

Beyond Cognitive Offloading: What Must Remain Internal in Generative AI-Supported Learning?

A Cognitive and Neuroeducation-Informed Conceptual Analysis

Hui Li

ABSTRACT

Generative artificial intelligence (GenAI) can now perform cognitive activities such as explanation, synthesis, generation, evaluation, and reasoning on learners' behalf. The central educational question is therefore no longer only which tasks can be delegated to GenAI, but which cognitive resources must genuinely remain with the learner after learning. Existing research has addressed this issue through distinctions such as scaffolding versus substitution, different forms of cognitive offloading, the locus of GenAI participation, and task allocation. Building on this work, the present paper shifts the analytical focus to three internal cognitive resources that learners ultimately need to retain: internal knowledge, internal cognitive capability, and internal cognitive control. These correspond to what learners need to know, what they ultimately need to be able to do, and how they need to judge when to use, verify, or reject GenAI. Drawing on cognitive offloading, cognitive load theory, generative learning, the learning-performance distinction, and metacognitive theory, we argue that GenAI may reduce non-target processing and may temporarily perform some cognitive operations, but it should not replace the knowledge, capabilities, or regulatory control that the learning objective itself requires learners to develop. GenAI-supported learning should therefore begin with the learning objective, specify the internal resources that must be retained, determine how and when GenAI should intervene and be faded, and verify learning through retention, transfer, independent performance, and metacognitive calibration. The paper concludes with corresponding design principles and a research agenda, while treating neuroscience as an explanatory basis rather than direct evidence of neural effects caused by GenAI.

Keywords: generative AI; cognitive offloading; neuroeducation; cognitive load; generative learning; metacognition; learning versus performance

1. Introduction

Generative artificial intelligence (GenAI) is a form of artificial intelligence (AI) that changes not only how quickly learners access information, but also how cognitive work is distributed. Unlike earlier digital tools that mainly stored, retrieved, or organized information, GenAI can generate explanations, synthesize evidence, propose solutions, evaluate outcomes, and revise products. It can therefore provide personalized support and immediate feedback while also taking over cognitive processes that learners previously had to perform themselves. This makes a longstanding distinction especially important: completing a task with a tool is not the same as acquiring learning that can be retained and transferred. Soderstrom and Bjork (2015) differentiate between immediate performance during acquisition and long-term learning that supports later retention and transfer. The main educational issue posed by GenAI is therefore not simply what learners can do with AI, but what they can still do when similar support is reduced or unavailable.

Cognitive offloading is one reason why learners outsource part of their mental work to external tools. According to Risko and Gilbert (2016), it involves using actions in the body or environment to alter a task's information-processing requirements and reduce cognitive demand. Unaided cognition has limited capacity, whereas external resources can extend memory, computation, search, and organization. Cognitive load theory likewise shows that working memory is limited when dealing with novel information, while structured knowledge in long-term memory can greatly improve processing efficiency (Sweller et al., 2019). Reducing unnecessary processing is therefore not necessarily harmful to learning. The important issue is what learners still retain when part of the work is moved outside the learner's own cognitive system.

Recent GenAI research has already moved beyond a simple opposition between using and not using AI. Zhu et al. (2026) distinguish dependent offloading, in which AI takes over core thinking, from autonomous offloading, in which AI functions as a scaffold while the user retains cognitive agency. Cho (2026) analyzes the locus of generative work through Generation Partner, Generation Substitute, and Generation Catalyst roles. Lin and Al-Hada (2026) argue that the consequences of delegation depend on whether the outsourced operation is surface, intermediate, or deep relative to the focal learning objective. Lai et al. (2026) show empirically that offloading one apparently high-level operation can be productive when learners remain responsible for other learning-generative processes, whereas Wong and Qiu (2026) show that requiring learners to generate ideas before using ChatGPT can strengthen later independent creativity. In assessment, Chase and Galvin (2026) align GenAI decisions with learning outcomes, constructs, capabilities, and observable indicators. More recent conceptual work further narrows the available novelty space. Liubertaitė's (2026) Elicit-Formulate-Feedback-Organize-Reflect-Transfer (EFFORT-AI) module distinguishes target cognition from support cognition and makes sequencing, learner attempts, reflection, and fading central design variables. Vendrell and Johnston (2026) propose design principles for scaffolding critical thinking, including deliberate alternation between AI-free and AI-mediated phases. Du and Yuan (2026) frame the issue as epistemic dependence, asking whether reliance on AI preserves or displaces the epistemic work

through which judgment develops. These contributions make it difficult to claim novelty simply by stating that AI should scaffold rather than replace, that not all offloading is equal, that AI allocation should be based on learning objectives, or that learners sometimes need AI-free practice.

Existing research has already clarified the roles AI can play, the cognitive activities that may be delegated, and the conditions under which learner participation should be maintained. The present paper moves the analytical endpoint one step further: under conditions of pervasive external cognitive support, what must education ultimately ensure remains with the learner? The paper is therefore not primarily concerned with where cognition occurs in a single AI-supported episode, but with which learner-level cognitive resources a learning sequence should ultimately develop and preserve. We use internal cognitive resources as an analytical umbrella rather than a new psychological construct. The term refers to resources learners can access and deploy without equivalent cognitive substitution by GenAI when educational goals require independent use. This position does not reject distributed or extended cognition: tools can be legitimate parts of competent activity, but learners still need internal resources sufficient to understand, evaluate, and regulate those tools.

We distinguish three interdependent internal cognitive resources. Internal knowledge is representational and asks what the learner must know. Internal cognitive capability is operational and asks what the learner must ultimately be able to do. Internal cognitive control is regulatory and asks how the learner must be able to monitor, select, and allocate internal cognition and external support. The classification is based on cognitive function rather than presumed neural or modular boundaries. Motivation, affect, identity, and social agency remain important, but are treated here as adjacent or boundary constructs rather than additional internal-resource categories. The paper does not propose a new neurocognitive mechanism. Instead, it integrates established theories and emerging GenAI evidence into a target-relative, developmental framework and advances four testable propositions concerning knowledge, capability, control, and the verification of learning.

This article adopts a problem-driven conceptual approach that combines theory synthesis with model building (Jaakkola, 2020). Mature cognitive theories are used to establish mechanisms relevant to memory, processing, generative activity, and metacognitive regulation, while recent GenAI conceptual and empirical studies provide contrastive evidence, boundary conditions, and close-neighbour frameworks. The aim is not to provide an exhaustive or systematic review, but to integrate these literatures around a common learner-level endpoint and derive propositions for future empirical testing; neuroscience is used only as an explanatory constraint, not as evidence of direct neural effects of GenAI.

2. Why Cognitive Offloading Occurs

Cognitive offloading is not a learning failure, nor is it unique to the GenAI era. People have long used notes, calculators, search engines, navigation systems, and other external tools to do part of their thinking. Risko and Gilbert (2016) show that externalization can compensate for limited memory and processing capacity, reduce cognitive effort, and improve performance. Outsourcing

peripheral calculation, reminders, or information retrieval to AI can therefore indicate efficient use of cognitive resources rather than avoidance of learning.

Cognitive load theory also explains why external support may be useful. Working memory is limited in capacity and duration when processing new information, whereas structured knowledge in long-term memory makes processing more efficient (Sweller et al., 2019). Instruction should therefore minimize demands that are irrelevant to the learning objective and support the formation of target knowledge structures. Learners do not have to perform every part of a task on their own: when AI removes non-target processing, cognitive relief can be productive; when it removes processing needed to build the target schema, the same automation can be counterproductive (Cho, 2026).

Cognitive offloading is also a strategy-selection problem. Risko and Gilbert (2016) show that learners choose between internal and external strategies based on task difficulty, confidence, perceived reliability of the external resource, effort, and previous experience. Repeated success with external systems can also change perceptions of one's own capability and make future offloading more likely. This dynamic is particularly relevant to GenAI because delegation is inexpensive and its outputs are rapid, fluent, and easy to adopt directly.

Skulmowski's (2023) digital externalization model further suggests that information entering working memory may either become part of biological long-term memory or remain in an external system; this allocation is affected by cognitive load, perceived memory capacity, metacognition, and information value. Externalization can result in only the gist or a retrieval pointer. This may be efficient when precise retention is not a learning objective, but it becomes important when the externalized content must later be understood, transferred, or evaluated. Taken together, these perspectives explain why offloading is attractive and can be effective, while also showing that its educational value is contingent on what later cognition requires.

3. What Must Remain Internal?

We structure the question of what must remain internal around three related questions: What must learners know? What do they ultimately need to be able to do? What do they need to regulate? These correspond to internal knowledge, internal cognitive capability, and internal cognitive control: representational resources, object-level operational resources, and meta-level regulatory resources. Here, internal does not mean AI-free; it refers to resources owned by the learner that remain available for understanding, action, and regulation even during human-AI collaboration. These categories are not separate faculties and do not cover every educational outcome; rather, they distinguish three learner-level functions that strong AI-assisted performance can obscure.

Table 1

Internal cognitive resources in GenAI-supported learning

Resource	Core question	Primary theoretical anchors	Design and evaluation criterion
Internal knowledge (Representational resource)	What must the learner know without equivalent AI substitution?	Cognitive load; long-term memory; digital externalization	Identify prerequisite knowledge; verify retention, retrieval, and comprehension after support is reduced.
Internal cognitive capability (Operational resource)	What must the learner ultimately be able to do?	Generative learning; constructive engagement; generation effect	Allow AI to scaffold or model, but preserve practice of target operations; verify transfer and independent performance.
Internal cognitive control (Regulatory resource)	What must the learner be able to regulate?	Metacognition; cognitive offloading; self-regulated learning	Preserve monitoring, strategy selection, verification, and override; verify calibration and appropriate delegation.

Note. Verification under reduced or absent AI support is warranted where independent execution is part of the intended learning outcome.

3.1 Internal Knowledge: What Must Learners Know?

Internal knowledge is knowledge that learners themselves should be able to access and use later for understanding, reasoning, transfer, evaluation, or action. This does not mean that learners should learn everything GenAI can provide. The criterion is not whether AI can retrieve the information, but whether learners can still engage in later cognitive activity without having that knowledge internally available. When knowledge is a precondition for further cognition, external accessibility cannot automatically be treated as internal mastery.

Cognitive load theory supports this distinction. Sweller et al. (2019) stress that long-term memory is more than a storehouse of facts: structured knowledge changes how subsequent tasks are processed. Elements that novices must handle separately can later be treated as an integrated schema. In AI-mediated learning, fluent access to external answers can therefore conceal the absence of an internal knowledge base.

Skulmowski (2023) further distinguishes biological long-term memory from external digital storage. Cognitive offloading may leave learners with only the gist of information or a pointer to where it can be found. This may be entirely appropriate for peripheral facts or reference details. The problem arises when the externalized material consists of core concepts, relations, or domain structures needed to understand new information or evaluate external answers. Repeatedly asking GenAI for a definition of working memory, for example, is not equivalent to having a schema of cognitive load that supports reasoning in a new situation.

Emerging evidence is consistent with this distinction. In a preregistered randomized secondary-school experiment, Kreijkes et al. (2026) found that note-taking, either alone or combined with a large language model (LLM), produced better retention and comprehension than LLM use alone. Students nevertheless tended to perceive the LLM as easier and more helpful. One explanation is that note-taking required learners to identify, paraphrase, and organize information, whereas LLM use could reduce some of this active processing. At the same

time, students also used the LLM to ask questions and explore background information. The finding therefore does not imply that LLMs are unhelpful; it shows that external support can aid initial understanding and exploration without demonstrating that focal knowledge has been internally retained.

A neuroeducation-informed perspective also cautions against treating human memory and AI generation as interchangeable. Rolls (2024) shows that hippocampal-neocortical retrieval and token-based generation differ fundamentally in their computational organization. This does not show that GenAI use changes the hippocampus, nor does it justify direct neural claims. It simply means that an external model's ability to generate information is not equivalent to the learner having formed a retrievable internal representation of that knowledge.

Proposition 1 - Internal Knowledge Requirement. *When particular knowledge is necessary for subsequent understanding, transfer, evaluation, or independent performance, external accessibility through GenAI should not be assumed to be functionally equivalent to that knowledge being internally available to the learner. GenAI-supported instruction should distinguish information that may be productively externalized from knowledge that must be sufficiently established in long-term memory to support the intended learning outcome.*

3.2 Internal Cognitive Capability: What Must Learners Be Able to Do?

Possessing knowledge is not the same as possessing capability. Education also concerns whether learners can use knowledge to retrieve, explain, generate, compare, reason, evaluate evidence, diagnose errors, and revise. We define internal cognitive capability as the object-level operations learners are ultimately expected to execute without equivalent external substitution. A learner may know the criteria of a good argument but still be unable to construct or evaluate one independently; conversely, a learner may produce a high-quality argument with GenAI while the relevant capability remains largely external. Evaluation here refers to object-level judgment, such as deciding whether evidence supports a claim. Meta-level judgments about whether one can evaluate adequately, whether AI should be consulted, or whether an AI judgment should be accepted or overridden belong to internal cognitive control.

Generative learning theory explains why cognitive processes performed by learners can matter for learning. Wittrock (1992) describes learning as the active construction of relationships among new information, previous knowledge, experience, and purpose rather than passive recording. Chi (2009) likewise distinguishes simple manipulation of given information from constructive engagement that produces inferences, explanations, and integrations. Generation-effect research also indicates that self-generated information tends to be remembered better than information that is merely read (Bertsch et al., 2007). These traditions do not require learners to produce every product themselves; they show that some generative operations can themselves function as mechanisms of learning.

GenAI makes it more difficult to see who is actually performing those operations. Cho (2026) argues that a generated product no longer shows that the learner carried out the underlying generative work and distinguishes AI as a Generation Partner, Generation Substitute, or Generation Catalyst. Removing cognitive work is therefore not inherently beneficial or harmful: the same automation may eliminate unnecessary processing in one task but remove processing

required to construct the target schema in another. The present paper builds on this point by shifting the endpoint from who performs an operation in a single episode to which operations must ultimately become executable learner capabilities.

The issue is not whether every episode is completely independent, but whether the instructional sequence gives learners opportunities to learn, practice, refine, and finally perform the target operation. Carrying out an operation once, however, does not show that a durable capability has formed. Internal cognitive capability is therefore a developmental outcome built through support, practice, and increasing independence.

Recent GenAI studies reflect this logic. Wong and Qiu (2026) required one group to generate their own ideas before working with ChatGPT. Although unrestricted ChatGPT use improved immediate creative performance, its advantage disappeared when AI was removed; the regulated group later showed stronger independent creativity. Process analyses reported that using ChatGPT to improve self-generated ideas was related to later originality. When independent idea generation is itself a desired outcome, preserving learner generation before AI intervention can therefore support capability development.

Lai et al. (2026), however, show why the claim that ‘generation must remain human’ is too simple. In English as a foreign language (EFL) writing feedback, the condition in which GenAI drafted an essay and learners evaluated and revised it produced the strongest writing development among the tested designs. AI-generated text served as material for learner comparison, evaluative judgment, diagnosis, explanation, and revision. Offloading drafting did not eliminate learning-generative activity; it shifted that activity toward capabilities emphasized by the task design. The implication is not that drafting is unimportant, but that the acceptability of offloading depends on the capability the instructional sequence is trying to develop.

Programming research supports the same distinction. Liu et al. (2026) identify a separation between Domain Mastery and Tool Mastery in AI-assisted programming: learners may use AI efficiently without corresponding independent conceptual understanding, explanation, evaluation, or verification. Mastery of an AI-mediated workflow should therefore not be assumed to be the same as mastery of the underlying cognitive capability.

Proposition 2 - Internal Cognitive Capability Requirement. *The educational appropriateness of delegating a cognitive operation to GenAI depends on whether independent execution of that operation is an intended learner capability. GenAI may successfully carry out an operation when its output allows learners to practice other target capabilities, yet repeated substitution becomes problematic when it reduces opportunities to acquire and refine that capability through independent execution of the target operation.*

3.3 Internal Cognitive Control: What Must Learners Be Able to Regulate?

Having knowledge and capability does not necessarily lead to adaptive use of GenAI. A learner who can think independently may still revert to AI whenever effort is required, whereas another learner with limited expertise may seek support strategically, check the output, and reduce assistance as competence grows. The third resource is therefore internal cognitive control: monitoring one’s understanding, deciding when to seek help, choosing internal or external strategies, evaluating AI output, and revising or overriding it when necessary. We use the term as

an umbrella for metacognitive monitoring and regulation of delegation, rather than for the narrower executive-function construct of cognitive control.

This layer is grounded in metacognitive theory. Nelson and Narens (1990) describe metacognition as a monitoring-control relationship between an object level and a meta level: the meta level receives information about current cognitive performance and regulates object-level processing in response. Fleur et al. (2021) likewise emphasize metacognitive knowledge, monitoring, and control in learning. In learning with GenAI, decisions about whether to consult, trust, verify, or reject AI are themselves meta-level regulatory decisions.

Risko and Gilbert's (2016) account of cognitive offloading also shows that learners choose between internal and external strategies based on confidence, effort, previous experience, and external reliability. Offloading can reshape later judgments as well: repeated success with an external system can increase preference for the external route. Because GenAI is fast, broad, and linguistically fluent, this feedback loop can occur even when outputs are incomplete or incorrect.

With GenAI, it is also possible to outsource not only execution but governance. Zhu et al. (2026) define cognitive agency transfer as the gradual relinquishing of decision-making authority over what information matters, how evidence should be assessed, and what conclusions should be drawn. They distinguish this from trust: a learner can rely on AI while retaining control through verification and override, or can defer governance without fully trusting the system. In their three-wave study, dependent offloading was associated with greater cognitive agency transfer, while metacognitive monitoring attenuated that pathway. The evidence is correlational, but it shows why educationally adaptive offloading involves more than tool-use skill. Du and Yuan (2026) reach a complementary conclusion from an epistemic perspective: the problem is not reliance itself, but whether AI-mediated reliance preserves or replaces the epistemic work through which learners develop judgment.

This is not meant to introduce cognitive governance as a new construct. Monitoring, strategy selection, verification, override, and regulation of delegation are established metacognitive processes that become especially visible in GenAI contexts. Cho's (2026) idea of metacognitive displacement similarly implies that AI can assist monitoring and control, yet scaffolding can turn into substitution when the system begins to replace the process being evaluated or monitored. For novices, early external orchestration by teachers or structured tasks may still be necessary because limited prior knowledge constrains error detection. Cognitive control can therefore be scaffolded, but it should progressively become the learner's own capacity to regulate future cognitive delegation.

Neuroeducation research supports treating metacognitive control as developable while providing no basis for direct claims about GenAI's neural effects. Fleur et al. (2021) show that metacognitive processes have identifiable neurocognitive correlates and change across childhood and adolescence, while educational and laboratory operationalizations are not fully aligned. It is therefore reasonable to treat metacognitive control as something that develops with knowledge and expertise, but not to argue that GenAI directly strengthens or weakens particular brain regions.

Proposition 3 - Internal Cognitive Control Requirement. *As cognitive processes become increasingly distributable between learners and GenAI, adaptive GenAI-supported learning requires*

learners to develop the metacognitive capacity to monitor understanding, decide when external support is needed, evaluate and override AI output, and regulate which cognitive operations are delegated. Such control may be externally scaffolded at first, especially for novices, but instruction should gradually enable the learner to govern future cognitive delegation rather than routinely delegating that governance to AI.

4. A Target-Relative Developmental Framework for GenAI-Supported Learning

The three internal resources shift the focus from how tasks are divided to what learners ultimately develop. A conventional allocation question asks which part of a task the learner should perform and which part AI should perform. A developmental question first specifies the knowledge, capability, and control learners should possess at the end of learning, and then works backward to determine how AI should intervene. The same AI action can therefore be productive, neutral, or substitutive depending on the learning objective, prior knowledge, timing of support, and whether responsibility is progressively returned to the learner.

The resulting design logic can be stated as a sequence: learning objective -> required internal resource -> AI role and allocation -> scaffolding and fading -> learner-level verification. The five-part sequence specifies the overall design logic, whereas the four considerations below explain how that logic should be applied in instructional decisions. This sequence does not replace existing frameworks; it connects them around a common endpoint. Cho's (2026) Partner-Substitute-Catalyst typology helps describe the role AI occupies in a learning episode. Lin and Al-Hada's (2026) differentiated framework helps determine whether delegated processing is surface, intermediate, or deep relative to the focal learning objective. Zhu et al. (2026) distinguish between autonomous and dependent forms of offloading. Chase and Galvin (2026) show how learning outcomes, constructs, capabilities, and observable indicators can be unpacked to make assessment-valid decisions. Liubertaitė (2026) explicitly describes phase-sensitive sequencing and preservation of target cognition, whereas Vendrell and Johnston (2026) show how AI-free and AI-mediated stages can be intentionally combined to scaffold critical thinking. Du and Yuan (2026) add an epistemic criterion by asking whether reliance preserves the work through which judgment develops. The internal-resource account complements these views by defining the learner-level endpoint in representational, operational, and regulatory terms: which resources should ultimately be learner-available, and how can an instructional sequence build rather than merely display them?

The first step is to define the educational endpoint more specifically than task completion. If the learning objective is access to factual information, a high degree of externalization may be acceptable. If the learning objective is a schema needed for later reasoning, relevant knowledge has to be established internally. If the learning objective is independent proof construction, source evaluation, argumentation, or debugging, the corresponding operation should eventually become a learner capability. If the learning objective also includes professional judgment or responsible AI use, learners need to gain control over when to delegate, what to verify, and when

to override AI. Knowledge, capability, and control targets may therefore occur within the same task.

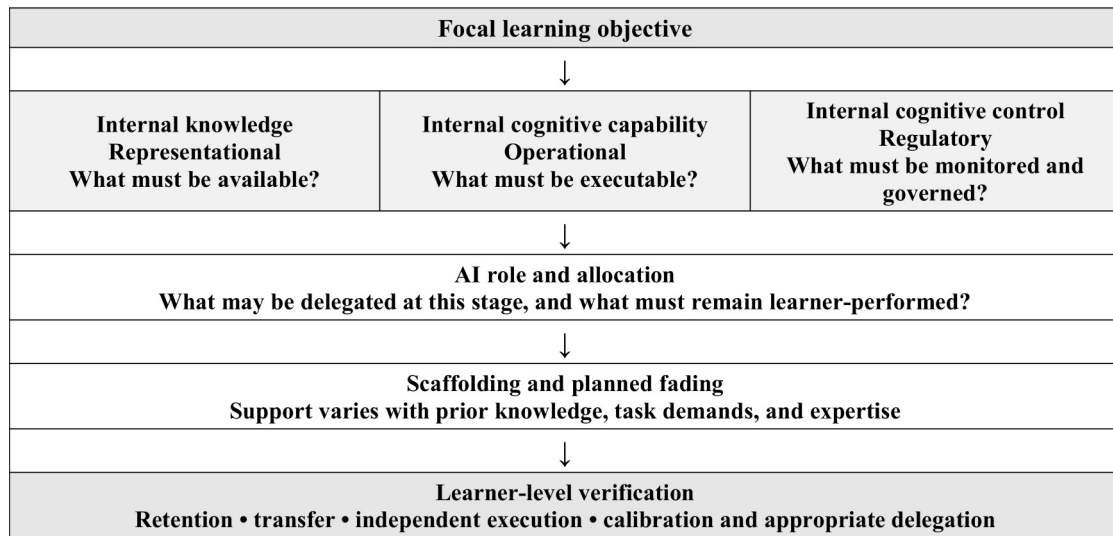
The second step is to understand AI support functionally rather than through surface labels. Generation, evaluation, and feedback are not inherently safe or unsafe to offload. In Lai et al. (2026), a draft generated by AI became an object for human evaluation and revision. In Wong and Qiu (2026), AI generation before learner ideation risked replacing the very capability under development. According to Cho (2026), AI output can function as a Substitute when it replaces target generative work, but as a Catalyst when the product becomes the focus of additional learner reasoning. The same technology, and even the same nominal operation, must therefore be judged in relation to the internal resource targeted by instruction.

The third step is to view AI assistance as scaffolding that must evolve over the course of learning. Initial explanations, examples, guided prompts, or AI-generated representations can help novices overcome excessive cognitive demands. As knowledge and capability increase, the same support may become redundant or substitutive. Planned fading, independent practice, and transfer tasks should therefore form part of GenAI-supported learning. Lai et al. (2026) recommend both scaffolded and independent practice, while Liu et al. (2026) propose critique-the-AI activities, planned fading, and verification journals as practices that require further testing.

The fourth step is to distinguish delegation of execution from delegation of regulatory authority. Even when AI performs an operation, learners can still decide why it is delegated, how the returned output should be evaluated, and what should happen next. Zhu et al.'s (2026) autonomous offloading reflects this pattern, whereas dependent offloading more closely resembles direct acceptance of AI-generated structures and conclusions. Instruction can therefore allow substantial task-level outsourcing while preserving learner responsibility for epistemic and strategic decisions.

Accordingly, what must remain internal does not mean that educated cognition should be isolated from tools. Expertise has always included appropriate use of external representations, collaborators, reference systems, and technologies. The claim is instead that external support does not eliminate the importance of learner-owned resources, because those resources make tool outputs understandable, evaluable, and usable. Internal knowledge, capability, and control should therefore be defined relative to learning objectives rather than treated as a fixed inventory that technology must never touch.

Figure 1
Target-relative developmental logic for GenAI-supported learning



Note. The three internal-resource classes are functional categories rather than independent faculties. Boundary conditions include learner age and expertise, prior knowledge, task and domain, AI reliability and availability, and whether independent execution is part of the intended learning outcome. Arrows indicate the proposed instructional design sequence rather than established causal pathways.

5. Evaluating Learning Beyond AI-Assisted Performance

If education is focused on developing cognitive resources that belong to learners, assessment must distinguish AI-assisted performance from learning that has actually occurred. Soderstrom and Bjork (2015) show that short-term performance may differ from long-term retention and transfer: conditions that make practice easier can improve performance without producing durable learning. Salomon, Perkins, and Globerson (1991) similarly distinguish between effects with technology and effects that persist after technology use. This distinction is especially important with GenAI because it can improve product quality, fluency, and speed without demonstrating that the corresponding knowledge or capability belongs to the learner.

Wong and Qiu (2026) provide a straightforward example. Unrestricted ChatGPT use increased creativity on the AI-supported task, but this advantage was not maintained when learners later worked without ChatGPT. The regulated ‘think first, ChatGPT later’ condition showed the reverse profile: less immediate gain but stronger later independent creativity. Zhu et al. (2026) likewise found that dependent and autonomous offloading were associated with similar immediate benefits despite divergent downstream perceived cognitive correlates. Liu et al. (2026) observed weakened metacognitive calibration in which perceived readiness could exceed independent capability. Together, these findings show why an AI-assisted artifact is an incomplete indicator of learning.

Evaluation should therefore be matched to the targeted internal resource. Internal knowledge can be examined through delayed retention, recall, comprehension, and retrieval of prerequisite

knowledge without equivalent AI support. Internal cognitive capability can be examined through transfer to novel problems, independent generation or evaluation, and performance after support is reduced. Internal cognitive control requires measures of metacognitive calibration, strategy selection, verification, error detection, and appropriate decisions about when to use or withhold AI. AI-assisted product quality can remain a performance indicator, but it cannot substitute for these learner-level measures.

The same logic applies to assessment validity. Chase and Galvin (2026) argue that educators should identify the constructs and capabilities an assessment is intended to represent before deciding which forms of GenAI use are compatible with those inferences. The present framework extends this principle beyond formal assessment: if we claim that learning has occurred, the targeted knowledge, capability, or control should be observable at the learner level rather than inferred only from a human-AI product.

Proposition 4 - Internal-Resource Verification. *Claims that GenAI enhances learning should not rely solely on performance observed while AI support is present. Where independent execution is part of the intended learning outcome, evaluation should include outcomes that test whether target knowledge, cognitive capabilities, and regulatory control remain available when equivalent AI support is reduced or absent, using measures such as delayed retention, transfer, independent performance, error detection, strategy selection, and metacognitive calibration.*

6. Neuroeducation-Informed Implications and Research Agenda

Neuroeducation can provide explanatory constraints for this framework, but should not be used to infer direct neural consequences of GenAI. Current evidence does not support educational claims such as GenAI offloading reducing hippocampal activity or weakening particular prefrontal networks. We use neuroscience only to support three more modest points: working memory is limited, long-term memory has important biological and cognitive functions, and metacognitive regulation depends on developable cognitive systems. Rolls (2024) cautions against treating human memory and generative AI as computationally interchangeable, while Fleur et al. (2021) show both the promise and the operational differences involved in linking educational metacognition with cognitive neuroscience.

Uden and Hwang's (2026) Lifelong Learning, Engagement, Active Processing, Reflection, and Neuro-based Design (LEARN) framework provides a similar precedent: neuroplasticity, effortful cognition, metacognitive regulation, motivation, and socio-emotional engagement are used to explain instructional design rather than treated as demonstrated neural outcomes of the framework. We adopt the same position. Neuroscience helps explain why durable memory, meaningful cognitive effort, and metacognitive regulation matter educationally; causal effects of GenAI on neural systems remain questions for direct empirical study.

For instruction, GenAI integration should begin by asking which learner resources must ultimately be retained rather than how much AI should be permitted. Educators should first identify the knowledge needed for later reasoning, the operations learners must eventually perform themselves, and the regulatory judgments they must master. AI can then be used to reduce non-target processing, provide models, generate comparison cases, or create material for

critique. Its educational value should be evaluated by whether it facilitates development of the target internal resources, not merely by the quality of the answer generated.

Progressive transfer of responsibility is also a key part of the framework. When novices lack prior knowledge or face excessive task demands, initial assistance can be substantial. As expertise increases, learners should receive more opportunities to retrieve, generate, evaluate, and make decisions with less support. AI-free checkpoints, planned fading, oral explanation, independent transfer tasks, and post-AI verification can make learner-owned resources visible. The aim is not to exclude AI, but to create diagnostic moments in which genuine learning can be observed.

Future studies should therefore go beyond simple AI-versus-no-AI comparisons and test different support sequences. For example, studies could compare unrestricted AI use, scaffolded AI use that preserves important learner operations, and scaffolded use followed by fading. Outcomes should be assessed both during GenAI-supported activity and after support is reduced, with delayed tests when durable learning is claimed. These designs can show whether immediate product quality, retention, transfer, and metacognitive calibration follow different trajectories.

Process measurement is a second priority. Interaction logs may reveal whether learners request complete answers, hints, critiques, alternatives, or verification, but these actions should not be treated as equivalent to learning. Think-aloud protocols, confidence judgments, error-detection tasks, revision traces, and explanations of why a particular operation was delegated can help distinguish strategic offloading from habitual dependence. Zhu et al. (2026) provide initial measures of dependent and autonomous offloading and cognitive agency transfer, but stronger psychometric validation and longitudinal work are still required; the verification and fading mechanisms suggested by Liu et al. (2026) likewise need testing across contexts.

Developmental and disciplinary variation is a third priority. What must remain internal changes with age, expertise, and domain. Beginners often need more background knowledge to judge external output, whereas experts may reasonably externalize routine operations and devote attention to higher-level analysis. The transferability of metacognitive control may also increase as expertise develops (Fleur et al., 2021). Future research should therefore specify the learning objective, assumptions about prior knowledge, the operation entrusted to AI, the internal resource expected to remain learner-owned, and the reduced- or no-AI condition used to verify it.

The framework also reframes AI literacy. Effective prompting and understanding system limitations are clearly important, yet an internal-resource approach emphasizes three additional questions: Is the learner sufficiently knowledgeable to challenge AI output? Can the learner act when AI is unavailable or wrong? Is the learner able to determine what should be delegated in the first place? AI literacy should therefore be judged not only on the basis of tool proficiency, but also by whether tool use remains subordinate to the learner's educational and epistemic goals.

Several limitations characterize this account. First, the paper is a problem-driven theoretical synthesis rather than an exhaustive review and should not be read as a comprehensive map of GenAI learning research. Second, internal knowledge, internal cognitive capability, and internal cognitive control are analytical categories and do not exclude the importance of motivation, affect, social interaction, or identity. Third, the propositions need to be empirically tested across

disciplines, age groups, task structures, and AI systems. Fourth, what should remain internal is always relative to the learning objective: in some authentic professional tasks, sustained human-AI collaboration is itself part of competent performance, so reduced- or no-AI verification is necessary only when the intended learning outcome includes independent execution. Finally, GenAI capabilities are evolving quickly, so particular applications of the framework will need continued testing as tools and educational settings change.

7. Conclusion

Generative AI takes cognitive offloading beyond memory and calculation to explanation, synthesis, generation, evaluation, and judgment. The problem is not that this externalization is inherently harmful, but that AI-assisted performance is becoming increasingly weak evidence of what the learner has actually learned. Existing research has already clarified differences in AI roles, task types, timing, scaffolding, offloading modes, and epistemic dependence. Building on that work, the present paper shifts the endpoint to three resources learners ultimately need to retain: internal knowledge, internal cognitive capability, and internal cognitive control.

The main argument is simple: the learning objective determines what must remain with the learner. Knowledge that supports later learning should be internally available; cognitive activities that learners are expected to perform independently must eventually become their own capabilities; and even when AI performs part of the work, decisions about when to delegate, how to verify, and when to reject AI should progressively remain with the learner. GenAI can therefore be used extensively, but its educational value should be judged by whether it helps develop and verify these internal resources. The question is not simply how much cognition can be outsourced to AI, but what learners should still be able to know, do, and regulate for themselves after learning.

References

- Bertsch, S., Pesta, B. J., Wiscott, R., & McDaniel, M. A. (2007). The generation effect: A meta-analytic review. *Memory & Cognition*, 35(2), 201–210. <https://doi.org/10.3758/BF03193441>
- Chase, A.-M., & Galvin, K. (2026). Thinking to learn: Managing the risks of outsourcing to GenAI. *Assessment & Evaluation in Higher Education*. Advance online publication. <https://doi.org/10.1080/02602938.2026.2620055>
- Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. <https://doi.org/10.1111/j.1756-8765.2008.01005.x>
- Cho, H. (2026). Relocating the locus of generation: The AI-GLM framework for generative learning in AI-mediated mathematics education. *Educational Psychology Review*, 38, 116. <https://doi.org/10.1007/s10648-026-10210-y>
- Du, Y., & Yuan, Y. (2026). Epistemic dependence in AI-mediated learning. *AI & Society*. Advance online publication. <https://doi.org/10.1007/s00146-026-03294-1>
- Fleur, D. S., Bredeweg, B., & van den Bos, W. (2021). Metacognition: Ideas and insights from neuro- and educational sciences. *npj Science of Learning*, 6, 13. <https://doi.org/10.1038/s41539-021-00089-5>
- Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1–2), 18–26. <https://doi.org/10.1007/s13162-020-00161-0>
- Kreijkes, P., Kewenig, V., Kuvalja, M., Lee, M., Hofman, J. M., Vitello, S., Sellen, A., Rintel, S., Goldstein, D. G., Rothschild, D., Tankelevitch, L., & Oates, T. (2026). Effects of LLM use and note-taking on reading comprehension and memory: A randomised experiment in secondary schools. *Computers & Education*, 243, 105514. <https://doi.org/10.1016/j.compedu.2025.105514>
- Lai, C., Pan, M., Guo, K., & Cui, Y. (2026). What task to offload to GenAI in the writing feedback process? Effects of task-offloading approaches on EFL learners' writing skill development. *Computers & Education*, 252, 105675. <https://doi.org/10.1016/j.compedu.2026.105675>
- Lin, J., & Al-Hada, N. M. (2026). The critical-thinking paradox in generative AI-integrated learning: Distinguishing efficiency from cognitive depth: A differentiated framework and testable propositions. *Frontiers in Psychology*, 17, 1906070. <https://doi.org/10.3389/fpsyg.2026.1906070>
- Liu, D., Fan, G., & Pan, L. (2026). Tool, tutor, or crutch? A grounded theory of cognitive scaffolding and offloading in AI-assisted programming education. *International Journal of STEM Education*, 13, 10. <https://doi.org/10.1186/s40594-025-00592-w>
- Liubertaitė, A. D. (2026). EFFORT-AI: An effort-preserving module for cognitively aligned human-AI learning. *Frontiers in Education*, 11, 1849821. <https://doi.org/10.3389/educ.2026.1849821>
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation*, 26, 125–173. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5)
- 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>
- Rolls, E. T. (2024). The memory systems of the human brain and generative artificial intelligence. *Heliyon*, 10, e31965. <https://doi.org/10.1016/j.heliyon.2024.e31965>
- Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, 20(3), 2–9. <https://doi.org/10.3102/0013189X020003002>
- Skulmowski, A. (2023). The cognitive architecture of digital externalization. *Educational Psychology Review*, 35, 101. <https://doi.org/10.1007/s10648-023-09818-1>
- Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10(2), 176–199. <https://doi.org/10.1177/1745691615569000>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31, 261–292. <https://doi.org/10.1007/s10648-019-09465-5>

- Uden, L., & Hwang, G.-J. (2026). The LEARN framework for responsible use of generative AI in education: A neuroscience-informed model for problem-based learning. *Journal of Computers in Education*. Advance online publication. <https://doi.org/10.1007/s40692-026-00398-x>
- Vendrell, M., & Johnston, S.-K. (2026). Scaffolding critical thinking with generative AI: Design principles for integrating large language models in higher education. *Computers and Education: Artificial Intelligence*, 10, 100572. <https://doi.org/10.1016/j.caeai.2026.100572>
- Wittrock, M. C. (1992). Generative learning processes of the brain. *Educational Psychologist*, 27(4), 531–541. https://doi.org/10.1207/s15326985ep2704_8
- Wong, S. S. H., & Qiu, S. X. (2026). Think first, ChatGPT later: Guiding human-AI collaboration for learning gains in independent human creativity. *Educational Psychology Review*, 38, 45. <https://doi.org/10.1007/s10648-026-10118-7>
- Zhu, Q., Li, X., Dong, Y., Chang, P., & Fan, M. (2026). Not all cognitive offloading is equal: Distinguishing dependent and autonomous offloading to generative AI. *Frontiers in Psychology*, 17, 1878629. <https://doi.org/10.3389/fpsyg.2026.1878629>