



Paper-first knowledge engineering as a pedagogical framework for cognitive agency in AI-assisted learning

Gary Wayne Collins ^{1*}

 0000-0001-8053-8529

¹ Department of Applied Languages, Tshwane University of Technology, Pretoria, SOUTH AFRICA

* Corresponding author: collinsgw@tut.ac.za

Citation: Collins, G. W. (2026). Paper-first knowledge engineering as a pedagogical framework for cognitive agency in AI-assisted learning. *Contemporary Educational Technology*, 18(4), Article ep700. <https://doi.org/10.30935/cedtech/19638>

ARTICLE INFO

Received: 8 May 2026

Accepted: 11 Sep 2026

ABSTRACT

This study investigates how a structured “paper-first” approach can support cognitive engagement during artificial intelligence (AI)-assisted academic work in higher education. Growing reliance on generative AI has increased concerns about surface-level learning, trial-and-error prompting, and reduced student ownership of academic writing. Using a two-year design-based research approach involving nine iterative cycles, the study developed and refined a framework that required students to construct manual logic diagrams before engaging with AI tools and expert system environments. Participants included a six-member interdisciplinary design team and 150 communication for academic purposes students from the faculty of information and communications technology at a South African university. Data were collected through iterative observations and focus group interviews and analyzed using grounded theory and constant comparative analysis. Findings indicate that the framework supported more structured reasoning, reduced inefficient prompting behaviors, and improved students’ ability to identify weaknesses in their own arguments before digital interaction. Students also reported increased confidence and greater ownership of their academic work. The study argues that structured pre-AI planning activities can strengthen critical thinking and academic literacy by positioning AI as a support tool rather than a substitute for cognitive effort. The findings contribute practical guidance for educators seeking to integrate AI into higher education while preserving student agency and deep learning.

Keywords: generative AI, academic literacy, design-based research, cognitive agency, higher education, AI-supported learning

INTRODUCTION

The rapid integration of generative artificial intelligence (GenAI) into higher education has significantly transformed how students engage with academic tasks such as writing, problem solving, information synthesis, and content generation. While these technologies offer substantial opportunities for improving access to information and supporting learning processes, their increasing use has also generated growing concern regarding the potential impact on student cognition, critical engagement, and academic authorship (Bearman & Ajjawi, 2023; Selwyn, 2022). In particular, the ease and speed with which GenAI systems can produce sophisticated academic outputs raise important questions about the extent to which students remain actively involved in the intellectual processes underlying knowledge construction.

Traditional academic writing and problem-solving practices often require sustained engagement through drafting, revision, reflection, and conceptual restructuring. These processes play an important role in supporting deeper learning and the development of disciplinary understanding. However, AI-assisted environments may unintentionally encourage students to prioritize rapid task completion over deliberate cognitive engagement, particularly when interactions with AI systems rely primarily on iterative prompting and automated content generation (Bender et al., 2021). Although these tools can enhance productivity,

This article is based on research originally conducted as part of the author’s PhD thesis.

concerns have emerged regarding the possibility that excessive reliance on automated outputs may reduce opportunities for critical evaluation, conceptual clarification, and independent reasoning.

Within this context, the concept of cognitive agency becomes increasingly important. Cognitive agency refers to the learner's capacity to intentionally direct and regulate their own reasoning processes while maintaining ownership over the logic and structure of their work (Luckin & Bagshaw, 2025). Rather than functioning as passive recipients of artificial intelligence (AI)-generated content, students must be supported in developing the ability to critically evaluate, organize, and verify the reasoning that informs academic outputs. This challenge is particularly relevant in AI-mediated learning environments where the underlying decision-making processes of large language models (LLMs) remain largely opaque to users (Bearman & Ajjawi, 2023).

A growing concern within contemporary digital pedagogy is that many students engage with GenAI through forms of reactive "trial-and-error" prompting. In these interactions, students frequently submit broad or underdeveloped prompts and subsequently attempt to refine AI-generated responses through repeated iterative adjustments. While such practices may produce acceptable surface-level outputs, they may not necessarily support the development of robust mental models or deliberate reasoning processes (Sweller, 2023). Without a clearly articulated conceptual framework guiding the interaction, students may struggle to evaluate the accuracy, coherence, or limitations of AI-generated information.

In response to these concerns, this paper presents a pedagogical framework termed the "paper-first" approach, developed through a multi-year design-based research (DBR) study conducted within higher education academic literacy contexts. The framework requires students to externalize their reasoning processes through manual flow-diagramming and logic-gate architectures before engaging with AI-assisted digital tools. Drawing on principles associated with knowledge engineering (KE), cognitive load theory (CLT), and desirable difficulty (DD), the framework aims to support students in developing structured forms of "architectural prompting," where AI interaction is guided by deliberate logical planning rather than reactive prompting practices.

The paper argues that structured analogue scaffolding may help students maintain cognitive agency during AI-assisted academic work by encouraging explicit reasoning, conceptual clarification, and critical engagement with automated outputs. In doing so, the study contributes to emerging discussions regarding the role of pedagogical design in supporting meaningful human-AI collaboration in higher education.

To address these concerns, this study was guided by the following overarching research question: How can a structured, "paper-first" KE framework be designed to support and sustain students' cognitive agency during GenAI-assisted academic tasks?

To answer this question, the research pursued three primary design objectives:

1. To design and iteratively refine a pre-digital analogue workflow that scaffolds structured reasoning before digital execution.
2. To investigate how externalized logic blueprints affect students' capacity to critically evaluate and verify probabilistic AI outputs.
3. To evaluate the impact of intentional cognitive friction on the behavioral efficiency and independence of students' interactions with AI systems.

LITERATURE REVIEW

The increasing integration of GenAI into higher education has generated substantial debate regarding its implications for student learning, cognition, and academic practice. While AI-assisted systems offer new opportunities for supporting productivity and access to information, scholars have also raised concerns regarding cognitive offloading, reduced critical engagement, and the changing nature of academic authorship (Bearman & Ajjawi, 2023; Selwyn, 2022). This literature review examines research related to cognitive offloading, deep processing, CLT, KE, and human agency in AI-assisted learning environments. Together, these perspectives provide the theoretical foundation for the proposed "paper-first" framework.

Automation, Cognitive Offloading, and Deep Processing

Research on automation has long demonstrated that increasing technological efficiency may unintentionally reduce opportunities for skill development and critical oversight (Bainbridge, 1983). Within educational contexts, this concern is increasingly discussed through the concept of cognitive offloading, where individuals rely on external technologies to perform cognitive tasks that would otherwise require internal mental effort (Risko & Gilbert, 2016). Although cognitive offloading can support efficiency and reduce cognitive burden in some contexts, excessive reliance on automated systems may limit opportunities for deeper conceptual engagement and independent reasoning.

Recent scholarship on GenAI in higher education suggests that some students may engage primarily in forms of “surface curation,” where AI-generated content is lightly edited rather than critically constructed through iterative reasoning and synthesis (Lodge et al., 2020). This trend raises concerns regarding reduced opportunities for deep processing, which involves organizing, integrating, and reflecting on information in ways that support long-term learning. Sweller et al. (1998) argues that when automated systems provide excessive scaffolding, learners may engage less actively in the cognitive processes necessary for meaningful knowledge construction.

An additional concern involves the “black box” nature of LLMs. Unlike traditional rule-based systems, LLMs generate outputs through probabilistic associations that are often inaccessible to users (Bender et al., 2021). As a result, students may struggle to evaluate how conclusions were produced or to identify inaccuracies and biases within generated outputs (Bearman & Ajjawi, 2023). From a pedagogical perspective, this lack of transparency highlights the need for instructional approaches that encourage students to externalize and verify their own reasoning processes during AI-assisted work.

Intentional Friction and Analogue Scaffolding

The concept of DD provides an important theoretical lens for understanding how learning may be strengthened through effortful cognitive engagement. Bjork and Bjork (2020) argue that instructional conditions requiring greater mental effort often support stronger long-term retention and knowledge transfer than conditions emphasizing immediate fluency or efficiency. Within AI-assisted learning environments, the increasing ease of automated content generation may therefore reduce opportunities for the productive struggle associated with deeper learning processes.

CLT further supports the importance of carefully structured learning environments that minimize unnecessary cognitive distractions while supporting meaningful problem solving (Sweller, 2020). Although digital systems can assist learning, highly interactive AI interfaces may also encourage rapid iterative behavior that fragments attention and promotes reactive decision making. In contrast, paper-based activities may provide a more focused environment for organizing reasoning and managing intrinsic cognitive load.

Research on handwriting and haptic engagement also suggests that manual drawing and writing activities may support memory formation and conceptual processing. Studies have demonstrated that handwriting activates broader neural networks associated with sensory-motor processing and memory than typing alone (Ose Askvik et al., 2020; Van der Weel & Van der Meer, 2024). Within educational contexts, externalizing reasoning through diagrams and handwritten structures may therefore support deeper conceptual engagement and reflection.

Together, these perspectives suggest that intentional forms of analogue scaffolding may provide pedagogical value within AI-assisted learning environments by slowing down the reasoning process and encouraging more deliberate forms of cognitive engagement.

Knowledge Engineering and Externalized Reasoning

KE emerged within AI research as an approach focused on representing human expertise through explicit rule-based systems (Giarratano & Riley, 2024). Traditional expert systems rely on structured “if-then” logic and clear relationships between knowledge bases and inference mechanisms. Although contemporary AI systems increasingly rely on probabilistic machine learning approaches, the principles of explicit logical representation remain relevant within educational settings.

The concept of “mindtools” proposed by Jonassen (2000) is particularly important in this regard. Rather than functioning solely as information delivery systems, “mindtools” encourage learners to construct, organize, and represent their own understanding through interaction with structured technological environments. In educational contexts, structured systems such as flow-diagrams, decision trees, and expert system shells may therefore support higher-order reasoning by requiring students to externalize and formalize their thinking processes.

Research on visual thinking and logical mapping further suggests that external representations can help learners identify inconsistencies and gaps in reasoning (Davies, 2011; Tversky, 2019). By translating abstract reasoning into visible structures, students are better positioned to evaluate relationships between concepts and refine their arguments systematically. These approaches align with the current study’s emphasis on logic-gate architectures and flow-diagramming as tools for supporting deliberate reasoning during AI-assisted academic work.

Human Agency and Artificial Intelligence-Assisted Learning

The rapid adoption of AI technologies in higher education has also generated renewed interest in questions of human agency and authorship. Current scholarship increasingly emphasizes the importance of maintaining meaningful human oversight during AI-assisted academic processes (Luckin, 2018; Selwyn, 2022). Rather than positioning AI as a replacement for human cognition, many researchers advocate for models of human-AI collaboration that preserve student intentionality, critical judgment, and ethical responsibility.

These concerns are particularly significant within multilingual and culturally diverse educational contexts. Because LLMs are trained in dominant linguistic and cultural datasets, they may reproduce biases or fail to represent local forms of knowledge adequately (Bender et al., 2021). In the South African context, Kwet and Prinsloo (2020) caution that increasing dependence on platform-driven technologies may contribute to the marginalization of local epistemologies and linguistic diversity.

Within this context, pedagogical approaches that encourage students to externalize and critically evaluate their reasoning processes may support stronger forms of cognitive agency during AI-assisted learning. DBR provides an appropriate methodological framework for investigating such interventions because it emphasizes iterative refinement within authentic educational environments (Amiel & Reeves, 2008; McKenney & Reeves, 2019). The current study therefore draws on DBR to explore how paper-based KE activities may support more deliberate and critically engaged forms of AI interaction in higher education contexts.

THEORETICAL UNDERPINNINGS

The proposed “paper-first” methodology is informed by an interdisciplinary synthesis of cognitive psychology, educational technology, and knowledge representation theory. Rather than positioning digital tools as the primary drivers of learning, this framework emphasizes the importance of structured cognitive engagement prior to AI-assisted execution. In this view, learning is strengthened when learners are required to externalize and organize their reasoning before interacting with generative systems. The framework therefore draws on three complementary theoretical perspectives: KE, DD, and externalized cognitive representations. Together, these perspectives provide a foundation for understanding how structured analogue activities may support deeper cognitive engagement in AI-mediated learning environments.

Knowledge Engineering as Pedagogical Structuring

KE provides a useful theoretical lens for understanding how complex reasoning can be represented and structured in educational contexts. Traditionally associated with AI systems, KE focuses on the explicit representation of knowledge through formal rules, relationships, and logical structures (Giarratano & Riley, 2024). In educational applications, these principles can be adapted to support learners in articulating the structure of their reasoning processes more explicitly.

Rather than relying solely on natural language prompts or informal explanations, students are encouraged to represent their thinking through structured formats such as decision trees, flow diagrams, or rule-based logic sequences. This process shifts emphasis from surface-level content generation toward the systematic

organization of reasoning. In this way, KE becomes a pedagogical strategy for supporting clearer articulation of conceptual relationships and more deliberate problem decomposition.

Desirable Difficulty and Structured Cognitive Engagement

The concept of DD, as articulated by Bjork and Bjork (2020), suggests that learning is strengthened when learners engage in tasks that require effortful cognitive processing. Although such conditions may slow initial performance, they contribute to improved retention and transfer of knowledge over time.

Within the context of AI-assisted learning environments, where GenAI systems can rapidly produce coherent outputs, there is a risk that cognitive engagement becomes overly streamlined. From this perspective, the introduction of structured pre-digital tasks may serve to reintroduce productive cognitive effort into the learning process.

Paper-based activities such as manual diagramming and structured planning introduce a form of cognitive resistance that requires learners to slow down and articulate their reasoning more explicitly. This process can be understood as a form of structured cognitive engagement that supports deeper processing of information before digital execution. Rather than functioning as a limitation, this form of constraint may support more deliberate and reflective engagement with complex tasks.

Externalizing Mental Models through Representational Design

A recurring challenge in educational contexts is the difficulty students experience in making their implicit reasoning processes visible and examinable. Mental models are often internal, fragmented, and difficult to articulate, which can limit opportunities for reflection and revision. External representation provides a means of addressing this challenge by translating internal thought processes into observable structures.

Research on visual thinking and representational design suggests that diagrams, flowcharts, and other external cognitive artefacts can support learners in organising and evaluating their reasoning more effectively (Tversky, 2019). By externalizing thought, learners are able to identify inconsistencies, gaps, and alternative pathways within their reasoning structures.

In this context, flow-diagramming can be understood as a representational tool that supports cognitive transparency. It enables learners to examine the structure of their reasoning before it is translated into digital form. This aligns with findings in educational research which indicate that external representations can support metacognitive awareness and improve the quality of problem solving processes (Davies, 2011).

Rather than serving as a replacement for digital tools, these external representations function as preparatory cognitive structures that support more deliberate engagement with AI systems. They provide a visible framework against which automated outputs can be interpreted, evaluated, and refined.

METHODOLOGY

This study employed a DBR methodology to develop and evaluate a structured “paper-first” framework aimed at improving student reasoning in AI-assisted academic writing. DBR was selected because it supports iterative refinement of educational interventions within authentic classroom environments while simultaneously generating practical and theoretical insights.

Research Design

The research was conducted over a two-year period and consisted of nine DBR cycles. Each cycle followed a structured sequence of design, implementation, observation, and refinement. The intervention focused on introducing a mandatory paper-based planning phase prior to students engaging with GenAI tools and an expert system environment.

Across the cycles, task complexity was progressively increased while maintaining a consistent requirement for students to produce a manual logic diagram before any digital execution. This allowed the study to examine how structured pre-planning influenced students’ reasoning processes, their interaction with AI tools, and the quality of their academic writing outcomes.

Participants and Setting

The study was situated within a foundation-level academic literacy module, communication for academic purposes, within the faculty of information and communications technology at a South African university of technology. A purposive sampling strategy was employed to select participants who were actively learning academic writing protocols while simultaneously navigating digital literacy requirements in an information and communications technology (ICT) curriculum.

The participant group comprised two distinct cohorts:

1. An interdisciplinary design team ($n = 6$) consisting of two instructional designers, and four lecturers in communication and applied languages (average teaching experience = 14 years).
2. A student cohort ($N = 150$) consisting of first-year, foundation-level ICT students.

Demographically, the student cohort reflected a diverse multilingual profile typical of the institution, with over 85% speaking English as an additional language alongside home languages such as Sepedi, Setswana, IsiZulu, and Tshivenda. The cohort had uniform access to institutional computing facilities but demonstrated varying levels of prior experience with digital tools, ranging from regular informal use to absolute non-familiarity.

Intervention: "Paper-First" Framework

The intervention required students to complete a structured paper-based planning task before using any digital tools. This planning phase involved the manual construction of flow diagrams, the sequencing of logical relationships using decision structures, and the identification of key concepts and argument components prior to digital engagement. Only after completing this structured planning phase were students permitted to use GenAI tools and expert system software to complete their academic tasks.

Data Collection

Data were collected through multiple qualitative streams to provide a comprehensive, triangulated account of the framework's iteration across the DBR cycles. Focus group interviews served as the primary empirical source, comprising 14 distinct sessions (coded T1 to T14) targeting both the design and student cohorts:

- Ten semi-structured focus group sessions were conducted with the six members of the interdisciplinary design team. These sessions were held at a centralized institutional site over an extended six-month period. Each session lasted between 45 and 120 minutes, allowing for deep, exhaustive collaborative reflections, though institutional lecturing obligations occasionally required individual faculty members to excuse themselves mid-session.
- Four focus group sessions were conducted with a purposively selected sample from the student population, organized into two distinct groups. These sessions were hosted at a regional campus facility during the instructional implementation phase. To capture the evolution of student reasoning, interviews were administered longitudinally: two sessions occurred midway through the framework implementation, and two sessions were conducted at the conclusion of the training.

All focus group interactions were facilitated by the researcher using an emergent, semi-structured funneling approach. Sessions opened with expansive, general prompts (e.g., *"What are your impressions of the learning environment?"* or *"How did you experience what we did today?"*) before narrowing into more specific inquiries as concrete thematic insights emerged. The focus groups were recorded using handheld audio devices and subsequently transcribed verbatim to establish a stable textual database for subsequent qualitative analysis.

Data Analysis

Data were analyzed using grounded theory principles, with a constant comparative method guiding the analytical process. The analysis began with open coding of focus group transcripts, design team reflections, and observational notes. These codes were then systematically grouped into categories, which were subsequently refined into broader themes related to student reasoning processes, AI interaction patterns,

and task execution strategies. Comparison across DBR cycles allowed for continuous refinement of both the emerging categories and the intervention itself.

Trustworthiness and Analytical Rigor

To ensure the credibility and dependability of the analytical process, the study implemented strict qualitative controls. The raw dataset, consisting of the 14 verbatim focus group transcripts alongside continuous classroom observational notes, was subjected to a data triangulation protocol where student experiential accounts were constantly cross-checked against the design team's specialized reflections.

Because the researcher served in a dual capacity as both the facilitator and primary analyst, single-coder vulnerabilities were explicitly checked. An independent secondary coder evaluated the coding architecture against the raw text files. Coding discrepancies were discussed and resolved through a collaborative review session. Furthermore, the analytical process followed an iterative path from open coding to thematic consolidation, systematically moving from literal observations to high-level conceptual categories. This progression ensured that the ultimate design principles remained grounded directly in the physical and contextual realities documented across the distinct institutional research sites.

Data saturation was systematically monitored across the iterative cycles. By the conclusion of cycle 7 (spanning focus group T11 and group T12), no new conceptual codes or properties emerged regarding student prompting behaviors or verification strategies; subsequent cycles (cycle 8 and cycle 9) served to validate and consolidate the established thematic structure.

Transferability is addressed through thick, contextual descriptions of the institutional setting, participant demographics, and the explicit structural phases of the intervention layout. While the findings are context-bound, these detailed parameters allow educators in similar multilingual higher education environments to evaluate the applicability of the design principles to their own settings.

Ethical Considerations

This study was conducted in strict compliance with institutional ethical research protocols and received formal clearance from the university Research Ethics Committee (Ref: SM 10/10/02J). Prior to data collection, all participants were provided with written information sheets detailing the purpose, scope, and voluntary nature of the study. Signed informed consent was obtained from all participating students and design team members. To protect participant confidentiality, all data were fully anonymized using a structured transcript alphanumeric key, and all digital records were securely stored on password-protected cloud servers accessible only to the research team.

The "Paper-First" Framework (Intervention Design)

The "paper-first" framework is a structured instructional intervention developed through iterative DBR cycles. It requires students to complete an analogue-to-digital workflow in which problem-solving is first constructed on paper before any interaction with digital tools or GenAI systems.

The framework is designed to ensure that students develop a verified logical structure prior to AI use, thereby improving reasoning quality and reducing reliance on iterative prompting. It consists of three sequential phases:

- (1) manual logic construction,
- (2) logical validation, and
- (3) digital execution and verification.

Logic Blueprint Phase (Analogue Construction)

In the first phase, students construct a hand-drawn logic blueprint representing their solution to a given academic problem. This involves translating textual tasks into structured diagrams using:

- decision nodes
- IF-THEN rules
- Boolean operators (AND, OR, NOT)

Table 1. Transcript coding key

Component	Code	Description
Group	DT/S	Design team/students
Session	T1-T14	Focus group session
Participant	R/P	Respondent or participant ID

- flow-direction mapping

The purpose of this phase is to externalize student reasoning in a structured visual form before any digital input is used. Students are required to represent the full logical pathway of their solution, ensuring that all assumptions and decision points are explicitly documented.

This phase functions as a pre-digital structuring activity, ensuring that AI tools are used only after the logic of the task has been clearly defined.

Logical Validation and Ambiguity Resolution

Once the logic blueprint is completed, students review their diagrams to identify incomplete reasoning, contradictions, or missing logical links.

This phase focuses on detecting:

- unresolved decision branches
- inconsistent IF-THEN conditions
- incomplete causal pathways

Errors are corrected manually at this stage, with revisions made directly to the diagram. This creates a low-stakes iterative refinement process before digital implementation.

This step ensures that students enter the digital phase with a complete and logically coherent model, reducing reliance on AI to resolve conceptual gaps.

Cognitive Anchoring During Artificial Intelligence Interaction

In the final phase, students translate validated logic blueprints into digital tools, including structured rule-based environments and GenAI systems.

The completed paper-based model functions as a reference structure (“cognitive anchor”) during AI interaction. Students compare AI outputs against their predefined logic to verify correctness and identify inconsistencies.

This process supports systematic evaluation of AI-generated responses and reduces unstructured iterative prompting. The paper-based blueprint therefore acts as a verification tool throughout digital execution, ensuring that AI output remains aligned with student-defined logic structures.

FINDINGS: FROM PROMPTING TO STRUCTURED ARTIFICIAL INTELLIGENCE-ASSISTED REASONING

Findings from the nine DBR cycles indicate a consistent shift in student performance across three qualitatively distinct learning dimensions:

- (1) pre-AI reasoning structure,
- (2) critical evaluation of AI outputs, and
- (3) efficiency of AI interaction.

Across iterations, the “paper-first” framework functioned as a cognitive scaffold that improved the quality, coherence, and control of student engagement with GenAI.

To ensure analytical transparency and traceability, qualitative data were coded using a structured transcript identification system ([Table 1](#)), enabling consistent mapping of insights across DBR cycles.

Improved Pre-Artificial Intelligence Reasoning Structure

A key observed pattern was improved student ability to structure reasoning prior to AI interaction. In early DBR cycles, student prompts were broad and unstructured, requiring multiple revisions before producing usable outputs. The design team described this as a “poke and see mentality” (DT-T3), reflecting reactive rather than planned reasoning.

Following the introduction of mandatory paper-based logic diagrams, students demonstrated clearer pre-task organization and improved decomposition of problems prior to digital execution. This was reflected in more precise prompt construction and reduced ambiguity at the input stage. Students reported that the manual planning phase exposed gaps in reasoning before AI use, with one participant noting, “the paper makes the holes in your logic really obvious” (S-T4-P2). This experience was corroborated by another participant who remarked, “drawing the layout forces you to settle your own argument before the AI tries to write it for you” (S-T7-P5). This indicates improved cognitive structuring prior to task execution.

Enhanced Critical Evaluation of Artificial Intelligence Outputs

A second thematic shift was improved critical evaluation of AI-generated outputs. In early cycles, students tended to accept outputs with limited scrutiny, often relying on repeated prompting rather than systematic verification.

After the intervention, students increasingly used their logic blueprints as reference frameworks to assess alignment between intended arguments and AI-generated content. This enabled more targeted correction of errors and reduced reliance on regeneration cycles. One participant captured this shift, stating, “I’m building the rules, the computer is just filling in the words” (S-T5-P4). This developing sense of evaluative control was echoed in a later session: “having the paper open next to the screen meant I could catch exactly where the machine started introducing errors” (S-T12-P1). This reflects a transition from passive acceptance to structured verification behavior anchored in pre-defined reasoning structures.

Increased Interaction Efficiency with Artificial Intelligence Systems

A third noteworthy qualitative development was improved efficiency in AI interaction, observed through reduced prompt iteration cycles and more direct task completion. Prior to the intervention, students frequently engaged in iterative trial-and-error prompting loops, leading to inefficient task resolution.

Post-intervention data indicated fewer prompt revisions and improved first-attempt accuracy when AI interaction was informed by prior paper-based planning. Students attributed this improvement to clearer conceptual structuring before digital engagement. Reflexive thematic coding revealed that participant narratives naturally bridged the epistemic shift in pre-prompting conceptualization with downstream operational mechanics. For instance, as one participant explained, “you can’t hide a lack of logic when you have to draw it out” (S-T4-P2), illustrating a conceptual continuum where analogue planning simultaneously exposes internal reasoning gaps during initial design and directly reduces input ambiguity during prompt execution. This operational efficiency, corroborated across independent participant accounts (e.g., S-T9-P3), demonstrates how “paper-first” planning translates directly into streamlined digital interaction.

Disconfirming Cases

Despite the prevailing positive results, the qualitative analysis revealed important disconfirming cases characterized primarily by initial student resistance to the mandatory analogue (paper based) constraint. In earlier DBR cycles, a subset of participants viewed the “paper-first” requirements as an unnecessary burden that disrupted the expediency or convenience of digital interaction. This tension was particularly evident among students who prioritized speed over depth, with one participant noting: “Writing it down on paper felt like an unnecessary thing when I could just type on the computer and get answers” (S-T3-P6). This resistance highlights a critical threshold in the framework: for highly digitally fluent students, the intentional introduction of analogue friction can initially be perceived as an efficiency penalty rather than a cognitive scaffold, requiring explicit pedagogical mediation to reframe the value of pre-computation planning.

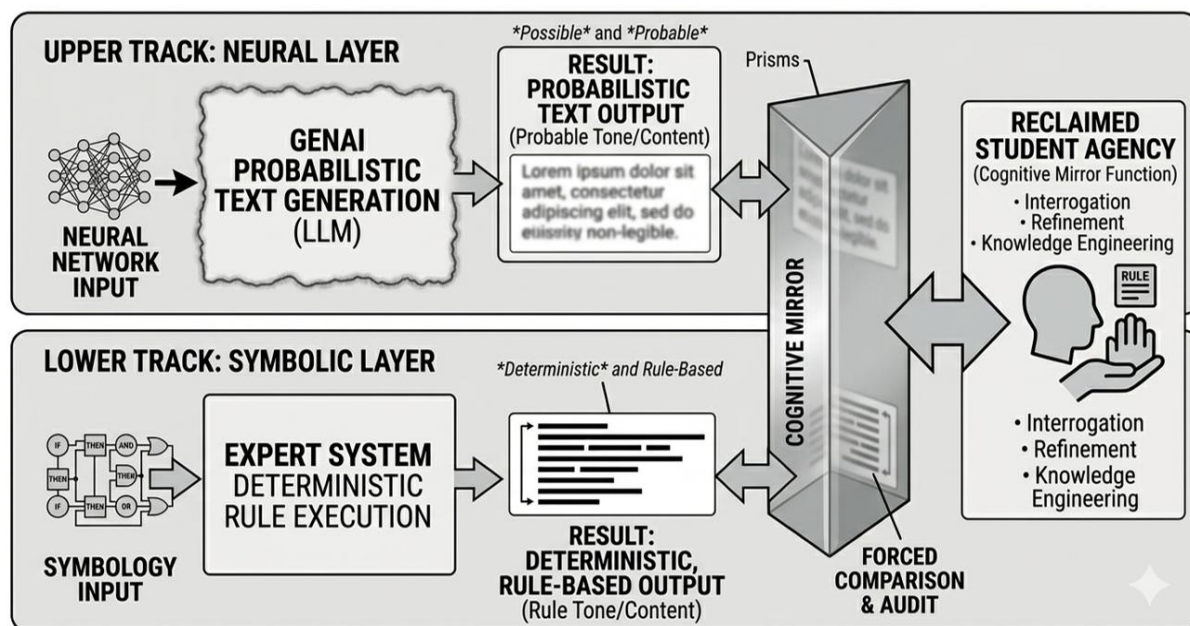


Figure 1. Dual-track reasoning model (developed by the author)

Integrated Outcome: “Paper-First” as a Cognitive Scaffolding Mechanism

Across all DBR cycles, the “paper-first” framework functioned as a cognitive scaffolding mechanism that improved three aligned learning outcomes: structured reasoning prior to AI use, enhanced evaluative control over outputs, and increased interaction efficiency.

Collectively, the findings demonstrate that structured analogue planning externalizes reasoning before digital execution, repositioning GenAI from an autonomous content generator to an execution tool governed by student-designed logic structures.

Figure 1 illustrates the dual-track structure in which students’ paper-based logic (symbolic layer) is used to guide and evaluate AI-generated outputs (probabilistic layer). This comparison enables systematic verification of AI responses against pre-defined reasoning structures.

DISCUSSION: PEDAGOGICAL IMPLICATIONS AND CRITICAL POSITIONING

The findings indicate that structured pre-AI planning through the “paper-first” framework can improve student reasoning structure, evaluative accuracy, and interaction efficiency when engaging with GenAI. However, these outcomes should be interpreted as context-dependent design effects rather than universal cognitive claims, consistent with DBR principles (McKenney & Reeves, 2019).

A key contribution of this study lies in repositioning GenAI from a primary content generator to a secondary execution tool governed by student-authored logic structures. This shift addresses a growing concern in educational technology literature: that frictionless AI use may weaken students’ ability to externalize reasoning before production (Lodge et al., 2020). The observed improvements in student performance suggest that externalized planning can function as a stabilizing mechanism for cognitive control during AI-mediated tasks, validating recent frameworks that call for explicit human-centered controls over automated writing tools (Lodge et al., 2020).

Importantly, the study does not claim that AI use inherently reduces cognitive agency. Rather, it demonstrates that agency is mediated by task design. Where students were required to construct explicit logic before AI interaction, they demonstrated stronger verification behavior and reduced reliance on iterative prompting. This finding aligns with established work on cognitive offloading and structured scaffolding (Risko & Gilbert, 2016; Sweller, 2023), while extending these paradigms directly into contemporary GenAI-mediated academic writing contexts (Luckin & Bagshaw, 2025; Sweller et al., 1998).

Addressing Validity and Generalizability Concerns

As a DBR study, the primary aim was not statistical generalization but design refinement in authentic learning environments. The nine iterative cycles across a foundation-level academic literacy context provide sufficient depth to support analytical generalizability, where findings inform principles transferable to similar instructional settings rather than populations.

However, limitations must be acknowledged. First, the study was conducted within a single institutional context and within a specific disciplinary module. Second, qualitative data relied on self-reported perceptions, which may introduce social desirability bias. Third, while improvements in efficiency and reasoning structure were observed, the study does not isolate individual causal variables due to the inherently iterative DBR design.

These limitations do not undermine the findings but rather define their scope of applicability. The framework should therefore be interpreted as a context-sensitive pedagogical intervention rather than a universal instructional model.

Addressing Concerns of Theoretical Overreach

While the study draws on KE, CLT, and DD, these frameworks function primarily as interpretive lenses rather than predictive mechanisms. Their purpose is to support the design rationale of the intervention rather than to assert direct neurocognitive causation. This distinction is critical in maintaining methodological integrity and avoiding theoretical over-extension.

CONCLUSION

This study set out to examine whether structured pre-AI reasoning, operationalized through a “paper-first” framework, could improve the quality and efficiency of student engagement with GenAI in an academic literacy context. The findings from nine DBR cycles suggest a consistent, context-specific improvement in three areas: the structuring of reasoning prior to AI use, the critical evaluation of AI-generated outputs, and the efficiency of AI interaction measured through reduced reliance on iterative trial-and-error prompting.

Importantly, the study does not position these outcomes as universal effects of AI use in education. Rather, they are interpreted as the result of a designed pedagogical condition, where cognitive performance was shaped by the requirement to externalize reasoning before digital execution. This distinction is critical for appropriately framing the contribution within DBR methodology, which prioritizes design refinement and contextual validity over statistical generalization.

A central conclusion is that GenAI does not inherently diminish or enhance cognitive agency. Instead, agency is redistributed according to task design structures. Where students were required to construct explicit logic models prior to AI engagement, they demonstrated stronger control over output verification and reduced dependence on iterative prompting. This suggests that the pedagogical challenge lies not in restricting AI use, but in structuring the conditions under which AI use occurs.

Implications for Pedagogical Design

The study offers three primary implications for higher education practice in AI-integrated learning environments.

First, pre-AI analogue structuring should be considered a viable instructional strategy for strengthening student reasoning. The findings indicate that requiring students to externalize logic before digital interaction supports clearer problem decomposition and more precise AI prompting. This has direct relevance for disciplines where argument construction and analytical clarity are core learning outcomes.

Second, AI literacy should extend beyond prompt engineering to include logic engineering competencies. Students in this study demonstrated improved evaluative capacity when they had a pre-constructed reasoning framework against which to assess AI outputs. This suggests that effective AI literacy includes the ability to design, test, and audit reasoning structures prior to automation.

Third, the study supports the pedagogical value of introducing deliberate cognitive friction in digitally mediated tasks. Contrary to assumptions that efficiency equates to better learning design, the findings

indicate that slowing down the initial reasoning phase can reduce downstream inefficiencies and improve task accuracy. This aligns with broader educational principles of structured difficulty but extends them into AI-mediated academic writing environments.

Limitations and Scope of Application

As a DBR study, the intention was not to produce statistically generalizable claims but to develop and refine a design intervention within a real educational setting. The findings are therefore best understood as analytical generalizations, offering transferable design principles rather than universal prescriptions.

The study is also limited by its single-institution context and reliance on qualitative self-report data alongside observational design team analysis. While triangulation across DBR cycles strengthens internal coherence, future research could extend this work through comparative institutional studies or mixed-methods validation of learning outcome measures.

Final Position

The “paper-first” framework contributes a structured pedagogical response to the increasing integration of GenAI in academic writing environments. Its core contribution is not technological but instructional: it demonstrates that the effectiveness of AI in education is contingent on the quality of prior human reasoning design. In this sense, the study reframes GenAI not as a replacement for cognitive work, but as a dependent system whose value is determined by the clarity and structure of student-authored logic.

Funding: The author received no financial support for the research and/or authorship of this article.

Acknowledgments: The author would like to thank the design team for their technical and pedagogical contributions: Prof. G. Genis, Mr. A. Harvey, Dr. G. Wissing, Dr. S. Roussow, and Dr. E. Rautenbach. The author would also like to thank the late Prof. J.G. Knoetze for his foundational guidance.

Ethics declaration: This study was approved by the Research Ethics Committee at Tshwane University of Technology on 9 April 2012 with approval number SM 10/10/02. Prior to data collection, all participants were provided with written information sheets detailing the purpose, scope, and voluntary nature of the study. Signed informed consent was obtained from all participating students and design team members. To protect participant confidentiality, all data were fully anonymized using a structured transcript alphanumeric key, and all digital records were securely stored on password-protected cloud servers accessible only to the research team.

AI statement: Generative AI tools were used for language refinement and proofreading of portions of the manuscript. All intellectual content, analysis, interpretation and final editorial decisions remain the responsibility of the author.

Declaration of interest: The author declared no competing interest.

Data availability: Data generated or analyzed during this study are available from the author on request.

REFERENCES

- Amiel, T., & Reeves, T. C. (2008). Design-based research and educational technology: Rethinking technology and the research agenda. *Educational Technology & Society*, 11(4), 29-40.
- Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775-779. [https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8)
- Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, 54(5), 1160-1173. <https://doi.org/10.1111/bjet.13337>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). ACM. <https://doi.org/10.1145/3442188.3445922>
- Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, 9(4), 475-479. <https://doi.org/10.1016/j.jarmac.2020.09.003>
- Davies, S. (2011). Concept mapping, mind mapping and argument mapping: What are the differences and do they matter? *Higher Education*, 62(3), 279-301. <https://doi.org/10.1007/s10734-010-9387-6>
- Giarratano, J. C., & Riley, G. (1994). *Expert systems: Principles and programming*. PWS Publishing Co.
- Jonassen, D. H. (2000). *Computers as mindtools for schools: Engaging critical thinking*. Prentice Hall.
- Kwet, M., & Prinsloo, P. (2020). The ‘smart’ classroom: A new frontier in the age of the smart university. *Teaching in Higher Education*, 25(4), 510-526. <https://doi.org/10.1080/13562517.2020.1734922> .

- Lodge, J. M., Kennedy, G., & Lockyer, L. (2020). Digital learning environments, the science of learning and the relationship between the teacher and the learner. In A. Carroll, R. Cunningham, & A. Nugent (Eds.), *Learning under the lens: Applying findings from the science of learning to the classroom* (pp.1-15). Routledge. <https://doi.org/10.4324/9780429027833-11>
- Luckin, R. (2018). *Machine learning and human intelligence: The future of education for the 21st century*. UCL IOE Press.
- Luckin, R., & Bagshaw, P. (2025). Charting an AI course: A framework for higher education strategy development. In S. Popenici, J. Rudolph, F. Ismail, & S. Tan (Eds.), *Handbook of artificial intelligence in higher education* (pp. 30-53). Edward Elgar Publishing. <https://doi.org/10.4337/9781035338764.00009>
- McKenney, S., & Reeves, T. C. (2019). *Conducting educational design research* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315105642>
- Ose Askvik, E., Van der Weel, F. R., & Van der Meer, A. L. H. (2020). The importance of cursive handwriting over typewriting for learning in the classroom: A high-density EEG study of 12-year-old children and young adults. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.01810>
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676-688. <https://doi.org/10.1016/j.tics.2016.07.002>
- Selwyn, N. (2022). *Education and technology: Key issues and debates* (3rd ed.). Bloomsbury. <https://doi.org/10.5040/9781350145573>
- Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*, 68, 1-16. <https://doi.org/10.1007/s11423-019-09701-3>
- Sweller, J. (2023) The development of cognitive load theory: Replication crises and incorporation of other theories can lead to theory expansion. *Educational Psychology Review*, 35, Article 95. <https://doi.org/10.1007/s10648-023-09817-2>
- Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296. <https://doi.org/10.1023/A:1022193728205>
- Tversky, B. (2019). *Mind in motion: How action shapes thought*. Basic Books. <https://doi.org/10.1145/3325480.3325525>
- Van der Weel, F. R., & Van der Meer, A. L. H. (2024). Handwriting but not typewriting leads to widespread brain connectivity: Highlighting the importance of handwriting in the classroom. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1219945>

