussion, such as 'intention and sincerity,' because direct dialogue and discussion were more effective." (P8, Interview).

When AI was used as a design tool, participants generally described it as a way to make planning more efficient through brainstorming, content creation, organisation, and prototyping. However, much less attention was given to the systematic evaluation of AI-generated outputs. In some cases, participants incorporated AI-generated content into the design process, but there was limited evidence that it had been carefully checked for accuracy, suitability, potential bias, or alignment with the intended educational purpose before use. There was also little evidence of systematic prompt refinement or other procedures for checking whether AI-generated material was appropriate for the intended learning task.

When AI was embedded as a feature within the application, the regulatory demands were different. Beyond ensuring that the AI-supported function worked as intended, par-

ticipants also had to consider how it might shape learners' interaction with the content. In some cases, there was evidence that participants recognised the need for human guidance or oversight. In others, the AI-supported feature was incorporated with limited consideration of how learners would question, interpret, or evaluate its outputs.

When participants questioned, checked, or refined AI-generated outputs, there was stronger evidence of regulatory activity than when AI was used mainly to save time or increase efficiency. Across the cases examined, AI appeared to support brainstorming and content development more readily than sustained monitoring, verification, and evaluation.

4.4. RQ4: What Design Features and Regulatory Conditions Were Associated with Stronger Metacognitive Profiles, and How Do These Inform Teacher-Education Coursework?

Cross-case analysis showed that stronger regulatory profiles were more closely associated with embedded monitoring structures and evidence-based evaluation, while procedural debugging remained relatively limited and AI pedagogical reasoning was less developed (Table 5). Participants also frequently linked stronger regulatory engagement with the need to validate AI-generated outputs.

Table 5. Integrated Regulatory Differentiation Profiles Linked to RQ1 Analytical Intensity Ratings.

Regulatory Condition	Case-Level Analytical Indicator (RQ1)	Cross-Source Qualitative Evidence (RQ4)	Regulatory Interpretation
Evidence-based evaluation	EV $\geq$ 4 (G2–G6)	Pre/post measures; performance indicators; explicit evaluation criteria	Stronger evidence of evaluative regulation
Embedded monitoring structures	MN = 4 (G4, G6)	Dashboards; analytics; feedback loops; revision cycles	More balanced regulatory profile
Procedural debugging	DB $\leq$ 3	Retry mechanisms; technical fixes; problem correction	Regulation focused mainly on immediate problem resolution
AI pedagogical reasoning	Not represented by a separate analytical rating in RQ1	Limited explicit reasoning about how AI should support learning	Conceptually underdeveloped area

Note. MN = monitoring; DB = debugging; EV = evaluation. G refers to the six design-team cases, while interview excerpts are reported using participant codes (P1–P9).

In the stronger regulation case, planning, monitoring, revision, and evaluation were particularly visible. The participants described checking whether the application remained aligned with the intended learning outcomes and making changes when needed. As one participant explained, “I returned to the learning outcomes and reviewed each part of the application to ensure that it served the objectives. I conducted small tests with students to observe their interaction and adjust any errors.” (P7, Interview). This interpretation was also consistent with the evidence of small-scale testing, feedback-based adjustment, and review of each design component against the intended learning outcomes, suggesting that monitoring and evaluation were considered at more than one stage of the design process.

In the contrasting lower regulation case, planning was clearly evident, but there was much less evidence of monitoring and evaluation once development was underway. The participants described the process mainly in terms of implementing what had been planned, with limited reference to systematic checking, revision, or evaluation. As one participant noted, “I used the application in very simple steps, such as inserting only short texts and illustrative images. I enlisted the help of a more experienced colleague to help me design the interactive activity and adjust the page layout.” (P9, Interview). This interpretation was

consistent with the limited presence of revision procedures and explicit evaluation criteria in both the application and the Teacher Guide.

In the stronger-regulation cases, the monitoring structures included feedback loops, conditional reasoning, and revision processes that supported iterative improvement. This was also reflected in the interviews, where participants described checking alignment, reviewing outputs, and making adjustments during development.

Cases with stronger regulatory profiles were also characterised by more explicit evaluative reasoning. Participants in these cases referred to criteria, needs, alignment, and evidence-based justification. In cases with lower regulation, planning strength was apparent, but monitoring and evaluation were weak or implicit.

In this study, the use of AI on its own was not associated with stronger metacognitive regulation. Greater regulatory depth was evident in cases where monitoring and evaluation were built into the design and also reflected in participants' explanations. Planning appeared to be the easiest metacognitive process to support, while sustained regulation was more evident in cases where the design process included opportunities for checking, revision, validation, and reflection.

5. Discussion

The study examined how metacognitive regulation became visible as graduate teacher education students designed AI-supported educational applications. Across the six cases, planning was the most consistently evident regulatory process. Monitoring, debugging, and evaluation were less visible and became clearer when the design process included opportunities for checking, revision, and reflection. The role of AI also varied across cases and depended on how participants engaged with AI-generated outputs, particularly whether they questioned, checked, revised, and justified them.

5.1. Metacognitive Regulation in AI-Supported Design

The findings provide evidence of metacognitive regulation across all three data sources—the educational applications, Teacher Guides, and reflective interviews—although the visibility of these processes varied across sources. Planning was the most clearly visible dimension. Across the six cases, participants linked learning objectives, pedagogical sequencing, and technology choices in ways that showed deliberate planning during the design process. Information management was also evident, particularly in the way content was organised and structured within the applications. However, this form of regulation was less useful for distinguishing between cases because it appeared across almost all of the designs.

This pattern is consistent with foundational theories of metacognition and self-regulated learning, which describe regulation as a cyclical process involving planning, monitoring, adjustment, and evaluation rather than as a single act of reflection (Flavell, 1979; Pintrich, 2002; Zimmerman, 2002). The uneven visibility of these processes across the six cases also supports the view that metacognitive activity does not necessarily become explicit simply because learners are engaged in a demanding task. As Veenman et al. (2006) argue, regulation often needs to be prompted and supported before it becomes visible in learners' actions and explanations.

A possible reason for the emphasis on planning is the nature of the design task itself. Within the TIP framework, participants were expected to define learning outcomes, select an instructional approach, explain the role of technology, and consider implementation before developing the final product. These requirements may have encouraged more deliberate planning while also making this process more visible in the data. This interpretation is

consistent with recent research suggesting that generative AI can help students organise steps and plan their approach to course-related tasks (Liu et al., 2025; S. C. Yang, 2026).

In addition, the findings showed that monitoring, debugging, and evaluation were less consistently visible than planning. Monitoring and debugging became clearer in cases where the applications included feedback mechanisms, revision pathways, or opportunities to respond to problems during development. This suggests that these regulatory processes may become more visible when the task itself includes clear opportunities for checking and revision. Similar patterns have been reported in recent reviews, suggesting that AI-supported learning environments are more likely to encourage metacognitive activity when learners are prompted to monitor and reflect rather than expected to self-regulate spontaneously (Tsakeni et al., 2025).

Evaluation showed a somewhat different pattern. Evaluation elements were often present in the applications and Teacher Guides, but participants did not always explain the reasoning behind them during the interviews. This suggests that evaluation can be built into a product without necessarily being accompanied by explicit evaluative reflection. Previous studies also indicate that deeper regulation is more likely when learners actively examine and reflect on AI-generated content (Iqbal et al., 2025; Tomisu et al., 2025). Research comparing different types of AI-generated feedback further suggests that metacognitive prompts can encourage deeper engagement than answer-focused feedback alone (Alsaiani et al., 2026). In the present study, this may help explain why AI use appeared more clearly connected to planning and content development than to sustained monitoring and evaluation in some cases.

5.2. AI as a Cognitive Scaffold and a Potential Cognitive Shortcut

These findings clarify the scaffold–shortcut distinction by showing that the role of AI did not depend simply on whether participants chose to use it. What mattered more was how they responded to AI-generated outputs afterwards. In the stronger-regulation cases, participants examined, revised, or rejected AI-generated material and connected their decisions to pedagogical goals. In the weaker-regulation cases, AI was used mainly to generate ideas or content quickly, with less evidence of further analysis, checking, or revision.

The findings of this study are in line with earlier research suggesting that AI can support learners when it is used as a thinking partner rather than treated as a source of ready-made answers. Previous studies show that AI can help learners explore alternatives, question assumptions, and make their thinking more visible (An & Park, 2026; Hsieh et al., 2026; Iqbal et al., 2025; Alam & Mohanty, 2023). A similar pattern appeared in this study, where deeper engagement was more evident when participants questioned, revised, or adapted AI-generated outputs. This interpretation also aligns with recent work in AI-supported Qur'anic education, which argues that AI is most valuable when positioned as pedagogical support rather than as a replacement for teachers, especially where accuracy, ethical alignment, and human verification remain essential (Almajdoub et al., 2026).

The weaker-regulation cases, however, show the risks of using AI mainly for speed and convenience. Earlier research has warned that easy access to AI-generated responses may reduce opportunities for deeper thinking when learners do not question or evaluate the system's output (Gerlich, 2025; Sitepu et al., 2025). The present study adds to this discussion by showing that the difference between scaffold and shortcut became evident in how participants monitored, checked, revised, and evaluated AI-generated outputs.

Taken together, the findings suggest that the scaffold–shortcut distinction is better understood as a spectrum of regulatory engagement rather than as a fixed characteristic of AI itself. In this study, AI functioned more as a scaffold when participants verified,

compared, revised, and justified its outputs. It functioned more as a shortcut when outputs were accepted mainly for speed or convenience, with limited evidence of monitoring or evaluation. This interpretation extends the study's original framing by identifying the regulatory conditions associated with each form of AI use. For teacher education, this means that AI-supported tasks should require students to demonstrate not only what AI contributed, but also how its outputs were examined, modified, rejected, or justified.

5.3. Accuracy and Human Verification in AI-Supported Learning Design

The findings show that AI can support learning design when students remain actively involved in the process. Many of the applications had clear learning goals, a logical sequence of activities, and purposeful use of technology. In these cases, participants' accounts suggest that AI-supported aspects of idea organisation, activity planning, and the coordination of different elements of the design.

At the same time, the findings suggest that the important issue is not simply whether AI is used, but how its outputs are handled. In the stronger-regulation cases, participants showed more evidence of checking, verifying, and refining AI-generated material. In the weaker-regulation cases, there was less evidence of systematic review. This highlights the continuing importance of human judgement, particularly when AI-generated content becomes part of an educational product. Previous research has raised similar concerns, showing that AI can support idea generation and organisation while still producing outputs that need careful review (Gerlich, 2025; Dergaa et al., 2023).

The findings also indicate that AI-supported design is stronger when participants do more than accept what the system produces. Participants who reviewed, questioned, and revised AI-generated material showed clearer evidence of validation and pedagogical alignment. Earlier studies have also highlighted the importance of pedagogical judgement, trust, and thoughtful integration when AI is introduced into teaching and learning (Hsieh et al., 2026). AI can therefore be useful in the design process, but its outputs still need to be examined rather than treated as complete or automatically reliable sources of knowledge.

Taken together, the findings suggest that the value of AI-supported learning design depends not only on access to AI tools, but also on what learners do with the outputs they receive. Human verification remains important, especially when AI-generated material is incorporated into educational products.

5.4. Implications for Pedagogical Guidance and Teacher AI Competence

The findings also highlight the role of guidance in AI-supported design activities. In this study, stronger regulatory patterns tended to appear when participants showed clearer instructional alignment, explained more explicitly why they were using AI, and paid greater attention to evaluating AI-generated outputs. Teacher AI competence was not measured directly, but these patterns suggest that the design and support of AI-supported tasks may influence how critically learners engage with AI.

From this perspective, teacher AI competence emerges as an implication of the findings rather than as a directly measured outcome. Helping learners use AI critically requires more than knowing how the technology works. Teachers also need to design activities that ask students to question AI-generated outputs, compare alternatives, explain their decisions, and return to their work for further revision. Teacher education programmes, therefore, need to prepare educators not only to use AI tools but also to create learning conditions in which students remain responsible for judging and improving AI-supported work.

Recent research on teacher AI competence supports this interpretation. Teachers' AI competence has been associated with their readiness to integrate AI into teaching and learning (Ruiz & Yambao, 2026). Current frameworks also treat AI readiness as broader

than technical competence alone, bringing together pedagogy, assessment literacy, ethical awareness, and human-centred approaches (Chiu et al., 2025). Other studies emphasise capabilities such as working with data, collaborating with AI systems, and thinking critically about how AI should be used in different contexts (J. Fan et al., 2026). This is consistent with evidence that AI may contribute more strongly to academic achievement than to social-skill development, reinforcing the need for human-centred teacher education approaches that preserve interaction, guidance, and social learning (Almajdoub et al., 2025). At the same time, research suggests that teachers may be able to use AI tools technically without yet feeling prepared to integrate them meaningfully into teaching and learning (Rajapakse, 2026). This points to the need for teacher education programmes to go beyond tool use and help teachers develop the judgement needed to interpret AI-generated outputs critically and consider their ethical implications.

Taken together, the findings of this study and the wider literature suggest that meaningful AI integration depends not only on what learners do with AI-generated outputs, but also on the pedagogical decisions that shape those interactions. Preparing teachers for AI-supported learning, therefore, means helping them design activities that make room for monitoring, evaluation, reflection, and human judgement.

5.5. Implications for Practice and Future Research

The results of this study provide several implications for educational practice and future research on AI-supported learning environments. One implication is that the findings of this study support incorporating AI tools into instructional design to create structured pedagogical designs, rather than treating them as independent technologies. Learning activities need to provide opportunities for students to evaluate, refine, and critically examine AI-generated outputs in order to sustain metacognitive engagement through deeper learning processes (Chiu et al., 2025; Tsakeni et al., 2025). Another implication of the study is that teachers must have the competencies required to support meaningful integration of AI into learning environments. Therefore, teacher preparation programmes must include opportunities to develop AI literacy, ethical awareness, and pedagogical decision-making skills needed to use AI in learning environments (J. Fan et al., 2026; Daher, 2025).

These practices can be incorporated into AI-supported design activities in fairly simple ways. At key points in the task, students can be asked to pause, review their progress, and consider whether the design is still moving towards the intended learning outcomes. Problems and changes can be documented in short error logs or revision journals, in which students explain what went wrong, what they did about it, and why they accepted, changed, or rejected particular AI-generated suggestions. Evaluation can also be made part of the task by asking students to check AI-generated outputs against agreed criteria, subject knowledge, pedagogical goals, or other reliable sources before deciding whether to use them.

Instructional designers can support the same processes by building feedback loops, revision pathways, and opportunities for learners to reconsider their decisions within AI-supported activities. When learners receive AI-generated outputs, they can be prompted to check their accuracy, compare alternatives, decide whether they agree with the response, and explain any changes they make. These design features can help keep human judgement visible and reduce the risk of learners accepting AI-generated outputs without reflection.

The findings also point to several areas for future research. Larger studies across a wider range of teacher education contexts could help determine whether the regulatory patterns identified here also appear in other settings and task designs. Longitudinal research could follow learners through the design process to examine how planning, monitoring,

debugging, and evaluation develop over time rather than relying mainly on retrospective accounts. Another useful direction would be to compare different forms of instructional support, such as monitoring prompts, validation checkpoints, peer review, and reflective design logs, to examine which approaches are most effective in supporting sustained regulatory engagement with AI. Further research could also explore how teachers' AI competence and pedagogical decision-making influence learners' willingness to question, verify, and revise AI-generated material.

6. Limitations

There are several limitations of the study that should be considered when interpreting the findings. First, the study was conducted within a single graduate teacher education programme at one institution and involved a relatively small number of participants. These features reflect the exploratory and qualitative nature of the multiple-case design but limit the transferability of the findings to other educational settings.

The findings should therefore be understood within the context in which the study was carried out. The design work was completed collaboratively in student pairs, likely shaping how participants planned their work, made decisions, used AI, and reflected on the design process. The detailed description of the course context, design task, and procedures used for data collection and analysis is intended to help readers judge whether the findings may be relevant to similar teacher education or educational design settings. Because the dyad was the unit of analysis, the study focused on shared regulatory processes within each pair. As a result, differences in how individual partners planned, monitored, or evaluated the task may not be fully visible in the findings.

Second, the analysis was based on Teacher Guides, application artefacts, and retrospective interviews. These data sources cannot provide a complete account of every cognitive process that occurred during the design process. Future research could therefore incorporate additional sources, such as design logs, think-aloud protocols, or longitudinal observations, to examine more directly how educators regulate their thinking while designing AI-supported learning environments.

Third, the retrospective nature of the interviews may have introduced social desirability or recall bias, as participants may have described their decision-making as more systematic than it was during the actual design process. Similarly, cognitive processes cannot be observed directly through artefacts alone, so interpretations of metacognitive regulation necessarily involve a degree of researcher judgement. Think-aloud protocols or design journals could provide more immediate evidence of regulatory thinking in future studies.

Finally, the MAI framework was adapted for qualitative coding, which goes beyond the way it was originally intended to be used. This should be kept in mind when interpreting the findings. Percentage agreement was also used as a consistency check rather than a chance-corrected reliability coefficient, which should be considered when interpreting the coding procedure.

7. Conclusions

This study examined how graduate students in teacher education regulate their thinking as they design AI-supported educational applications. By analysing design artefacts, Teacher Guides, and reflective interviews, the study explored how metacognitive processes such as planning, monitoring, debugging, and evaluation emerged during technology-enhanced instructional design.

The results indicate that AI can support idea generation, the organisation of learning tasks, and the exploration of alternative options, provided that participants remain actively

engaged with its outputs. Across the six cases examined in this study, AI functioned more as a scaffold when participants questioned, checked, revised, and justified AI-generated outputs. It functioned more as a shortcut when outputs were used mainly for speed or convenience, with less evidence of monitoring or evaluation.

The findings suggest that the scaffold–shortcut distinction is better understood as a pattern of engagement rather than as a fixed feature of AI itself. Across the cases, the role of AI was associated with the regulatory and pedagogical conditions surrounding its use, particularly the extent to which participants questioned, verified, revised, and evaluated AI-generated outputs.

The study makes three contributions to the growing literature on AI-supported teacher education. First, it highlights the uneven visibility of metacognitive regulation across the design process. Planning was consistently evident, while monitoring and debugging became more visible when the task included opportunities for checking, revision, and reflection. Second, it shows the value of combining educational applications, Teacher Guides, and reflective interviews, as different regulatory processes became visible through different sources. Finally, the study clarifies the scaffold–shortcut distinction by showing that, across these six cases, the role of AI was linked to the quality of learners' regulatory engagement, particularly their attention to monitoring, validation, refinement, and evaluation.

Author Contributions: Conceptualisation, Z.A.; methodology, Z.A., R.H. and Y.A.; validation, P.F., R.H. and Y.A.; formal analysis, Z.A.; investigation, Z.A. and Y.A.; resources, P.F.; data curation, Z.A., R.H. and Y.A.; visualisation, Z.A.; writing—original draft preparation, Z.A.; writing—review and editing, P.F., R.H. and Y.A.; supervision, P.F.; project administration, Z.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Emirates College for Advanced Education (Project No. RP-612-2025, Version 2; approved on 19 October 2025).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy and ethical restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K–12 settings. *AI and Ethics*, 2, 431–440. [\[CrossRef\]](#)
- Alam, A., & Mohanty, A. (2023). Educational technology: Exploring the convergence of technology and pedagogy through mobility, interactivity, AI, and learning tools. *Cogent Engineering*, 10(2), 2283282. [\[CrossRef\]](#)
- Almajdoub, Z., Fidalgo, P., Alramamneh, Y., & Al Armoti, A. (2025). The impact of artificial intelligence on quality assurance in education: Academic achievement and social skills dimensions. *Educational Process: International Journal*, 17, e2025298. [\[CrossRef\]](#)
- Almajdoub, Z., Fidalgo, P., Alramamneh, Y., AlHakmani, R., & Kasasbeh, N. (2026). Artificial intelligence in Qur'anic education: Pedagogical promise, practical challenges, and ethical concerns. *Frontiers in Education*, 11, 1876380. [\[CrossRef\]](#)
- Alsaiani, O., Baghaei, N., Lodge, J. M., Noroozi, O., Gašević, D., Boden, M., & Khosravi, H. (2026). Directive, metacognitive, or a blend of both? A comparison of AI-generated feedback types on student engagement, confidence, and outcomes. *Computers and Education: Artificial Intelligence*, 10, 100553. [\[CrossRef\]](#)
- An, H., & Park, M. (2026). AI tools and fashion design education: Instructor perspectives on student challenges and design process tool support. *Fashion and Textiles*, 13(1), 3. [\[CrossRef\]](#)
- Biagini, L. (2025). Towards an AI-Literate Future: A systematic literature review exploring education, ethics, and applications. *International Journal of Artificial Intelligence in Education*, 35, 2616–2666. [\[CrossRef\]](#)
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. [\[CrossRef\]](#)

- Chai, C. S., & Koh, J. H. L. (2017). Changing teachers' TPACK and design beliefs through the scaffolded TPACK lesson design model (STLDM). *Learning: Research and Practice*, 3(2), 114–129. [\[CrossRef\]](#)
- Chiu, T. K. F., Ahmad, Z., & Çoban, M. (2025). Development and validation of the teacher artificial intelligence competence self-efficacy scale. *Education and Information Technologies*, 30, 6667–6685. [\[CrossRef\]](#)
- Daher, R. (2025). Integrating AI literacy into teacher education: A critical perspective paper. *Discover Artificial Intelligence*, 5(1), 217. [\[CrossRef\]](#)
- Dergaa, I., Chamari, K., Zmijewski, P., & Ben Saad, H. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40(2), 615–622. [\[CrossRef\]](#) [\[PubMed\]](#)
- Dignath, C., & Veenman, M. V. J. (2021). The role of direct strategy instruction and indirect activation of self-regulated learning—Evidence from classroom observation studies. *Educational Psychology Review*, 33, 489–533. [\[CrossRef\]](#)
- Fan, J., Wang, H., & Gu, X. (2026). Development and validation of a framework and scale for primary and secondary school teachers' data-artificial intelligent competence. *Frontiers in Psychology*, 16, 1756893. [\[CrossRef\]](#) [\[PubMed\]](#)
- Fan, Y., Tang, L., Le, H., & Zhang, Y. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *arXiv*. [\[CrossRef\]](#)
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10), 906–911. [\[CrossRef\]](#)
- Geitz, G., Joosten-ten Brinke, D., & Kirschner, P. A. (2019). Design-based education, sustainable teaching, and learning. *Cogent Education*, 6(1), 1647919. [\[CrossRef\]](#)
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. [\[CrossRef\]](#)
- Ho, W. W. Y., & Lau, Y. H. Y. (2025). Role of reflective practice and metacognitive awareness in the relationship between experiential learning and positive mirror effects: A serial mediation model. *Teaching and Teacher Education*, 157, 104947. [\[CrossRef\]](#)
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promise and implications for teaching and learning*. Center for Curriculum Redesign.
- Hsieh, C.-C., Bali, S., & Li, H.-C. (2026). Untangling pathways from digital instructional leadership to AI pedagogy: A diffusion of innovation theory perspective. *BMC Psychology*, 14, 63. [\[CrossRef\]](#) [\[PubMed\]](#)
- Iiskala, T., Vauras, M., Lehtinen, E., & Salonen, P. (2011). Socially shared metacognition of dyads of pupils in collaborative mathematical problem-solving processes. *Learning and Instruction*, 21(3), 379–393. [\[CrossRef\]](#)
- Iqbal, J., Hashmi, Z. F., Asghar, M. Z., & Abid, M. N. (2025). Generative AI tool use enhances academic achievement in sustainable education through shared metacognition and cognitive offloading among pre-service teachers. *Scientific Reports*, 15(1), 16610. [\[CrossRef\]](#) [\[PubMed\]](#)
- Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual- and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning*, 11, 263–280. [\[CrossRef\]](#)
- Koehler, M. J., & Mishra, P. (2012). What is technological pedagogical content knowledge (TPACK)? *Journal of Education*, 193(3), 13–19. [\[CrossRef\]](#)
- Liu, X., Xiao, Y., & Li, D. (2025). Assessing strategic use of artificial intelligence in self-regulated learning: Instrument development and evidence from Chinese university students. *International Journal of Educational Technology in Higher Education*, 22, 69. [\[CrossRef\]](#)
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE Publications.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Shen, M. Q. (2021). Conceptualising AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. [\[CrossRef\]](#)
- Novoa-Echaurren, Á. (2025). Reflective practice and digital technology use in a Chilean university context. *Education Sciences*, 15(6), 643. [\[CrossRef\]](#)
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422. [\[CrossRef\]](#) [\[PubMed\]](#)
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, 41(4), 219–225. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rajapakse, C. P. (2026). A TPACK-assisted descriptive analysis of the Sri Lankan ICT teachers' readiness to teach AI in schools. *International Journal of Information and Education Technology*, 16(2), 557–568. [\[CrossRef\]](#)
- Roblyer, M. D., & Hughes, J. E. (2019). *Integrating educational technology into teaching* (8th ed.). Pearson.
- Ruiz, C. C., & Yambao, F. M. (2026). AI literacy and professional development engagement as predictors of mathematics teachers' readiness for AI-enhanced instruction. *International Journal of Learning, Teaching and Educational Research*, 25(2), 655–675. [\[CrossRef\]](#)
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460–475. [\[CrossRef\]](#)

- Sitepu, M. S., Prasajo, L. D., Hermanto, H., Salido, A., Nurhakim, L., Setyorini, E., Disnawati, H., & Wiratsongko, B. (2025). Mapping and exploring strategies to enhance critical thinking in the artificial intelligence era: A bibliometric and systematic review. *European Journal of Educational Research*, 15(1), 305–322. [\[CrossRef\]](#)
- Stanton, J. D., Neider, X. N., Gallegos, I. J., & Clark, N. C. (2021). Fostering metacognition to support student learning and performance. *CBE—Life Sciences Education*, 20(2), fe3. [\[CrossRef\]](#) [\[PubMed\]](#)
- Tomisu, H., Ueda, J., & Yamanaka, T. (2025). The cognitive mirror: A framework for AI-powered metacognition and self-regulated learning. *Frontiers in Education*, 10, 1697554. [\[CrossRef\]](#)
- Tondeur, J., Scherer, R., Siddiq, F., & Baran, E. (2020). Enhancing pre-service teachers' technological pedagogical content knowledge (TPACK): A mixed-method study. *Educational Technology Research and Development*, 68(1), 319–343. [\[CrossRef\]](#)
- Tsakeni, M., Nwafor, S. C., Mosia, M., & Egara, F. O. (2025). Mapping the scaffolding of metacognition and learning by AI tools in STEM classrooms: A bibliometric–systematic review approach (2005–2025). *Journal of Intelligence*, 13(11), 148. [\[CrossRef\]](#) [\[PubMed\]](#)
- Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, 1(1), 3–14. [\[CrossRef\]](#)
- Viberg, O., Grönlund, Å., & Andersson, A. (2023). What explains teachers' trust of AI in education across six countries? *arXiv*. [\[CrossRef\]](#)
- Wang, M., Chen, Z., Xu, Y., Maheshi, B., & Gašević, D. (2025). Intelligent teaching analytics for collaborative reflection: Investigating pre-service teachers' experiences and shared regulation processes. *International Journal of Educational Technology in Higher Education*, 22, 45. [\[CrossRef\]](#)
- Wass, R., Rogers, T., Brown, K., Smith-Han, K., Berg, D., & Gallagher, S. (2023). Pedagogical training for developing students' metacognition: A comparative study of educators' practice. *International Journal for Academic Development*, 1–14. [\[CrossRef\]](#)
- Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support. *British Journal of Educational Technology*, 56(5), 1842–1863. [\[CrossRef\]](#)
- Yang, S. C. (2026). Integrating technology acceptance, self-determination, and self-regulation: A structural model of generative AI-supported learning and competence. *Computers in Human Behavior*, 179, 108933. [\[CrossRef\]](#)
- Yang, Y., & Xia, N. (2023). Enhancing students' metacognition via AI-driven educational support systems. *International Journal of Emerging Technologies in Learning (iJET)*, 18(24), 133–148. [\[CrossRef\]](#)
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, 16, 39. [\[CrossRef\]](#)
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70. [\[CrossRef\]](#) [\[PubMed\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.