

Article

Not peer-reviewed version

The GPS for Thinking: How AI Partners Are Reshaping Student Metacognition

[Sayed Mahbub Hasan Amiri](#)^{*}, Naznin AKter, Marzana Mithila, Md. Mainul Islam

Posted Date: 25 May 2026

doi: 10.20944/preprints202605.1603.v1

Keywords: artificial intelligence in education; cognitive offloading; GPS metaphor; metacognition; self regulated learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

The GPS for Thinking: How AI Partners Are Reshaping Student Metacognition

Sayed Mahbub Hasan Amiri ^{1,*}, Naznin Akter ², Marzana Mithila ³ and Md. Mainul Islam ¹

¹ Department of CS, Dhaka Residential Model College, Bangladesh

² Department of English, Shamlapur Ideal Academy, Bangladesh

³ Department of Field Work, Unique Personnel (UK) Limited, UK

* Correspondence: amiri@drmc.edu.bd

Abstract

Generative artificial intelligence is rapidly becoming a cognitive partner in education, capable of planning tasks, monitoring progress, and evaluating solutions on a learner's behalf. This conceptual synthesis paper examines the risk that such AI tools, while improving immediate performance, may erode students' metacognitive abilities, their capacity to plan, monitor, and evaluate their own thinking. Drawing a parallel with GPS navigation research, where habitual turn-by-turn guidance has been shown to impair spatial memory and hippocampal engagement, we introduce the metaphor of AI as a "GPS for thinking." Through an integrative review of literature spanning cognitive psychology, neuroscience, and the learning sciences, we synthesise evidence that AI-assisted learning can lead to a form of cognitive disuse atrophy, specifically by short-circuiting the metacognitive loop. Emerging studies reveal that students who rely heavily on AI tutors often perform worse when the tool is removed, suffer from an illusion of explanatory depth, and struggle to articulate the reasoning behind their answers. To counter these effects, we propose a shift from a GPS model where the tool issues commands to a compass model, where the tool provides orientation while preserving learner agency. Five evidence-informed design principles are advanced: prompting planning before assistance, delaying and fading feedback, embedding mandatory reflection pauses, making AI reasoning visible, and calibrating learners' confidence. The article argues that the long-term goal of educational AI must be to strengthen, not supplant, the student's inner compass.

Keywords: artificial intelligence in education; cognitive offloading; GPS metaphor; metacognition; self-regulated learning

1. Introduction

In a sunlit classroom in the spring of 2026, a sixteen-year-old student named Lena sits before a laptop, wrestling with a physics problem about conservation of momentum. She types the question into an AI tutoring interface, and within seconds the screen populates with a sequence of flawless steps: identify the system, define initial and final states, apply the conservation equation, solve for the unknown velocity. Lena watches the reasoning unfold, nods along, and copies the final answer onto her worksheet. When her teacher kneels beside her and asks, "Can you explain how you solved it?", Lena's confidence falters. She can repeat the steps, but she cannot say why she began with that particular equation, what alternative approaches she considered, or what she would do if the problem were framed differently. She has arrived at the right destination, but she has no map of the terrain she crossed.

This scene, multiplied across millions of classrooms, captures a defining tension of the artificial intelligence era in education. Generative AI large language models and their multimodal successors has moved from speculative novelty to classroom infrastructure with breathtaking speed. Intelligent

tutoring systems, writing assistants, and problem-solving partners now promise to democratise expertise, personalise instruction, and close stubborn achievement gaps. Their ability to break complex tasks into manageable steps, offer instant feedback, and generate coherent explanations can reduce cognitive load for novice learners, making rigorous content accessible to students who might otherwise flounder. Yet as these tools become more capable and more seamlessly integrated, a quiet question grows louder: when we outsource the moment-by-moment navigation of a learning task to an external intelligence, what happens to the learner's own ability to navigate?

The question is not merely philosophical. It draws on a half-century of cognitive science that identifies *metacognition* the capacity to plan, monitor, and evaluate one's own thinking as one of the most potent predictors of academic achievement. First defined by Flavell (1979) as "cognition about cognitive phenomena," metacognition functions as the mind's internal guidance system, the quiet executive that asks, *What do I already know? Is this making sense? Should I change strategies? Did that work?* When students engage metacognitively, they build what might be called a cognitive map of the problem space: a mental representation that allows them to orient themselves not only in the present task but in future, unfamiliar ones. Hattie's (2009) monumental synthesis of educational research places metacognitive strategies among the highest-impact interventions known, with an effect size that rivals the influence of prior ability itself. And yet metacognition is not automatic. It develops slowly, through deliberate practice, in environments that require learners to articulate their reasoning, confront confusion, and make strategic choices (Azevedo & Gašević, 2019). It is, in other words, precisely the kind of cognitive function most vulnerable to being quietly replaced by an obliging AI.

This vulnerability becomes vivid when examined through an unlikely but instructive parallel: the way satellite navigation technology reshapes the human brain. Over the past two decades, cognitive neuroscientists have built a compelling body of evidence showing that habitual reliance on GPS devices degrades the neural and cognitive systems that support spatial navigation (Dahmani & Bohbot, 2017; Ishikawa et al., 2008). In a landmark study, participants who navigated an unfamiliar environment with a GPS-based system reached their destinations efficiently but later produced strikingly impoverished sketch maps compared to those who used paper maps or direct experience (Ishikawa et al., 2008). They had moved through space without constructing a mental model of it. Subsequent neuroimaging research confirmed that frequent GPS users exhibited reduced grey matter volume in the hippocampus and performed worse on self-guided navigation tasks (Dahmani & Bohbot, 2017). The mechanism at work is a form of *cognitive disuse atrophy*: when an external tool consistently performs a cognitive function, the brain, ever economical, reduces its investment in that function (Risko & Gilbert, 2016). The taxi drivers of London, who famously grow their posterior hippocampi through years of mapless navigation (Maguire et al., 2006), demonstrate the opposite principle what you use, you build; what you outsource, you lose.

Generative AI, we argue, presents education with a cognitive offloading challenge of analogous kind and greater scope. Where GPS offloads spatial reasoning, AI offloads *cognitive* reasoning the very processes of planning, monitoring, and evaluating that constitute metacognition. When a student like Lena receives a solution that has already been decomposed into logical steps, her own planning is preempted. When an AI silently corrects her errors in real time, her own monitoring is rendered unnecessary. When the final answer appears polished and accompanied by a fluent explanation, the reflective evaluation that consolidates learning is easily skipped. The AI becomes a GPS for thinking, navigating the intellectual terrain on the student's behalf. And the risk, as the navigation literature would predict, is not merely that students will become dependent on the tool in the moment, but that their internal cognitive mapping capacity their metacognitive skill will atrophy over time.

Early evidence from AI-in-education research already lends empirical weight to this concern. A recent study found that high school students who practiced mathematics with a ChatGPT-based tutor outperformed peers during AI-assisted sessions but scored significantly worse on a subsequent unassisted assessment, leading the researchers to conclude that generative AI without safeguards can

harm learning (Bastani et al., 2025). Studies in computing education reveal a related pattern: students who use AI code generators can describe what the generated code does but struggle to articulate the strategic reasoning behind their approach a clear breakdown in metacognitive monitoring and evaluation (Prather et al., 2024). These findings align with cognitive research showing that learners often mistake the fluency of an external explanation for their own understanding, an “illusion of explanatory depth” (Rozenblit & Keil, 2002) that AI’s smooth, authoritative prose may uniquely amplify. Roelle et al. (2017) demonstrated that such fluency experiences can inflate confidence while leaving actual comprehension unchanged, a dynamic that directly undermines the self-monitoring component of metacognition. The emerging picture is one of fragile performance gains that mask deeper developmental costs.

Yet the appropriate response to this evidence is not a retreat from AI. The GPS analogy, properly understood, does not indict navigation technology; it indicts *unexamined* navigation technology that asks nothing of the user. Some cities have redesigned navigation interfaces to display compass orientation, show whole-route previews, and prompt drivers to identify landmarks small acts of cognitive friction that keep the hippocampus engaged. The same principle can and should guide the design of AI learning partners. The goal is to shape tools that scaffold metacognition rather than supplant it, transforming the AI from a GPS that issues commands into a *compass* that provides orientation while preserving the learner’s responsibility for the journey.

This paper pursues that transformation through three guiding questions. First, how does AI offloading affect student metacognition? We review the cognitive architecture of metacognition and the ways in which current AI tools bypass its core components. Second, what can we learn from the cognitive effects of GPS navigation? We mine the spatial cognition literature for a framework cognitive disuse atrophy that illuminates the neural and behavioural consequences of systematic offloading, and we map that framework onto the domain of intellectual learning. Third, how can AI be designed to strengthen rather than weaken metacognitive skills? Drawing on research in self-regulated learning, feedback timing, and metacognitive scaffolding, we derive five evidence-informed design principles and advance a “compass model” of educational AI that prioritises durable human capability over momentary task completion.

The paper is structured as follows. Section 2 synthesises the interdisciplinary literature across four domains: GPS and spatial cognition, metacognition and academic achievement, AI in education and cognitive disuse, and design principles for metacognitive scaffolding. Section 3 details our integrative review methodology. Section 4 presents our synthetic framework, including the compass model and the five design principles, and discusses implications for practice, policy, and future research. A concluding section restates the central argument: that in learning, as in navigation, the destination matters less than the map, and that the most important thing an educational tool can do is leave the learner with a stronger internal compass.

2. Literature Review

2.1. Navigation Technology and Cognitive Offloading: The GPS Precedent

The cognitive consequences of offloading spatial reasoning to navigation devices provide the central analogy for this paper. A robust body of research demonstrates that habitual GPS use diminishes both behavioural and neural markers of spatial competence. Ishikawa et al. (2008) compared wayfinding performance among pedestrians using GPS-based mobile navigation, paper maps, or direct experience. Those in the GPS condition reached destinations with the fewest errors during navigation but subsequently produced significantly less accurate sketch maps, indicating impoverished mental representations of the traversed environment. Their efficient movement had not translated into durable spatial knowledge.

This behavioural finding has been corroborated and extended by neuroimaging evidence. Dahmani and Bohbot (2017) reported that individuals who habitually use GPS to navigate exhibit reduced grey matter volume in the hippocampus and lower hippocampal activity during self-guided

navigation tasks. Moreover, they found a dose-response relationship: more frequent GPS use predicted poorer spatial memory performance and greater difficulty forming cognitive maps. The study demonstrated that the hippocampus, a structure famously shown to enlarge in London taxi drivers who acquire extensive spatial knowledge through years of mapless navigation (Maguire et al., 2006), can also atrophy when systematic spatial demands are removed.

The underlying mechanism has been termed *cognitive offloading* the use of physical action or external tools to reduce the cognitive demands of a task (Risko & Gilbert, 2016). Offloading becomes cognitively hazardous when it is not strategic but habitual, such that the internal capacity it replaces undergoes disuse atrophy. Risko and Gilbert (2016) note that offloading changes the nature of cognitive processing rather than simply reducing it; the brain adapts to the reduced demand, resulting in a long-term decline in the offloaded capacity. Crucially, these effects are not inevitable consequences of tool use but of tool design: navigation interfaces that incorporate orientation cues, route previews, and active decision points can mitigate hippocampal disengagement. This distinction is central to our subsequent analysis of AI in education.

2.2. Metacognition: Components, Development, and Educational Impact

Metacognition, the ability to monitor and regulate one's own cognitive processes, was first systematically theorised by Flavell (1979), who defined it as "knowledge and cognition about cognitive phenomena." Contemporary frameworks typically decompose metacognition into two broad components: metacognitive knowledge (declarative understanding of one's own cognition and of strategies) and metacognitive regulation (the active processes of planning, monitoring, and evaluating during task performance) (Schraw & Moshman, 1995). The regulatory loop is of particular interest here because it constitutes the moment-by-moment cognitive navigation that AI tools are increasingly positioned to automate.

Planning involves the selection of appropriate strategies and the allocation of cognitive resources before engaging with a task. *Monitoring* refers to the ongoing awareness of comprehension and performance, enabling a learner to detect confusion, errors, or impasses. *Evaluation* involves appraising the outcomes and processes after task completion, feeding forward into future strategic decisions (Schraw & Dennison, 1994). Empirical research consistently shows that stronger metacognitive skills correlate with higher academic achievement across age groups and domains. Hattie's (2009) synthesis of over 800 meta-analyses reported a large effect size for metacognitive strategies, making them one of the most powerful educational interventions identified. More recent meta-analytic work confirms this finding and extends it to digital learning environments, where metacognitive prompting has been shown to improve learning outcomes significantly (Zheng et al., 2019).

Critically, metacognition does not develop automatically through content exposure. It requires deliberate practice in environments that prompt learners to articulate their thinking, confront discrepancies, and reflect on strategy use (Azevedo & Gašević, 2019). This developmental requirement makes metacognition vulnerable to offloading in technology-rich settings: if an AI partner performs planning, monitoring, and evaluation on the learner's behalf, the learner may have fewer opportunities to practice and internalise these regulatory processes.

2.3. AI in Education and the Risk of Metacognitive Disuse

The rapid deployment of generative AI in educational contexts has produced an early but instructive body of empirical evidence. Bastani et al. (2025) conducted a field experiment in which high school students practiced mathematics with access to a ChatGPT-based tutor. The AI group outperformed the control group during practice sessions. However, on a subsequent unassisted assessment, the pattern reversed: the AI group scored significantly lower, with the negative effect concentrated among students who had used the AI most extensively. The authors concluded that generative AI, when deployed without safeguards, can harm independent learning.

Qualitative studies in computing education reveal a complementary pattern: the erosion of metacognitive monitoring and evaluation. Prather et al. (2024) found that students who used AI code generators such as ChatGPT and GitHub Copilot could accurately describe the function of AI-generated code but struggled to articulate their own problem-solving strategies, justify design choices, or debug errors independently. This dissociation between being able to recognise a correct solution and being able to generate or evaluate one's own reasoning is the hallmark of bypassed metacognition. The AI had handled the planning and monitoring, leaving students with a surface-level understanding that collapsed under independent demands.

This phenomenon is consistent with the broader cognitive literature on the *illusion of explanatory depth* (Rozenblit & Keil, 2002), whereby individuals systematically overestimate their understanding of phenomena until they are required to produce a detailed causal explanation. The fluency of an AI-generated explanation its coherence, structure, and polished language can create a powerful feeling of knowing that masks genuine gaps in comprehension. Roelle et al. (2017) demonstrated that informing learners about the risks of cognitive offloading and building in strategic friction reduced overconfidence and improved learning outcomes, suggesting that metacognitive awareness of offloading risks can partly mitigate them.

2.4. Design Principles for Metacognitive Scaffolding in AI Tools

If the risk is that AI tools will passively replace metacognitive processes, the solution lies in designing tools that actively scaffold them. Research on self-regulated learning with technology offers a foundation. Azevedo and Gašević (2019) have argued that advanced learning technologies should prompt learners to set goals, monitor understanding, and engage in strategic reflection, rather than merely delivering content. In a similar vein, Wise and Hsiao (2019) demonstrated that embedding structured reflection prompts in online discussion platforms significantly increased the quality of student contributions and metacognitive engagement.

More specific design insights emerge from the feedback literature. The timing and form of feedback are critical moderators of whether it supports or undermines metacognitive development. Immediate, complete feedback can reduce errors during practice but produce worse long-term retention and transfer than delayed or partial feedback because it removes the demand for self-monitoring and error detection (Butler & Winne, 1995). This suggests that AI systems should provide graduated, "faded" support: starting with hints rather than full solutions, and progressively reducing assistance as learner competence grows. Such fading aligns with established cognitive load principles and with the broader finding that productive difficulty desirable difficulties enhances durable learning (Bjork & Bjork, 2011).

Confidence calibration is another promising target. The persistent gap between learners' perceived and actual understanding can be narrowed by asking them to explicitly rate their confidence before receiving feedback (Dunlosky & Rawson, 2012). An AI partner that tracks calibration over time and feeds it back to the learner could function as a metacognitive mirror, training the very self-monitoring skill that unreflective AI use tends to erode.

Finally, research on explanation and self-explanation provides a design principle for making AI reasoning processes visible. When learners are asked to compare their own reasoning with an external model rather than simply receiving the model as a finished product, they engage in deeper monitoring and evaluation (Chi, 2000). An AI that reveals its decision tree or alternative considered paths, and invites the learner to compare, becomes a tool for metacognitive dialogue rather than passive reception.

Synthesising these threads, we identify five evidence-informed design principles for metacognitive AI: (1) prompt planning before providing assistance; (2) delay and fade feedback to preserve self-monitoring demands; (3) embed mandatory reflection pauses after task completion; (4) make the AI's reasoning process visible; and (5) calibrate learner confidence systematically. These principles form the basis of the "compass model" that we articulate and discuss in Section 4.

3. Methodology

3.1. Research Design and Rationale

The study is designed as an integrative literature review, a method particularly suited to examining emerging, interdisciplinary topics where empirical findings from disparate fields must be synthesised into a novel theoretical framework (Torraco, 2005; Snyder, 2019). Traditional systematic reviews are optimised for well-defined, mature research domains in which a large body of homogenous empirical studies can be aggregated and effect sizes calculated. Integrative reviews, by contrast, are appropriate when a research area is nascent, fragmented, and conceptually diverse, requiring the creative synthesis of heterogeneous sources to generate new insights (Torraco, 2005). The intersection of generative AI, cognitive offloading, and metacognition is precisely such a domain: it draws on cognitive neuroscience, educational psychology, human-computer interaction, and learning design, and the most relevant bodies of evidence are neither fully overlapping nor organised under a single disciplinary heading.

An integrative review also permits the inclusion of both empirical and theoretical literature (Whittemore & Knafl, 2005), enabling the paper to juxtapose well-established findings from spatial cognition such as the neurological consequences of GPS use with emerging but still preliminary evidence on AI-assisted learning. This design choice is consistent with Snyder's (2019) argument that literature reviews can serve as a standalone research method when their aim is theory building rather than effect-size estimation.

The review's overarching goal is the construction of a conceptual framework the "compass model" and a set of design principles grounded in cross-domain pattern recognition. To achieve this, the methodology incorporates three integrated analytical strategies: thematic synthesis of within-domain findings, comparative analysis across domains (spatial navigation and AI-assisted learning), and metaphor-driven theory development.

3.2. Literature Search and Selection

The search strategy was designed to capture the interdisciplinary character of the research questions while ensuring transparency and replicability. Searches were conducted between January and March 2026 across five electronic databases selected for their disciplinary coverage: PsycINFO (cognitive and educational psychology), ERIC (education research), PubMed (neuroscience and cognitive science), Scopus (multidisciplinary), and Google Scholar (supplementary and grey literature). The choice of databases reflects the paper's ambition to integrate findings from neuroscience, education, and technology design.

Search terms were organised into four thematic clusters corresponding to the paper's core constructs. The first cluster targeted *GPS navigation and cognitive offloading*: ("GPS" OR "satellite navigation" OR "turn-by-turn navigation") AND ("spatial memory" OR "hippocampus" OR "cognitive map" OR "cognitive offloading"). The second cluster addressed *metacognition*: ("metacognition" OR "metacognitive" OR "self-regulated learning" OR "SRL") AND ("planning" OR "monitoring" OR "evaluation" OR "self-assessment"). The third cluster focused on *AI in education*: ("generative AI" OR "ChatGPT" OR "AI tutor" OR "intelligent tutoring system") AND ("learning outcomes" OR "cognitive offloading" OR "metacognition" OR "self-explanation"). The fourth cluster combined *design principles*: ("metacognitive scaffolding" OR "feedback timing" OR "confidence calibration" OR "reflection prompts") AND ("educational technology" OR "AI").

Boolean operators combined clusters to generate the full search corpus. For example, the central integrative search combined cluster one terms with cluster two and three terms: ("GPS" OR "cognitive offloading") AND ("metacognition" OR "self-regulated learning") AND ("AI" OR "intelligent tutoring system"). Snowball sampling from reference lists of key papers supplemented the database searches, as did forward citation tracking of seminal works such as Ishikawa et al. (2008), Maguire et al. (2006), and Flavell (1979).

Inclusion criteria were: (a) peer-reviewed empirical studies, meta-analyses, or major theoretical reviews published in English; (b) publication dates between 2000 and 2026, with the exception of foundational theoretical works from earlier periods; (c) relevance to at least one of the four thematic clusters; and (d) for empirical studies, clear reporting of methods and findings. Grey literature, such as the Bastani et al. (2025) working paper, was included when widely cited and methodologically transparent, given the rapid pace of AI research outstripping traditional publication cycles. Exclusion criteria comprised opinion pieces without empirical or theoretical grounding, studies focused exclusively on non-cognitive outcomes such as motivation or affect without a metacognitive or cognitive component, and papers addressing AI in non-educational contexts (e.g., industrial automation) without transferable cognitive implications.

The initial search yielded 647 records after duplicate removal. Title and abstract screening eliminated 392 records that were clearly off-topic or failed inclusion criteria. Full-text review of the remaining 255 records resulted in a final corpus of 78 sources that directly informed the synthesis. A PRISMA-style flow diagram (Moher et al., 2009) was maintained to document the selection process, though formal meta-analytic aggregation was not the method's objective. The final corpus comprised studies from spatial cognition (n = 14), metacognition and self-regulated learning (n = 23), AI in education (n = 18), and design principles and cognitive theory (n = 23).

3.3. Data Extraction and Quality Assessment

Given the integrative and theory-building nature of the review, data extraction followed a flexible but systematic protocol. For each included source, the following information was recorded: author(s) and year, disciplinary domain, study design (empirical, theoretical, review), key constructs investigated, principal findings, and relevance to the paper's guiding questions. A standardised extraction form was piloted on a subset of ten papers and refined before full application.

Quality appraisal was conducted using an adapted version of the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018) for empirical studies, supplemented by the AACODS checklist (Tyndall, 2010) for grey literature. Theoretical and review papers were assessed for conceptual clarity, internal consistency, and breadth of synthesis. No source was excluded solely on quality grounds; instead, quality ratings were used to weight the interpretive confidence assigned to different bodies of evidence. The well-replicated findings from spatial navigation neuroscience, for example, received stronger weight than the preliminary and largely correlational findings from AI-in-education studies, a distinction reflected in the paper's cautious language about the latter.

3.4. Analytical Framework

The analytical procedure unfolded in three interconnected phases: thematic synthesis, comparative cross-domain analysis, and metaphor-driven theory development.

Thematic synthesis. Following the procedures outlined by Thomas and Harden (2008), the synthesis began with inductive coding of findings within each thematic cluster. For example, within the GPS navigation cluster, recurring descriptive themes included "hippocampal grey matter reduction," "sketch map impoverishment," "dose-response relationship with offloading," and "interface design as moderator." Within the AI-in-education cluster, themes included "performance decline on unassisted tasks," "dissociation between recognition and generation," "fluency-induced overconfidence," and "reflection prompts as protective factor." These descriptive themes were then aggregated into analytical themes that cut across clusters, such as "cognitive disuse atrophy," "metacognitive bypass," and "strategic friction."

Comparative cross-domain analysis. The thematic outputs from the GPS and AI clusters were systematically compared to identify structural parallels and disanalogies. This comparative process was guided by the principle of analogical transfer (Gentner & Markman, 1997): the mapping of relational structures from a well-understood source domain (spatial navigation and the hippocampus) to a less-understood target domain (AI-assisted cognitive navigation and metacognition). Particular attention was paid to identifying the functional equivalents of key constructs spatial planning

mapped to cognitive planning, GPS turn-by-turn commands to AI step-by-step solutions, and hippocampal atrophy to metacognitive skill decline while noting boundary conditions where the analogy does not hold.

Metaphor-driven theory development. The final analytical phase employed the GPS-compass metaphor as an integrative device. Metaphors in scientific reasoning serve not merely as rhetorical ornament but as cognitive tools that structure problem representation and hypothesis generation (Lakoff & Johnson, 1980). The GPS metaphor was used diagnostically: it framed the problem by highlighting the structural similarity between offloading spatial navigation and offloading cognitive navigation. The compass metaphor was then developed prescriptively, as a target image for redesigned AI tools that provide orientation without removing agency. The five design principles were derived abductively from the intersection of the empirical evidence on metacognitive scaffolding and the normative ideal captured by the compass metaphor.

3.5. Derivation of Design Principles

The five design principles prompting planning before assistance, delaying and fading feedback, embedding mandatory reflection pauses, making AI reasoning visible, and calibrating learner confidence were not selected a priori. They emerged from the iterative interplay of literature review and cross-domain comparison. For each principle, the following reasoning was applied: a specific metacognitive vulnerability identified in the AI literature (e.g., bypassed planning) was matched with a protective intervention from the self-regulated learning or feedback literature (e.g., planning prompts) and translated into a concrete AI design specification. The principles were then cross-checked against the GPS-compass framework to ensure they moved the AI's functionality from commanding to orienting.

3.6. Trustworthiness and Rigour

In qualitative and conceptual synthesis research, trustworthiness is typically evaluated against criteria of credibility, transferability, dependability, and confirmability (Lincoln & Guba, 1985). Credibility was pursued through triangulation across disciplines, ensuring that claims are supported by evidence from multiple independent research traditions. Transferability was addressed by providing thick description of the theoretical constructs and design principles so that educators, developers, and researchers can assess their applicability to diverse contexts. Dependability and confirmability were supported by maintaining an audit trail of search strategies, inclusion decisions, and analytical memos, and by explicitly articulating the metaphorical lens that shapes the interpretation.

3.7. Methodological Limitations

Several limitations inhere in the chosen method. First, integrative reviews are inherently interpretive, and the cross-domain analogical reasoning that yields insight also introduces the risk of overextending the metaphor. The spatial navigation literature does not provide a perfect analogue for cognitive navigation, and differences between hippocampal and prefrontal systems must be respected. Second, the AI-in-education evidence base is still small, predominantly correlational, and subject to publication bias toward novel findings. As more rigorous longitudinal and experimental studies accumulate, the weight assigned to different strands of the argument may require adjustment. Third, the design principles advanced in this paper are theoretically grounded but empirically untested as an integrated set; their validation through classroom-based design experiments constitutes necessary future work. These limitations notwithstanding, the method provides a coherent, evidence-informed framework for understanding and responding to an urgent educational challenge.

4. Discussion: From GPS to Compass - A Framework for Metacognitive AI Design

The preceding sections have established a multi-disciplinary evidential chain: habitual GPS use provides a validated model of cognitive disuse atrophy (Dahmani & Bohbot, 2017; Ishikawa et al., 2008); metacognition constitutes an internal cognitive navigation system whose planning-monitoring-evaluation loop is highly susceptible to offloading (Flavell, 1979; Hattie, 2009); and early evidence from AI-in-education research confirms that unreflective use of generative AI tools can degrade independent performance and metacognitive accuracy (Bastani et al., 2025; Prather et al., 2024). This section moves from diagnosis to prescription. It synthesises these findings into a coherent framework the compass model and derives five evidence-informed design principles that can guide the development of AI learning partners that strengthen, rather than supplant, the learner's inner compass.

4.1. Synthesis of Cross-Domain Evidence

Before presenting the design principles, it is worth crystallising the key insight that emerges from the comparative analysis. The GPS literature demonstrates that when an external tool takes over the moment-by-moment planning and error-correction of a navigation task, the user's internal spatial mapping capacity declines. The AI-in-education literature, though still developing, reveals an analogous pattern: when an AI tutor decomposes a problem, supplies steps, corrects errors silently, and delivers a polished final explanation, the learner's internal cognitive mapping capacity metacognition is bypassed.

The structural parallels are striking. In both cases, performance during tool use is maintained or even enhanced, creating an illusion of competence. In both cases, the degradation becomes visible only when the tool is removed. In both cases, the mechanism is not passive forgetting but active adaptation: the brain reallocates resources away from functions that are consistently unused (Risko & Gilbert, 2016). And crucially, in both cases, the outcome is not inevitable but contingent on design. Just as navigation interfaces can incorporate orientation cues and decision points that preserve hippocampal engagement, AI learning tools can be designed to introduce strategic cognitive friction that keeps metacognitive faculties engaged.

The design challenge, then, is to identify the precise points in the learner-AI interaction where metacognitive bypass occurs and to insert evidence-based scaffolds that demand and develop the learner's own planning, monitoring, and evaluation. The five principles that follow are not exhaustive; rather, they represent a coherent, theoretically grounded starting point for redesigning AI partners as cognitive compasses.

4.2. Five Evidence-Informed Design Principles

Table 1. Five Design Principles for Metacognitive AI.

| Principle | Description | Metacognitive Phase Supported | Primary Risk Addressed | Key Supporting Evidence |
|-----------------------|--|-------------------------------|---|--------------------------------------|
| 1. Ask before you act | Prompt the learner to articulate a plan or initial strategy before the AI offers assistance. | Planning | Bypassed strategy selection; reception of AI-generated steps. | Azevedo & Gašević (2019); Chi (2000) |

| | | | | |
|--------------------------------|--|---------------------------------------|--|--|
| 2. Delay and fade feedback | Provide graduated hints rather than full solutions; reduce support over time as competence grows. | Monitoring | Suppression of error detection; learned helplessness without the tool. | Butler & Winne (1995); Renkl & Atkinson (2003) |
| 3. Mandatory reflection pauses | Insert structured post-task prompts that ask learners to evaluate their process and identify difficulties. | Evaluation | Skipped consolidation; failure to update strategic knowledge. | Wise & Hsiao (2019); Roelle et al. (2017) |
| 4. Make the process visible | Reveal the AI's reasoning path, decision points, and alternatives considered; invite comparison with the learner's own thinking. | Monitoring and evaluation | Fluency illusion; shallow understanding of the solution. | Rozenblit & Keil (2002); VanLehn (2011) |
| 5. Confidence calibration | Ask learners to rate their confidence before revealing answers; track calibration over time and feed it back. | Monitoring (self-assessment accuracy) | Overconfidence; illusion of explanatory depth. | Dunlosky & Rawson (2012); Hacker et al. (2008) |

Note. This table synthesises the five design principles derived from the integrative literature review. Each principle targets a specific phase of the metacognitive loop and is supported by empirical evidence discussed in the main text.

Principle 1: Ask before you act (prompt planning).

The first principle addresses the planning phase of the metacognitive loop. In many current AI tutoring interfaces, a student enters a problem and is immediately presented with a step-by-step solution. The AI has planned the approach on the student's behalf. To counter this, the AI should prompt the learner to articulate a plan before any assistance is provided for example, "Write one sentence describing how you think you might start." Even a brief, imperfect planning attempt activates the prefrontal executive networks responsible for goal-setting and strategy selection (Azevedo & Gašević, 2019) and ensures that the AI's subsequent guidance supplements rather than replaces the learner's own strategic thinking. This approach is consistent with the self-explanation effect, whereby generating explanations before receiving instruction enhances learning (Chi, 2000).

Principle 2: Delay and fade feedback.

Instant, complete feedback eliminates the need for the learner to monitor their own comprehension. As Butler and Winne (1995) established in their theoretical synthesis, feedback that is too immediate or comprehensive can short-circuit self-regulated learning by removing the opportunity for internal error detection. The second principle therefore requires that AI feedback be delayed allowing the learner time to detect confusion or errors independently and faded over time. Initially, the AI might provide a hint that points toward the nature of an error without correcting it; as the learner gains competence, the AI reduces its scaffolding, moving from worked examples to completion problems to independent practice (Renkl & Atkinson, 2003). This fading procedure, long established as effective in intelligent tutoring systems (VanLehn, 2011), preserves the desirable difficulty that builds durable learning (Bjork & Bjork, 2011).

Principle 3: Mandatory reflection pauses.

The evaluation phase of metacognition reflecting on what worked, what did not, and what might be done differently is the component most frequently omitted in the rush toward task completion. AI learning partners often present a seamless sequence of problems without any structural pause for reflection. The third principle mandates that after each task or session, the AI inserts a structured reflection prompt: “What part of this task was most confusing?” or “If you solved it again, what would you do differently?” Research by Wise and Hsiao (2019) demonstrates that embedding such prompts in digital learning environments significantly increases the quality of student contributions and metacognitive awareness. Roelle et al. (2017) further showed that reflection prompts combined with warnings about the risks of cognitive offloading improved both monitoring accuracy and learning outcomes.

Principle 4: Make the process visible.

The polished, authoritative surface of AI-generated explanations contributes to the fluency illusion learners mistake the ease of following a solution for genuine understanding (Rozenblit & Keil, 2002). To disrupt this illusion, the fourth principle requires the AI to make its own reasoning process visible. Rather than presenting a single, seamless chain of steps, the AI could display a decision tree showing alternative paths considered and rejected, or it could annotate its solution with explicit justifications. It would then ask the learner: “How does your own reasoning compare?” This turns the AI’s output from a monologue into a dialogue and invites the learner to monitor and evaluate not only their own thinking but the machine’s. VanLehn (2011) notes that expert tutors often externalise their reasoning precisely to model metacognitive processes, and this principle extends that practice to AI.

Principle 5: Confidence calibration.

A persistent finding in the metacognition literature is that learners, particularly novices, are poor at calibrating their confidence with their actual performance (Dunlosky & Rawson, 2012; Hacker et al., 2008). AI tools risk amplifying this miscalibration by providing fluent explanations that feel complete. The fifth principle embeds confidence calibration directly into the AI interaction: before revealing an answer or a step, the AI asks, “How confident are you that your answer is correct? (0–100%).” Over time, the system can track the discrepancy between confidence and accuracy and feed this information back to the learner, creating a metacognitive mirror. This approach trains the self-monitoring faculty that unreflective AI use tends to erode. Hacker et al. (2008) demonstrated that students who receive calibration feedback show significant improvements in both monitoring accuracy and academic performance over time.

4.3. The Compass Metaphor and Its Design Implications

The five principles, taken together, operationalise a broader philosophical shift in how we conceive of AI learning partners. The GPS metaphor captures the risk: a tool that issues commands (“Turn right. You have arrived.”), requires no cognitive engagement from the user, and produces dependence. The compass metaphor, by contrast, captures the aspiration: a tool that provides orientation (“North is that way.”), supports the user in making their own navigational decisions, and develops, through repeated use, a strengthened internal sense of direction.

A compass does not dictate the path; it provides a stable reference point that enables the traveller to read the terrain, choose a route, and correct course independently. An AI designed according to the five principles would function analogously. It would help the learner orient within a problem space identifying relevant concepts, flagging common pitfalls, suggesting possible strategies without walking every step. It would demand that the learner remain present and responsible. And it would, over time, become less necessary as the learner's internal cognitive map grows more detailed and more reliable.

This shift has implications that extend beyond individual design features to the architecture of AI tutoring systems. The compass model implies that an AI's success should be measured not only by the correctness of the learner's immediate answers but by growth in their independent cognitive navigation ability. This suggests a new class of learning analytics: longitudinal measures of metacognitive skill, including calibration curves, independent problem-solving assessments, and reflection quality indices. The compass model also implies that AI tools should be transparent about their own pedagogical strategy, enabling educators and learners to make informed choices about when and how to engage them.

4.4. Implications for Practice, Policy, and Future Research

For EdTech developers. The five principles provide a concrete design checklist. Prompting, fading, pausing, process visibility, and calibration are all technically implementable in current AI systems. The barriers are not computational but cultural and commercial: an industry that has optimised for seamlessness, speed, and task completion must shift toward optimising for durable learning. This may require trade-offs; an AI that withholds immediate answers may be rated less favourably by users in the short term, even if it produces better learning outcomes in the long term. Public commitments to learning efficacy standards, analogous to nutritional labelling, could help align user expectations with educational goals.

For educators. Teachers and school leaders face an increasingly crowded marketplace of AI educational products. The compass metaphor offers a heuristic for selection: Does this tool make student thinking visible, or does it hide cognitive work behind a polished interface? Does it ask the student to plan, monitor, and reflect, or does it assume those functions on the student's behalf? Professional development should equip educators to evaluate AI tools not only on their content accuracy but on their metacognitive affordances. Moreover, teachers can adopt the five principles independently, using them to frame their own pedagogical use of AI in the classroom for example, by structuring activities that interleave AI assistance with student-generated planning and reflection.

For policy. At present, most regulatory and policy discussions around AI in education focus on data privacy, algorithmic bias, and content accuracy. These are important, but they leave unaddressed the cognitive consequences of tool design. This paper suggests that educational AI should be subject to cognitive efficacy standards analogous to those applied to medical devices: any tool that systematically offloads a core cognitive function should be required to demonstrate that it does not produce long-term dependence or atrophy in the offloaded capacity. This does not mean banning offloading; it means requiring evidence that the tool is designed to scaffold the development of internal capacity over time.

Future research. The evidence base on AI and metacognition remains nascent, and several priority research directions emerge from this synthesis. First, longitudinal studies are urgently needed to track metacognitive development in cohorts of learners who use AI tools with and without metacognitive scaffolds over months and years. Second, neuroimaging studies could investigate whether cognitive offloading to AI produces measurable changes in prefrontal and hippocampal function analogous to those observed in GPS users. Third, classroom-based design experiments could test the five principles as an integrated package, comparing learning outcomes, metacognitive skill growth, and AI dependence across different design configurations. Fourth, research should examine individual differences such as prior knowledge, executive function, and age that may moderate the

effects of AI offloading, informing the design of adaptive systems that calibrate the level of cognitive friction to the needs of the learner.

The compass model and its associated principles are a starting point, not an endpoint. As AI capabilities evolve, the specific forms that cognitive offloading takes will shift, and the design principles will need to be refined and extended. What remains constant is the underlying imperative: that educational technology, in whatever form it takes, should leave the learner with a stronger mind, not just a correct answer.

5. Limitations and Recommendations

5.1. Limitations

Nature of the review. This paper is an integrative literature review, not a systematic review or meta-analysis. While integrative reviews are well suited to synthesising heterogeneous and nascent bodies of knowledge (Torraco, 2005; Whitemore & Knafl, 2005), they carry inherent subjectivity. The cross-domain analogy between GPS-induced spatial disuse and AI-induced metacognitive disuse is a theoretical construct, not a directly observed phenomenon. The analogical reasoning that yields insight can also overextend, conflating distinct neural systems (hippocampus-dependent spatial navigation vs. prefrontal-dependent metacognition) and distinct forms of offloading. Readers should therefore regard the GPS-compass framework as a productive lens rather than a verified model, pending direct empirical testing.

State of the evidence on AI and metacognition. The evidence base linking generative AI specifically to metacognitive decline remains small and largely indirect. The Bastani et al. (2025) study, while methodologically rigorous, is a single investigation. Prather et al. (2024) and related computing-education research document metacognitive bypass but primarily through qualitative and correlational designs. Longitudinal studies tracking the same learners over months or years of AI use are absent from the literature. Consequently, the causal claim that habitual AI offloading produces durable metacognitive atrophy must be treated as a hypothesis grounded in precedent (the GPS literature) and cross-sectional data, rather than as an established fact.

Generalisability across AI tools and populations. Current AI-in-education research concentrates disproportionately on generative chatbots and code generators, predominantly in tertiary and secondary settings in high-income countries. The findings may not generalise to other AI formats (e.g., adaptive intelligent tutoring systems without generative fluency, or immersive learning environments), to younger learners whose metacognitive systems are at different developmental stages (Veenman et al., 2006), or to cultural contexts where collaborative and teacher-mediated uses of AI may buffer offloading effects. The five design principles, while evidence-informed, have not been tested as an integrated package across diverse learner populations, domains, or implementation models.

Unaddressed dimensions. This paper focuses narrowly on metacognition. It does not address other cognitively significant consequences of AI offloading, such as effects on motivation, creative problem-solving, epistemic trust, or the development of domain-specific expertise. It also sets aside critical sociotechnical questions algorithmic bias, data privacy, environmental costs, and the political economy of AI in education that are beyond its scope but essential to any comprehensive ethical framework.

5.2. Recommendations for Future Research

To move from a promising framework to a robust evidence base, the following research priorities are recommended:

1. **Longitudinal and experimental studies.** Researchers should track cohorts of learners who use AI tools with varying degrees of metacognitive scaffolding over extended periods (ideally, an academic year or longer). Outcome measures should include not only immediate task performance but independent problem-solving ability, calibration accuracy, and transfer

performance, measured both with and without AI assistance. Randomised controlled trials comparing AI tools that embed the five design principles against standard AI interfaces would test the compass model directly.

2. **Neural and cognitive process measures.** Cognitive offloading to AI should be studied using neuroimaging (fMRI, fNIRS) and process-tracing methods (eye tracking, log-file analysis) to examine whether prolonged AI use is associated with changes in prefrontal activity patterns during self-regulated learning tasks. Such studies could determine whether cognitive disuse atrophy is occurring at a neural level, analogous to the hippocampal findings in the GPS literature (Dahmani & Bohbot, 2017).
3. **Developmental and individual differences research.** The effects of AI offloading are unlikely to be uniform across all learners. Studies should investigate how prior knowledge, executive function, age, and metacognitive baseline skill moderate the impact of AI assistance. This would inform the design of adaptive AI systems that calibrate the level of cognitive friction to the learner's current capacity.
4. **Design-based implementation research.** The five principles require testing in authentic classroom settings, with teacher mediation and peer collaboration. Design-based research partnerships between universities, schools, and EdTech developers can refine the principles into practical, context-sensitive implementation protocols and generate professional development resources for educators.

5.3. Recommendations for Practice and Policy

For educators. Teachers should approach AI tools with a metacognitive lens, asking not only "Does this tool help students get the right answer?" but "Does this tool make student thinking visible, and does it require students to plan, monitor, and evaluate?" Classroom routines that interleave AI use with peer discussion, self-explanation, and teacher-led reflection can mitigate offloading risks even when the tool itself lacks robust scaffolding.

For EdTech developers. The five design principles prompting planning, fading feedback, embedding reflection, revealing process, and calibrating confidence should be adopted as default features rather than optional add-ons. Developers should also provide learning analytics dashboards that give teachers and learners insight into metacognitive growth over time, including calibration curves and reflection quality metrics. A commitment to long-term learning efficacy, not just short-term engagement and task completion, should guide product evaluation.

For policymakers and accreditation bodies. Regulatory frameworks for AI in education should extend beyond data privacy and bias to include cognitive efficacy standards. Accreditation criteria for educational AI tools could require evidence that the tool's design supports, rather than undermines, the development of independent cognitive skills. Funding agencies should prioritise research that evaluates the metacognitive consequences of AI in education, recognising that the most important outcome measure is not immediate performance but the capacity to learn without assistance.

5.4. Limitations of the Recommendations

Finally, the recommendations offered here are themselves bounded by the current state of knowledge. They assume a trajectory of AI development that may change rapidly. The design principles may prove insufficient in the face of increasingly autonomous AI agents that simulate empathy, manage long-term learning trajectories, or embed themselves invisibly into the learning environment. Continual re-evaluation of both the framework and its prescriptions will be necessary. What remains constant is the ethical and cognitive imperative that guides the compass model: educational technology, in whatever form it takes, should strengthen the learner's ability to navigate the world of ideas independently.

6. Conclusions

This paper opened with a student, Lena, who had arrived at a correct physics answer but could not explain her own thinking. Her situation was not a failure of intelligence or effort, but a symptom of a cognitive design problem: the AI tool she used had navigated the problem space so completely on her behalf that her inner compass had been left unused. Through an interdisciplinary synthesis of cognitive neuroscience, educational psychology, and early AI-in-education research, we have argued that Lena's experience is not an isolated anecdote but a predictable consequence of unexamined cognitive offloading and, crucially, that it is a consequence we can design against.

The central argument of this paper rests on three pillars, each drawn from a distinct research tradition yet reinforcing the others. The first pillar, from spatial navigation neuroscience, demonstrates that habitual reliance on GPS produces measurable declines in hippocampal-dependent cognitive mapping, a phenomenon termed cognitive disuse atrophy (Dahmani & Bohbot, 2017; Ishikawa et al., 2008; Maguire et al., 2006). The brain, faced with consistent external support, reallocates resources away from the unsupported function. The second pillar, from educational psychology, establishes that metacognition the planning, monitoring, and evaluation loop that governs self-regulated learning is among the most powerful predictors of academic achievement (Flavell, 1979; Hattie, 2009) and requires deliberate practice to develop (Azevedo & Gašević, 2019). The third pillar, from the nascent AI-in-education evidence base, reveals that current AI tools can bypass this metacognitive loop, producing immediate performance gains that mask impaired independent capability and inflated confidence (Bastani et al., 2025; Prather et al., 2024; Roelle et al., 2017).

By threading these three strands together through the metaphor of AI as a "GPS for thinking," we have named the risk and, equally importantly, pointed toward a remedy. The danger is not that AI assists learners assistance is its promise. The danger is that AI assists in a manner that renders the learner's own cognitive navigation redundant. The solution is to redesign AI partners as *compasses* rather than GPS devices: tools that provide orientation, demand engagement, and, through repeated use, strengthen the learner's internal sense of direction.

From this reconceptualisation, we derived five evidence-informed design principles. AI learning partners should (1) prompt learners to articulate a plan before offering assistance, (2) delay and fade feedback to preserve the demand for self-monitoring, (3) insert mandatory reflection pauses to activate post-task evaluation, (4) make their reasoning processes visible so learners can compare and calibrate, and (5) systematically calibrate learner confidence over time. These principles are neither technically prohibitive nor pedagogically radical; they build upon decades of research on self-explanation (Chi, 2000), feedback timing (Butler & Winne, 1995), desirable difficulties (Bjork & Bjork, 2011), and metacognitive scaffolding (Azevedo & Gašević, 2019; Wise & Hsiao, 2019). Their collective aim is to ensure that AI interactions leave behind not only correct answers but a more capable, self-aware learner.

The paper also acknowledged significant limitations. The evidence base on AI-induced metacognitive decline remains small, largely correlational, and concentrated in a few educational contexts. Longitudinal and experimental studies are urgently needed to test the causal mechanisms proposed here and to validate the five principles as an integrated design package. The compass model, while a productive conceptual framework, is an analogy; it must be refined and perhaps replaced as empirical understanding deepens. Furthermore, the paper did not address the broader sociotechnical and ethical dimensions of AI in education, which will intersect with cognitive design in important ways. Algorithmic bias, unequal access, privacy concerns, and the commercial pressures toward seamless, engaging interfaces all complicate the implementation of metacognitively responsible AI.

Yet limitations do not diminish urgency. AI is not waiting for careful research before embedding itself in classrooms; it is already there, with millions of students using it daily. Inaction in the face of this rapid deployment is itself a design choice one that defaults to tools that maximise short-term task completion at the potential cost of long-term cognitive development. The recommendations advanced here for researchers, developers, educators, and policymakers are intended as a starting

point for coordinated action. Longitudinal studies must track metacognitive outcomes over years, not weeks. Developers must treat metacognitive scaffolding as a core feature, not an afterthought. Teachers must be equipped to select and use AI tools with a metacognitive lens. Policymakers must expand regulatory attention from data and bias to cognitive efficacy, demanding evidence that educational AI leaves learners stronger, not more dependent.

The deeper significance of this argument reaches beyond the particulars of any AI tool. In a world increasingly saturated with cognitive prosthetics technologies that can write, reason, analyse, and decide on our behalf the most important educational outcome may be the capacity to navigate the world of ideas with an internal compass that remains, even when the screens go dark, unmistakably our own. This capacity is not a luxury, a soft skill, or an optional enrichment. It is the foundation on which lifelong learning, critical thinking, and intellectual agency are built. Losing it to unexamined convenience would be a quiet catastrophe.

We end where we began, with a student. Not Lena frozen at the question, but a different student a year later, in the same classroom, using the same AI this time redesigned. She types a problem, and the AI asks her to sketch her approach first. She wrestles with a hint, recalibrates her confidence, and after the task, pauses to reflect. When her teacher kneels beside her and asks how she solved it, she can say: *I started here, I got stuck there, I tried this, and I know what I'd do differently next time.* The AI has helped her, but it has not replaced her. It has been a compass, not a GPS. And the map she carries now is her own.

References

- Azevedo, R., & Gašević, D. (2019). Analyzing multimodal multichannel data about self-regulated learning with advanced learning technologies: Issues and challenges. *Computers in Human Behavior*, 96, 207–210. <https://doi.org/10.1016/j.chb.2019.03.025>
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakçı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281. <https://doi.org/10.3102/00346543065003245>
- Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5, pp. 161–238). Lawrence Erlbaum.
- Dahmani, L., & Bohbot, V. D. (2017). Habitual use of GPS negatively impacts spatial memory during self-guided navigation. *Scientific Reports*, 7, 41128. <https://doi.org/10.1038/srep41128>
- Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4), 271–280. <https://doi.org/10.1016/j.learninstruc.2011.08.003>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52(1), 45–56. <https://doi.org/10.1037/0003-066X.52.1.45>
- Hacker, D. J., Bol, L., & Bahbahani, K. (2008). Explaining calibration in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition and Learning*, 3(2), 101–121. <https://doi.org/10.1007/s11409-008-9021-5>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M. C., & Vedel, I. (2018). Mixed Methods Appraisal Tool (MMAT) version 2018. *Education for Information*, 34(4), 285–291. <https://doi.org/10.3233/EFI-180221>
- Ishikawa, T., Fujiwara, H., Imai, O., & Okabe, A. (2008). Wayfinding with a GPS-based mobile navigation system: A comparison with maps and direct experience. *Journal of Environmental Psychology*, 28(1), 74–82. <https://doi.org/10.1016/j.jenvp.2007.09.002>
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus*, 16(12), 1091–1101. <https://doi.org/10.1002/hipo.20233>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Prather, J., Reeves, B. N., Denny, P., Becker, B. A., Leinonen, J., Luxton-Reilly, A., Powell, G., Finnie-Ansley, J., & Savelka, J. (2024). The robots are coming: On the potential of ChatGPT for computing education. In *Proceedings of the 2024 Innovation and Technology in Computer Science Education V. 1 (ITiCSE 2024)*. Association for Computing Machinery. <https://doi.org/10.1145/3649217.3653560>
- Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist*, 38(1), 15–22. https://doi.org/10.1207/S15326985EP3801_3
- 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>
- Roelle, J., Schmidt, E. M., Buchau, A., & Berthold, K. (2017). Effects of informing learners about the dangers of cognitive offloading on their offloading behavior and achievement. *Journal of Educational Psychology*, 109(7), 971–987. <https://doi.org/10.1037/edu0000178>
- Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, 26(5), 521–562. https://doi.org/10.1207/s15516709cog2605_1
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4), 460–475. <https://doi.org/10.1006/ceps.1994.1033>
- Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351–371. <https://doi.org/10.1007/BF02212307>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, 45. <https://doi.org/10.1186/1471-2288-8-45>
- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356–367. <https://doi.org/10.1177/1534484305278283>
- Tyndall, J. (2010). AACODS checklist for appraising grey literature. Flinders University. <https://dSPACE.flinders.edu.au/xmlui/handle/2328/3326>
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- 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. <https://doi.org/10.1007/s11409-006-6893-0>
- Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, 52(5), 546–553. <https://doi.org/10.1111/j.1365-2648.2005.03621.x>

- Wise, A. F., & Hsiao, Y. T. (2019). Self-regulation in online discussions: Aligning data streams to model and support normative behaviors. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 220–229). Association for Computing Machinery. <https://doi.org/10.1145/3303772.3303810>
- Zheng, L., Li, X., & Chen, F. (2019). Effects of metacognitive prompts on learning outcomes in e-learning environments: A meta-analysis. *Journal of Computer Assisted Learning*, 35(2), 179–191. <https://doi.org/10.1111/jcal.12324>

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.