METATIDE: BRINGING EVOLVING SEEDS INTO MULTIMODAL REASONING VIA METACOGNITIVE INTERFLOW

Anonymous authors
Paper under double-blind review

ABSTRACT

Multimodal reasoning increasingly combines visual evidence with multi-step inference, while metacognitive regulation helps models assess and adjust an ongoing solution. A central challenge is how the understanding gained from these judgments and adjustments can persist and guide future problems. We introduce `MetaTide`, a framework that brings evolving metacognitive seeds into current reasoning through *Metacognitive Interflow*. Stored in a *Metacognitive Undercurrent*, these seeds capture transferable relations among cognitive conditions, regulatory choices, and observed effects. Metacognitive Interflow matches the current Meta-State to an applicable seed and translates it into second-order control, which guides task-specific observation and inference in *Surface Reasoning*. In this way, seeds shape how the model observes, evaluates, and adjusts its reasoning in a new task. The outcomes of these actions refine its ongoing assessment, and completed solutions are distilled into new or consolidated seeds. This reciprocal flow allows prior self-understanding to shape present reasoning while present reasoning renews the basis for future regulation. Experiments across five multimodal benchmarks show gains of $+29.03$ and $+22.58$ percentage points over `ReasoningBank` on MATH-Vision and MuirBench, respectively. *Code is available at MetaTide.*

1 INTRODUCTION

Multimodal large language models have evolved from connecting visual perception with language generation to reasoning over images and text (Alayrac et al., 2022; Li et al., 2023; Liu et al., 2023). Intermediate rationales can incorporate visual evidence (Zhang et al., 2023), while visual search and focused inspection let models return to image regions relevant to a question (Wu & Xie, 2024; Shao et al., 2024). Visual scratchpads further allow intermediate visual representations to participate in inference (Hu et al., 2024). Together, these developments have expanded what a multimodal solver can observe, represent, and reason through.

However, richer visual evidence and longer reasoning chains do not by themselves tell a model whether its current judgment is warranted, which part of its process needs revision, or when it should stop (Von Zahn et al., 2025; Tankelevitch et al., 2024). To address this gap, recent work has turned to multimodal metacognition: the model assesses how its observation and reasoning are unfolding and uses that assessment to guide what happens next (Yin et al., 2026). This can help it decide when to inspect an image again, revise an inference, continue reasoning, or stop. For example, Metacognitive Harness uses self-assessment to guide retries (Cao et al., 2026a), while Reflect to Inform strengthens visual re-examination during reasoning (Lv et al., 2026).

Despite helping regulate an ongoing solution, these approaches leave a deeper question unanswered: what has the model learned about its own cognition that can guide a future problem? A visual check may reveal why an inference seemed convincing, which uncertainty it missed, and whether changing course actually helped (Flavell, 1979; Nelson, 1990). Factual memory preserves knowledge of the task world (Lewis et al., 2020; Bo et al., 2026), and first-order experience preserves ways of completing tasks (Wang et al., 2024b); metacognitive reflection instead reveals how the model judges, when it needs to adjust, and whether those adjustments worked (Flavell, 1979; Zhuang et al., 2026). For this understanding to guide later problems, it must be distilled from completed attempts into transferable seeds. What is needed, then, is a substrate in which such seeds persist until a fitting cognitive state awakens them.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

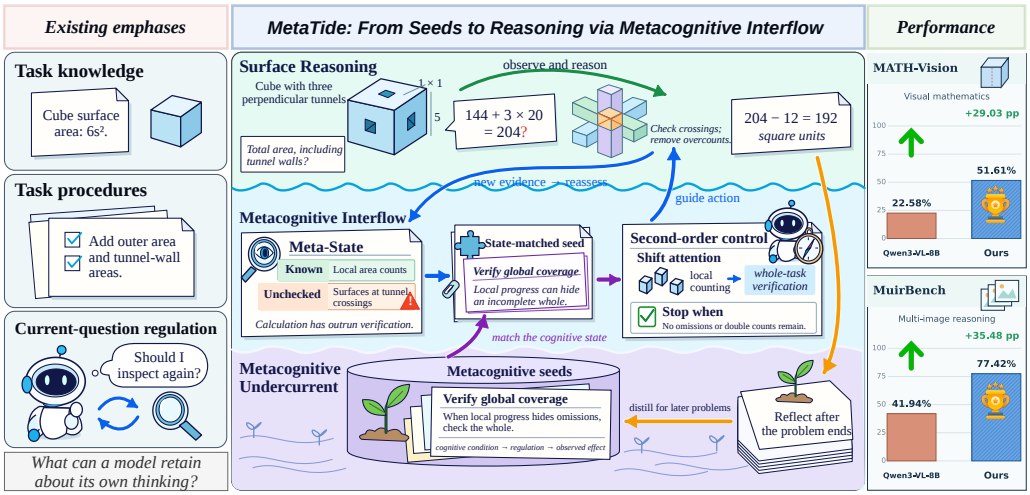


Figure 1: From task knowledge and within-question regulation to metacognition across problems. `MetaTide` matches the current Meta-State with persistent seeds, guides surface reasoning through second-order control, and distills completed episodes for later questions. Right: performance on MATH-Vision and MuirBench.

On this basis, we introduce `MetaTide`, a metacognitive framework comprising *Surface Reasoning*, *Metacognitive Interflow*, and the *Metacognitive Undercurrent*, as shown in Figure 1. These are not parallel modules, but a cognitive current moving from depth to surface and back again: metacognitive seeds held in the Metacognitive Undercurrent are stirred by the present solving situation. Through Metacognitive Interflow, they inform the model's assessment of its ongoing reasoning and how it should adjust, and shape the observation and reasoning unfolding in Surface Reasoning. The results of action flow back into this current, changing the model's understanding of its own solving at the next moment; once a problem is complete, the successes and failures of its judgments and adjustments can settle into new metacognitive seeds, ready to awaken a fitting cognitive state in future reasoning.

❶ **Surface Reasoning** → enacts that control through task-specific observation and inference, then supplies feedback. Each action yields new evidence that informs the next assessment of what is known, what remains uncertain, and how reasoning should proceed.
❷ **Metacognitive Interflow** → reads the Meta-State, activates a fitting seed, and forms second-order control. This control determines what to inspect, reconsider, or verify in the next step as the model's cognitive state changes.
❸ **Metacognitive Undercurrent** → stores transferable seeds distilled from completed solutions. New seeds are added and related seeds consolidated to guide subsequent problems.

Thus, `MetaTide` lets prior self-understanding continually enter present solving, while present solving gives rise to the beginnings of future cognition.

Our contributions are threefold:

① **Metacognitive seeds.** We introduce metacognitive seeds that capture a transferable relation between a cognitive condition, a regulatory choice, and its observed effect, preserving how reasoning should be regulated rather than task answers or fixed procedures. A seed can be reactivated when a similar cognitive condition arises in a different task.

② **Metacognitive Interflow.** We propose `MetaTide`, which matches evolving metacognitive seeds to the current Meta-State and translates them into second-order control that guides task-specific observation and reasoning. Completed solutions, in turn, renew these seeds for future problems.

③ **Effective metacognitive regulation.** Across five multimodal benchmarks, `MetaTide` achieves strong answer accuracy and shows improved premise recognition, requirement recovery, and action redirection across diverse visual reasoning settings. A cross-domain case further illustrates how the same seed guides distinct task-specific actions through metacognitive control.

2 Related Work

2.1 Multimodal Reasoning

Early multimodal language models connected visual perception with language generation through interleaved image–text learning, modality alignment, and visual instruction tuning (Alayrac et al., 2022; Li et al., 2023; Liu et al., 2023). As these models became capable of answering questions about images, attention turned to how visual evidence could support intermediate reasoning. Multimodal-CoT (Zhang et al., 2023) incorporated images into rationale generation, while visual search and region-focused reasoning enabled closer inspection of details relevant to a question (Wu & Xie, 2024; Shao et al., 2024). Visual scratchpads extended this interaction by allowing models to construct diagrams and feed them back into inference (Hu et al., 2024; Menon et al., 2024). Alongside improvements in visual understanding (Luo et al., 2025), staged reasoning and cross-modal formalization provided more structured paths from perception to an answer (Xu et al., 2025; Yang et al., 2025). CINEMA (Yin et al., 2026) brings this trajectory toward explicit action selection, organizing multi-image reasoning through operations such as global inspection, focused observation, and deliberation. Together, these developments enrich the evidence and operations available to a multimodal solver, providing the substrate for the seed-guided regulation studied here.

2.2 Metacognition in Foundation Models

The distinction between monitoring and control offers a useful lens for organizing work on model self-assessment, correction, and strategy selection (Nelson, 1990). Language models can estimate aspects of their knowledge and answer reliability (Kadavath et al., 2022), and self-generated feedback can support iterative refinement (Madaan et al., 2023). Yet intrinsic self-correction can also preserve errors or overturn correct answers (Huang et al., 2024), showing why producing a critique is insufficient. Tool-supported verification and reinforcement learning address the reliability of correction (Gou et al., 2024; Kumar et al., 2025); interleaved action feedback and search provide opportunities to revise a trajectory before committing to an answer (Yao et al., 2023; 2022). A further question is what kind of reasoning the current difficulty requires. Task-specific reasoning structures and explicit monitoring stages address this choice (Zhou et al., 2024; Oh & Gobet, 2025), while Chain of Mindset (Jiang et al., 2026) and Cognitive Demand Steering (Scoville et al., 2026) adapt cognitive modes or interventions to evolving reasoning states. `MetaTide` carries monitoring and control across problems: completed trajectories produce metacognitive seeds that guide later interventions.

2.3 Strategy-Based Memory

Experience can guide reasoning only if the agent retains useful information and recognizes when it applies. Memory architectures support retention through experience retrieval, reflection, and context management (Park et al., 2023; Packer et al., 2023), but the form of stored experience determines the guidance available. Trajectory exemplars preserve concrete action sequences, while verbal reflections capture lessons from previous attempts (Zheng et al., 2024; Shinn et al., 2023). Executable skills, workflows, and thought templates offer reusable procedures at different levels of abstraction (Wang et al., 2023; 2024b; Yang et al., 2024); ReMe (Cao et al., 2026b) and ReasoningBank (Ouyang et al., 2026) distill broader insights, including lessons from unsuccessful experience. Even a useful lesson, however, must remain relevant as a problem unfolds. MC^2 (Zhuang et al., 2026) improves the accumulated guidance through consolidation of metacognitive traces across instances, whereas AdaMEM (Zhang et al., 2026) regenerates short-term strategies to suit the evolving state within an episode. `MetaTide` stores the relation among a cognitive condition, a regulatory choice, and its observed effect, then reactivates that relation when a similar Meta-State arises.

3 Method

3.1 Overview

Consider an ordered stream of multimodal questions $x_n = (q_n, V_n)$, where q_n is the question and V_n its visual input. At step t of question n, $h_{n,t}$ contains the solver's observations and reasoning, while $g_{n,t}$ records the assessed states and planning decisions. The metacognitive seed bank $\mathcal{M}_n$ holds

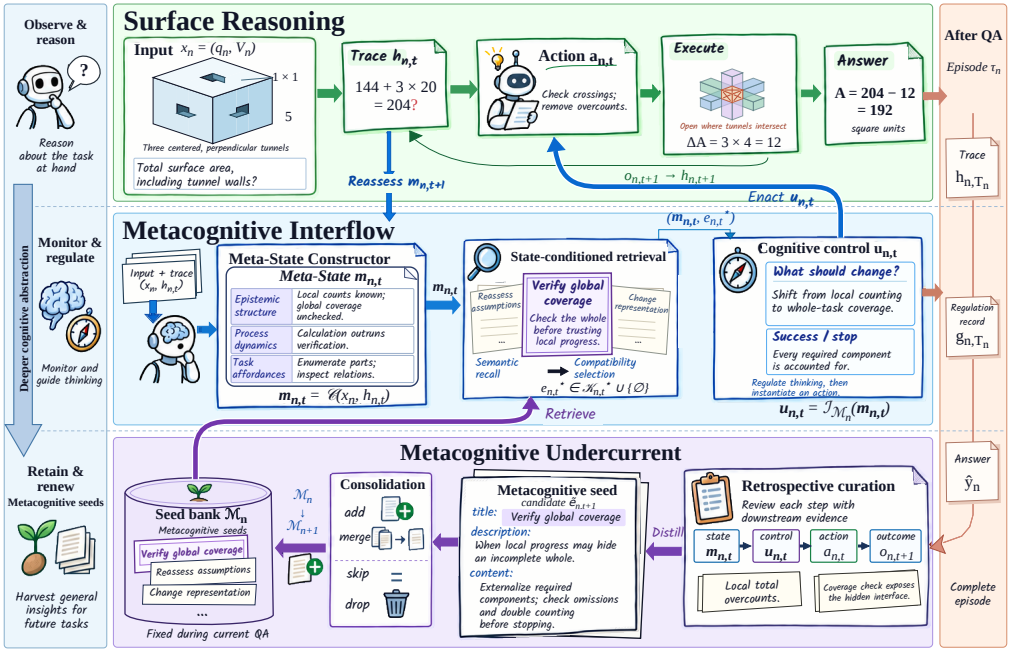


Figure 2: **MetaTide overview.** Stored metacognitive seeds inform state-dependent second-order control, which is instantiated as a local first-order action. Execution changes both the task trace and the next assessment of the solving process. The completed episode supports retrospective seed curation for later questions. The cube gives a schematic instance of this loop.

principles distilled from completed questions and remains fixed throughout question n. We index the T_n executed steps by $t = 0, \ldots, T_n - 1$ and write $\tau_n = (h_{n,T_n}, g_{n,T_n}, \hat{y}_n)$ for the completed episode, with $\hat{y}_n$ the final answer.

The three levels of MetaTide participate in one cognitive current (Figure 2). The *Metacognitive Undercurrent* holds transferable metacognitive seeds. *Metacognitive Interflow* brings these seeds into the present: it constructs a Meta-State $\boldsymbol{m}_{n,t}$, awakens an applicable seed $e^{\star}_{n,t}$, and forms a second-order control decision $\boldsymbol{u}_{n,t}$. *Surface Reasoning* realizes that decision as a task-specific action $a_{n,t}$. The action's consequence changes the evidence from which the next Meta-State is formed. After the question ends, reflection on τ_n can settle into new seeds for $\mathcal{M}_{n+1}$.

This current couples two timescales: regulation unfolds within a question, and the seed bank evolves between completed questions. Their joint process factorizes as:

$$p(\tau_{1:N}, \mathcal{M}_{2:N+1} \mid x_{1:N}, \mathcal{M}_1) = \prod_{n=1}^{N} p_{\mathrm{Q}}(\tau_n \mid x_n, \mathcal{M}_n)\, p_{\mathrm{U}}(\mathcal{M}_{n+1} \mid \tau_n, \mathcal{M}_n)\,, \tag{1}$$

where p_{Q} describes solving under a fixed bank and p_{U} describes retrospective curation from the completed episode. Deterministic stages are represented by point masses. The result of question n first changes the bank available to question $n+1$, while action outcomes update the current question. We next describe surface feedback, control, and seed renewal.

3.2 Surface Reasoning: Enacting and Reassessing Control

Enacting a control decision. At step t, the first-order planner $\mathcal{P}^{(1)}$ receives the task context $\xi_{n,t} = (x_n, h_{n,t}, \boldsymbol{m}_{n,t})$ and a control contract $\boldsymbol{u}_{n,t} = (\rho_{n,t}, \kappa_{n,t}, \omega_{n,t}, \sigma_{n,t}, \phi_{n,t})$ from Metacognitive Interflow. Here ρ names the regulatory target, κ the cognitive strategy, ω why it applies now, σ an observable success signal, and ϕ the switching or stopping condition. The planner chooses a feasible action that implements ρ, κ, and σ:

$$a_{n,t} = \mathcal{P}^{(1)}(\xi_{n,t}; \boldsymbol{u}_{n,t}) \in \big\{ a \in \mathcal{A}(x_n, h_{n,t}) \mid a \models (\rho_{n,t}, \kappa_{n,t}, \sigma_{n,t}) \big\}, \tag{2}$$

where $\mathcal{A}$ is the set of task-feasible actions and $a \models (\rho, \kappa, \sigma)$ denotes this alignment. Checking global coverage, for example, may require comparing image regions or enumerating visible geometric faces. If a contract is infeasible, the second-order planner revises it.

Observing the result. Let $\mathcal{S}$ execute the chosen action, yielding $o_{n,t+1} = \mathcal{S}(x_n, h_{n,t}, a_{n,t})$, a new observation or inference. The trace and regulation record advance:

$$(h_{n,t+1}, g_{n,t+1}) = \big(h_{n,t} \oplus o_{n,t+1},\ g_{n,t} \oplus (\boldsymbol{m}_{n,t}, \boldsymbol{u}_{n,t}, a_{n,t})\big), \tag{3}$$

Here $\oplus$ appends, $g_{n,0} = \langle\rangle$, and g records state, control, and action for review. If reasoning continues, $\mathcal{C}$ reassesses the trace, yielding $\boldsymbol{m}_{n,t+1} = \mathcal{C}(x_n, h_{n,t+1})$. The solver stops on completion or returns its best-supported answer at the inference limit.

3.3 Metacognitive Interflow: From State to Control

Monitoring the changing state. The constructor reads x_n and the actual first-order trace $h_{n,t}$ to produce a structured textual Meta-State:

$$\boldsymbol{m}_{n,t} = \mathcal{C}(x_n, h_{n,t}) = \big(m^{\mathrm{epi}}_{n,t}, m^{\mathrm{dyn}}_{n,t}, m^{\mathrm{aff}}_{n,t}\big), \tag{4}$$

where m^{epi} describes supported beliefs, unverified assumptions, remaining obligations, and decisive uncertainty; m^{dyn} describes progress, repetition, contradiction, and changes in the reasoning process; and m^{aff} describes cognitive operations available under the current evidence and resources. These components characterize the ongoing process. Earlier control becomes visible through its consequences in the actual trace.

Awakening applicable seeds. The full Meta-State queries seed bodies $b(e)$, each joining a title, applicability description, and content (formalized below). Semantic recall retains up to k entries; a small bank can be compared in full. With $k_n = \min(k, |\mathcal{M}_n|)$, the candidate set is:

$$\mathcal{K}^{\star}_{n,t} \in \operatorname*{arg\,max}_{\mathcal{K} \subseteq \mathcal{M}_n,\ |\mathcal{K}| = k_n} \sum_{e \in \mathcal{K}} s_{\mathrm{sem}}(\boldsymbol{m}_{n,t}, b(e))\,, \tag{5}$$

where s_{sem} scores semantic fit to the joint cognitive state and $b(e)$ concatenates the seed's three fields. Compatibility selection then returns $e^{\star}_{n,t} = \mathcal{R}_{\mathrm{cmp}}(\boldsymbol{m}_{n,t}, \mathcal{K}^{\star}_{n,t}) \in \mathcal{K}^{\star}_{n,t} \cup \{\varnothing\}$. The empty choice leaves a seed dormant when its cognitive condition does not fit. Task structure reaches this comparison through $m^{\mathrm{aff}}_{n,t}$, allowing similar regulatory needs to be recognized across different question types.

Forming a second-order decision. An activated seed acquires a present purpose through second-order planning. Its output is a control contract $\boldsymbol{u}_{n,t} = (\rho_{n,t}, \kappa_{n,t}, \omega_{n,t}, \sigma_{n,t}, \phi_{n,t})$. We express the state-to-control transformation of Interflow as:

$$\boldsymbol{u}_{n,t} = \mathcal{I}_{\mathcal{M}_n}(\boldsymbol{m}_{n,t}) := \mathcal{P}^{(2)}\big(\boldsymbol{m}_{n,t}, \mathcal{R}_{\mathrm{cmp}}\big(\boldsymbol{m}_{n,t}, \mathcal{K}^{\star}_{n,t}\big)\big)\,, \tag{6}$$

where $\mathcal{I}_{\mathcal{M}_n}$ maps the current state to cognitive control under the fixed bank and $\mathcal{P}^{(2)}$ forms the second-order decision from that state and the selected seed. A contract may call for testing a consequential assumption, recovering an overlooked requirement, or changing representation. When $e^{\star}_{n,t} = \varnothing$, planning proceeds from the Meta-State alone. The contract directs surface reasoning and supplies a criterion for judging its effect.

Mediating the surface action. Let Γ denote the set of control contracts and $\pi^{(2)}_{n,t}$ and $\pi^{(1)}_{n,t}$ the second- and first-order generation laws. Marginalizing control gives the surface-action law:

$$\pi^{\mathrm{surf}}_{n,t}(a) = \sum_{\boldsymbol{u} \in \Gamma} \pi^{(2)}_{n,t}\big(\boldsymbol{u} \mid \boldsymbol{m}_{n,t}, e^{\star}_{n,t}\big)\, \pi^{(1)}_{n,t}(a \mid \xi_{n,t}, \boldsymbol{u})\,, \tag{7}$$

where the selected seed affects a through control, and the first-order planner receives the task context and realized contract. Thus $a_{n,t} \perp e^{\star}_{n,t} \mid (\xi_{n,t}, \boldsymbol{u}_{n,t})$. Its outcome informs the next Meta-State.

3.4 The Metacognitive Undercurrent: Seeds and Renewal

Stored metacognitive seeds. A seed links a cognitive condition, a regulation principle, and an observable sign of its effect. Following the three-field organization of ReasoningBank (Ouyang et al., 2026), an entry and its retrieval body are:

$$e_i^{(n)} = \big(\ell_i^{(n)}, \delta_i^{(n)}, \chi_i^{(n)}\big) \in \mathcal{M}_n, \qquad b\Big(e_i^{(n)}\Big) = \ell_i^{(n)} \oplus \delta_i^{(n)} \oplus \chi_i^{(n)}, \tag{8}$$

where ℓ, δ, and χ denote the `title`, `description`, and `content`, respectively, and (n, i) identifies an entry within a bank version. The title names the principle, the description states when it applies, and the content gives transferable guidance. A global-coverage seed can guide different images or calculations; provenance, embeddings, and examples remain external.

Retrospective construction. After τ_n is complete, the extractor revisits each step with its later consequences and proposes a seed from the state, decision, action, and result:

$$\mathcal{C}_n = \Big[\mathcal{E}(\boldsymbol{m}_{n,j-1}, \boldsymbol{u}_{n,j-1}, a_{n,j-1}, o_{n,j}; \tau_n)\Big]_{j=1}^{T_n}, \tag{9}$$

where $\mathcal{E}$ abstracts a transferable condition–regulation–effect relation, $o_{n,j}$ is the outcome of step $j-1$, and $\widetilde{e}_{n,j} = \mathcal{C}_n[j]$ is the resulting three-field candidate. Later steps reveal whether progress persisted. Extraction uses the trace and answer without gold or scorer feedback.

Selective consolidation. The curator processes candidates sequentially from $\mathcal{B}_{n,0} = \mathcal{M}_n$. For candidate $\widetilde{e}_{n,j}$, it consults the current bank and external first-order examples $\mathcal{F}$, assigning $d_{n,j} \in \{\mathsf{A}, \mathsf{M}, \mathsf{S}, \mathsf{D}\}$ for add, merge, skip, or drop. A merge identifies one related entry $e_{n,j}^- \in \mathcal{B}_{n,j-1}$ and a consolidated successor $e_{n,j}^+$. The corresponding edit is:

$$\mathcal{U}_{n,j}(\mathcal{B}) := \big(\mathcal{B} \setminus \{e_{n,j}^- \mid d_{n,j} = \mathsf{M}\}\big) \cup \{\widetilde{e}_{n,j} \mid d_{n,j} = \mathsf{A}\} \cup \{e_{n,j}^+ \mid d_{n,j} = \mathsf{M}\}, \tag{10}$$

where conditional singletons are empty when their condition fails. Applying $\mathcal{B}_{n,j} = \mathcal{U}_{n,j}(\mathcal{B}_{n,j-1})$ inserts a new principle under ADD or replaces one related principle under MERGE. The bank stays unchanged under SKIP, which finds an equivalent already stored, and DROP, which rejects the candidate itself. With $\mathcal{M}_{n+1} = \mathcal{B}_{n,T_n}$, accepted writes and net growth are:

$$W_n = \sum_{j=1}^{T_n} \mathbf{1}_{\{\mathsf{A},\mathsf{M}\}}(d_{n,j}), \qquad \Delta_n = \sum_{j=1}^{T_n} \mathbf{1}_{\{\mathsf{A}\}}(d_{n,j}) = |\mathcal{M}_{n+1}| - |\mathcal{M}_n|, \tag{11}$$

Thus merges count as accepted writes without growing the bank. Continual runs apply edits before the next question; frozen-bank runs retain $\mathcal{M}_{n+1} = \mathcal{M}_n$. External examples support retrospective curation but never online seed selection.

4 Experiments

4.1 Benchmarks

To cover complementary visual and reasoning demands, we evaluate `MetaTide` on 5 benchmarks. ❶ MATH-Vision (Wang et al., 2024a) and ❷ MathVerse (Zhang et al., 2024) require mathematical reasoning grounded in visual inputs. ❸ M3CoT (Chen et al., 2024) tests multistep multimodal inference, while ❹ MMMU-Pro (Yue et al., 2025) spans academic disciplines and ❺ MuirBench (Wang et al., 2025) requires reasoning across multiple images. Together, they vary both the evidence a solver must inspect and the structure of the reasoning needed to answer. We report answer accuracy in percent throughout, allowing the results to show where the same regulation loop remains useful across these settings.

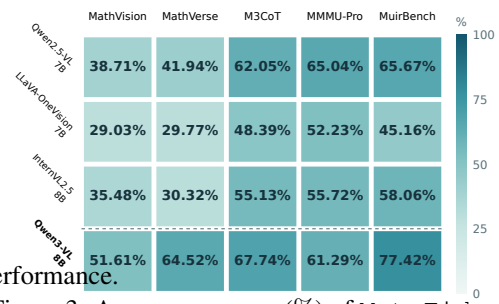
4.2 Main Results

Table 1 compares `MetaTide` with published multimodal systems and a local memory baseline by answer accuracy. ViLoMem's reported Qwen3-VL-8B-Instruct result is included as a recent memory reference (Bo et al., 2026); further memory-related results appear in Appendix A.

Table 1: Answer accuracy (%) across multimodal benchmarks.

Model / method	Size	❶ MATH-Vision Acc.↑	❷ MathVerse Acc.↑	❸ M3CoT Acc.↑	❹ MMMU-Pro Acc.↑	❺ MuirBench Acc.↑
Closed-source models						
GPT-4o	–	–	–	55.70$_{\uparrow 12.04}$	56.13$_{\uparrow 5.16}$	68.00$_{\uparrow 9.42}$
Open-source models						
LLaVA-OneVision	7B	–	26.20$_{\uparrow 38.32}$	–	–	41.80$_{\uparrow 35.62}$
InternVL2.5	8B	22.00$_{\uparrow 29.61}$	39.50$_{\uparrow 25.02}$	–	34.30$_{\uparrow 26.99}$	51.10$_{\uparrow 26.32}$
Qwen2.5-VL	7B	19.50$_{\uparrow 32.11}$	40.40$_{\uparrow 24.12}$	60.10$_{\uparrow 7.64}$	38.00$_{\uparrow 23.29}$	57.90$_{\uparrow 19.52}$
Qwen3-VL	8B	22.58$_{\uparrow 29.03}$	–	54.84$_{\uparrow 12.90}$	61.29	41.94$_{\uparrow 35.48}$
Reasoning models						
R1-Onevision	7B	29.90$_{\uparrow 21.71}$	46.40$_{\uparrow 18.12}$	57.30$_{\uparrow 10.44}$	33.90$_{\uparrow 27.39}$	–
VLAA-Thinker	7B	26.40$_{\uparrow 25.21}$	47.80$_{\uparrow 16.72}$	61.30$_{\uparrow 6.44}$	39.50$_{\uparrow 21.79}$	–
VisionR1	7B	29.90$_{\uparrow 21.71}$	52.40$_{\uparrow 12.12}$	53.20$_{\uparrow 14.54}$	30.30$_{\uparrow 30.99}$	–
MixedR1	7B	30.30$_{\uparrow 21.31}$	40.80$_{\uparrow 23.72}$	59.90$_{\uparrow 7.84}$	38.00$_{\uparrow 23.29}$	–
Memory-augmented methods						
ViLoMem	8B	49.34$_{\uparrow 2.27}$	–	–	–	–
ReasoningBank	8B	22.58$_{\uparrow 29.03}$	48.39$_{\uparrow 16.13}$	64.52$_{\uparrow 3.22}$	67.74$_{\downarrow 6.45}$	54.84$_{\uparrow 22.58}$
Ours						
MetaTide	8B	51.61	64.52	67.74	61.29	77.42

Observation 1: MetaTide improves over memory-based reasoning across diverse tasks. Table 1 shows that MetaTide exceeds ReasoningBank on four of five benchmarks. The largest gains occur on MATH-Vision and MuirBench, at +29.03 and +22.58 percentage points, spanning visual mathematics and multi-image reasoning. Improvements on MathVerse and M3CoT further extend this advantage to diagram-grounded and multistep inference. Together, these results demonstrate the value of seed-guided regulation across diverse visual reasoning demands.



Observation 2: Backbone choice shapes performance across tasks. Figure 3 compares MetaTide across four backbones. Qwen3-VL achieves the highest accuracy on four benchmarks, while Qwen2.5-VL leads on MMMU-Pro. Qwen2.5-VL also outperforms LLaVA-OneVision across all five tasks, with InternVL2.5 generally between them. The relative strengths therefore vary by task, with Qwen3-VL providing the strongest overall performance.

Figure 3: Answer accuracy (%) of MetaTide with four backbones across different benchmarks.

4.3 Online Learning and Seed Evolution

We track performance over ordered questions and the curation and use of metacognitive seeds. Figures 4 and 5 show cumulative accuracy, seed-bank evolution, and reasoning after seed injection.

Observation 3: The continual route builds a lead on three task streams. Figure 4 traces cumulative accuracy as questions arrive in each ordered stream. On MuirBench, the continual seed bank A pulls ahead after the early questions and maintains a clear lead over the frozen bank B; similar late-stream advantages appear on MATH-Vision and M3CoT. These trajectories illustrate the value of carrying metacognitive experience forward across sequential questions.

Observation 4: MetaTide grows its seed bank while deploying it selectively. Figure 5(a) separates post-question seed curation from live use. After QA, the curator makes add decisions in 66.8% and merge decisions in 30.2% of reviews, expanding the bank while consolidating related guidance. The bank grows across the stream, yet at least one seed is injected in only 25.8% of

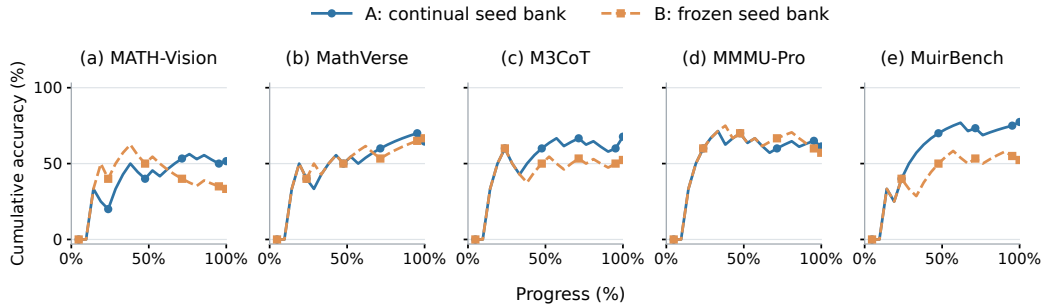


Figure 4: Cumulative online accuracy for A with an evolving seed bank and B with a frozen bank.

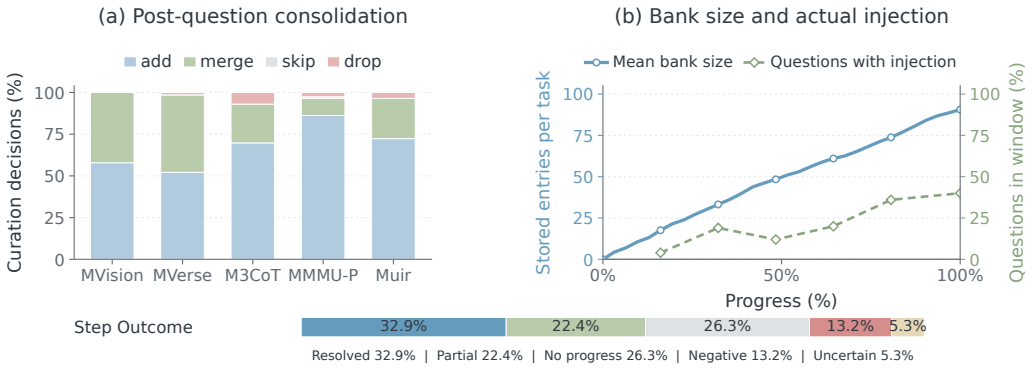


Figure 5: Seed curation, bank growth, selective injection, and step outcomes.

Table 2: Core ablations (accuracy, %). Arrows show the percentage-point margin of MetaTide.

Condition	❶ MATH-Vision Acc.↑	❷ MathVerse Acc.↑	❸ M3CoT Acc.↑	❹ MMMU-Pro Acc.↑	❺ MuirBench Acc.↑
A: MetaTide	51.61	64.52	67.74	61.29	77.42
B: Frozen seed bank	36.84$_{\uparrow 14.77}$	63.68$_{\uparrow 0.84}$	57.89$_{\uparrow 9.85}$	58.16$_{\uparrow 3.13}$	59.47$_{\uparrow 17.95}$
C: No seed retrieval	42.11$_{\uparrow 9.50}$	61.93$_{\uparrow 2.59}$	56.14$_{\uparrow 11.60}$	56.40$_{\uparrow 4.89}$	68.94$_{\uparrow 8.48}$
D: Seed to First-order	43.75$_{\uparrow 7.86}$	62.10$_{\uparrow 2.42}$	66.84$_{\uparrow 0.90}$	59.31$_{\uparrow 1.98}$	73.68$_{\uparrow 3.74}$

questions. MetaTide thus preserves lessons broadly and deploys them selectively: online retrieval brings a principle into current reasoning when it matches the constructed Meta-State.

Observation 5: Retrieved seeds advance local reasoning in most labeled steps. Figure 5 also examines the steps immediately following seed injection. The annotations indicate that 32.9% resolve the local gap and 22.4% make partial progress, yielding 55.3% with visible forward movement. Resolution addresses the missing requirement; partial progress adds relevant evidence while leaving the gap open. These labels expose intermediate progress beyond final-answer accuracy, showing retrieved seeds acting as control guidance.

4.4 Ablation

We compare the full continual route MetaTide with frozen-bank, no-retrieval, and direct-to-first-order versions. Table 2 reports answer accuracy, while Figure 6 examines regulation events.

Observation 6: Effects of seed-bank evolution, retrieval, and second-order control. Table 2 tests these components through frozen-bank (B), no-retrieval (C), and direct-to-first-order (D) ablations. Compared with the full route A, freezing the bank reduces MuirBench accuracy by 17.95↓ points, and removing retrieval reduces M3CoT accuracy by 11.60↓ points; direct first-order seed use also trails A. These declines show how seed-bank evolution, retrieval, and second-order use of seeds contribute to performance.

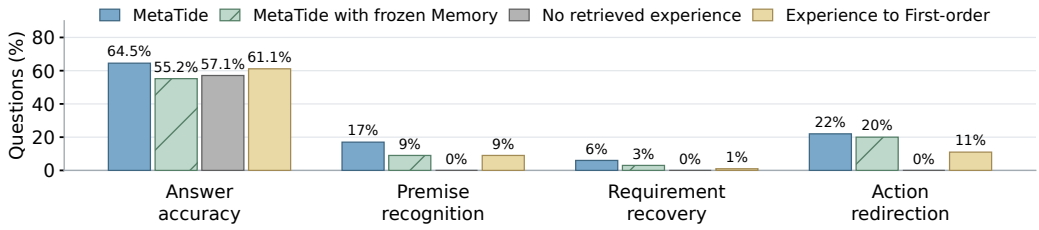


Figure 6: Answer accuracy and regulation events across different settings.

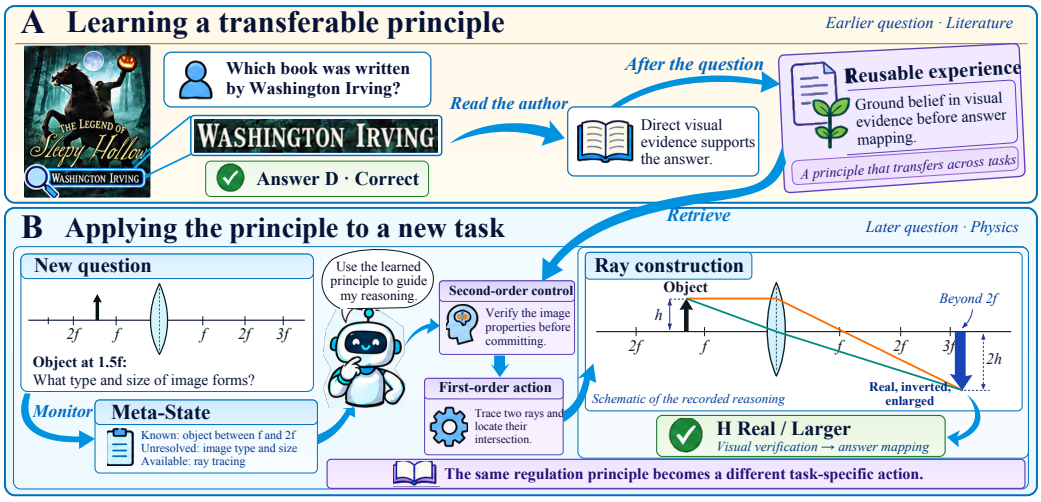


Figure 7: Cross-domain transfer of a visual-grounding principle. (A) A literature question yields reusable experience after the answer. (B) A later optics question applies that experience through metacognitive verification and ray construction.

Observation 7: `MetaTide` combines accuracy with stronger regulation. Figure 6 connects answer accuracy with regulation events. `A` achieves the highest average accuracy of $\sim$64.5%, ahead of `D` at $\sim$ 61.1%, `B` at $\sim$ 55.2%, and `C` at $\sim$ 57.1%. It recognizes mistaken premises in 17% of traces versus 9% for both `B` and `D`; one reviewed case corrects a stick-count maximum. Requirement recovery reaches 6% for `A` versus 3% and 1%, while action redirection reaches 22% versus 20% and 11%. The accuracy lead is accompanied by stronger diagnosis of what is wrong, what is missing, and what to do next.

4.5 Qualitative Case Study

Observation 8: A grounding principle becomes a task-specific verification action. Figure 7 shows transfer across tasks. The solver reads Washington Irving on a book cover; curation distills the rule: ground answers in visual evidence. For a later lens question, the `Meta-State` records an object at $1.5f$ but leaves image type and size unresolved. Retrieved experience prompts `Metacognitive Interflow` to verify before commitment. `Surface Reasoning` traces two rays; their intersection beyond $2f$ establishes a real, inverted, enlarged image and supports answer H. The seed guides a new action without carrying over the answer.

5 Conclusion

We presented `MetaTide`, a framework that brings evolving metacognitive seeds into multimodal reasoning through *Metacognitive Interflow*. By interpreting the current Meta-State and translating relevant seeds into second-order control, Interflow connects the *Metacognitive Undercurrent* with *Surface Reasoning*. Actions refine the ongoing assessment, while completed solutions yield new or consolidated seeds for later problems. Experiments demonstrate the value of this reciprocal flow for answer accuracy and metacognitive regulation. `MetaTide` thus connects regulation within a problem with the continued development of metacognitive understanding across problems.

ACKNOWLEDGMENTS

We thank colleagues with expertise in psychology and neuroscience for helpful discussions on metacognitive monitoring and control.

AI USE STATEMENT

Generative AI tools were used to assist with language editing and LaTeX formatting during manuscript preparation. All scientific content, including the methodology, experimental results, analyses, and conclusions, was reviewed and verified by the authors.

REPRODUCIBILITY STATEMENT

The reasoning framework and seed-update procedure are described in Section 3. The benchmarks, evaluation metric, comparative results, and ablations are presented in Section 4, with additional baseline context in Appendix A.

ETHICS STATEMENT

This work studies metacognitive regulation and experience reuse using existing multimodal benchmark datasets. No new human-subject data were collected for this study.

REFERENCES

Jean-Baptiste Alayrac, Jeff Donahue, Pauline Luc, Antoine Miech, Iain Barr, Yana Hasson, Karel Lenc, Arthur Mensch, Katherine Millican, Malcolm Reynolds, et al. Flamingo: a visual language model for few-shot learning. *Advances in neural information processing systems*, 35:23716–23736, 2022.

Weihao Bo, Shan Zhang, Yanpeng Sun, Jingjing Wu, Qunyi Xie, Xiao Tan, Kunbin Chen, Wei He, Xiaofan Li, Na Zhao, et al. Vilomem: Agentic learner with grow-and-refine multimodal semantic memory. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5476–5486, 2026.

Qi Cao, Yufan Wang, Peijia Qin, Shuhao Zhang, and Pengtao Xie. Llms know when they know, but do not act on it: A metacognitive harness for test-time scaling. *arXiv preprint arXiv:2605.14186*, 2026a.

Zouying Cao, Jiaji Deng, Li Yu, Weikang Zhou, Zhaoyang Liu, Bolin Ding, and Hai Zhao. Remember me, refine me: A dynamic procedural memory framework for experience-driven agent evolution. In *Findings of the Association for Computational Linguistics: ACL 2026*, pp. 16803–16822, 2026b.

Qiguang Chen, Libo Qin, Jin Zhang, Zhi Chen, Xiao Xu, and Wanxiang Che. M3cot: A novel benchmark for multi-domain multi-step multi-modal chain-of-thought. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 8199–8221, 2024.

John H Flavell. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, 34(10):906, 1979.

Zhibin Gou, Zhihong Shao, Yeyun Gong, Yujiu Yang, Nan Duan, Weizhu Chen, et al. Critic: Large language models can self-correct with tool-interactive critiquing. In *International Conference on Learning Representations*, volume 2024, pp. 57734–57811, 2024.

Yushi Hu, Weijia Shi, Xingyu Fu, Dan Roth, Mari Ostendorf, Luke Zettlemoyer, Noah A Smith, and Ranjay Krishna. Visual sketchpad: Sketching as a visual chain of thought for multimodal language models. *Advances in Neural Information Processing Systems*, 37:139348–139379, 2024.

Jie Huang, Xinyun Chen, Swaroop Mishra, Huaixiu Steven Zheng, Adams Yu, Xinying Song, and Denny Zhou. Large language models cannot self-correct reasoning yet. In *International conference on learning representations*, volume 2024, pp. 32808–32824, 2024.

Tianyi Jiang, Arctanx An, Hengyi Feng, Naixin Zhai, Haodong Li, Xiaomin Yu, Jiahui Liu, Hanwen Du, Shuo Zhang, Zhi Yang, et al. Chain of mindset: Reasoning with adaptive cognitive modes. *arXiv preprint arXiv:2602.10063*, 2026.

Saurav Kadavath, Tom Conerly, Amanda Askell, Tom Henighan, Dawn Drain, Ethan Perez, Nicholas Schiefer, Zac Hatfield-Dodds, Nova DasSarma, Eli Tran-Johnson, et al. Language models (mostly) know what they know. *arXiv preprint arXiv:2207.05221*, 2022.

Aviral Kumar, Vincent Zhuang, Rishabh Agarwal, Yi Su, John D Co-Reyes, Avi Singh, Kate Baumli, Shariq Iqbal, Colton Bishop, Rebecca Roelofs, et al. Training language models to self-correct via reinforcement learning. In *International Conference on Learning Representations*, volume 2025, pp. 54523–54549, 2025.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in neural information processing systems*, 33: 9459–9474, 2020.

Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. In *International conference on machine learning*, pp. 19730–19742. PmLR, 2023.

Junwei Liao, Haoting Shi, Ruiwen Zhou, Jiaqian Wang, Shengtao Zhang, Wei Zhang, Ying Wen, Zhiyu Li, Feiyu Xiong, Bo Tang, et al. Memq: Integrating q-learning into self-evolving memory agents over provenance dags. *arXiv preprint arXiv:2605.08374*, 2026.

Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. Visual instruction tuning. *Advances in neural information processing systems*, 36:34892–34916, 2023.

Kang Liu, Zijing Wang, Yongkang Liu, Mengjie Zhao, Xiaocui Yang, Shi Feng, Yifei Zhang, and Daling Wang. Tram: Enhancing multimodal reasoning with trajectory-derived auxiliary memory. *arXiv preprint arXiv:2608.01922*, 2026.

Gen Luo, Xue Yang, Wenhan Dou, Zhaokai Wang, Jiawen Liu, Jifeng Dai, Yu Qiao, and Xizhou Zhu. Mono-internvl: Pushing the boundaries of monolithic multimodal large language models with endogenous visual pre-training. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 24960–24971. IEEE, 2025.

Shuai Lv, Chang Liu, Feng Tang, Yujie Yuan, Aojun Zhou, Kui Zhang, Xi Yang, and Yangqiu Song. Reflect to inform: Boosting multimodal reasoning via information-gain-driven verification. *arXiv preprint arXiv:2603.26348*, 2026.

Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, et al. Self-refine: Iterative refinement with self-feedback. *Advances in neural information processing systems*, 36:46534–46594, 2023.

Sachit Menon, Richard Zemel, and Carl Vondrick. Whiteboard-of-thought: Thinking step-by-step across modalities. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 20016–20031, 2024.

Thomas O Nelson. Metamemory: A theoretical framework and new findings. In *Psychology of learning and motivation*, volume 26, pp. 125–173. Elsevier, 1990.

Nick Oh and Fernand Gobet. Monitor-generate-verify (mgv): Formalising metacognitive theory for language model reasoning. *arXiv preprint arXiv:2511.04341*, 2025.

Siru Ouyang, Jun Yan, I Hsu, Yanfei Chen, Ke Jiang, Zifeng Wang, Rujun Han, Long Le, Samira Daruki, Xiangru Tang, et al. Reasoningbank: Scaling agent self-evolving with reasoning memory. In *International Conference on Learning Representations*, volume 2026, pp. 94327–94354, 2026.

Charles Packer, Sarah Wooders, Kevin Lin, Vivian Fang, Shishir G Patil, Ion Stoica, and Joseph E Gonzalez. Memgpt: Towards llms as operating systems. *arXiv preprint arXiv:2310.08560*, 2023.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pp. 1–22, 2023.

John Scoville, Shengzhuang Chen, Yejin Bang, Stefan Winzeck, and Jonathan Richard Schwarz. Cognitive demand steering for adaptive meta-reasoning in large language models. *arXiv preprint arXiv:2608.01319*, 2026.

Hao Shao, Shengju Qian, Han Xiao, Guanglu Song, Zhuofan Zong, Letian Wang, Yu Liu, and Hongsheng Li. Visual cot: Advancing multi-modal language models with a comprehensive dataset and benchmark for chain-of-thought reasoning. *Advances in Neural Information Processing Systems*, 37:8612–8642, 2024.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. *Advances in neural information processing systems*, 36:8634–8652, 2023.

Lev Tankelevitch, Viktor Kewenig, Auste Simkute, Ava Elizabeth Scott, Advait Sarkar, Abigail Sellen, and Sean Rintel. The metacognitive demands and opportunities of generative ai. In *Proceedings of the 2024 CHI conference on human factors in computing systems*, pp. 1–24, 2024.

Moritz Von Zahn, Lena Liebich, Ekaterina Jussupow, Oliver Hinz, and Kevin Bauer. Knowing (not) to know: Explainable artificial intelligence and human metacognition. *Information Systems Research*, 2025.

Fei Wang, Xingyu Fu, James Y Huang, Zekun Li, Qin Liu, Xiaogeng Liu, Mingyu Derek Ma, Nan Xu, Wenxuan Zhou, Kai Zhang, et al. Muirbench: A comprehensive benchmark for robust multi-image understanding. In *International Conference on Learning Representations*, volume 2025, pp. 62624–62650, 2025.

Guanzhi Wang, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. Voyager: An open-ended embodied agent with large language models. *arXiv preprint arXiv:2305.16291*, 2023.

Jieke Wang, Tiancheng Shen, Yibo Yang, and Ming-Hsuan Yang. Dual-grained agent memory and shapley context attribution for multimodal agentic learner. *arXiv preprint arXiv:2608.23268*, 2026a.

Ke Wang, Junting Pan, Weikang Shi, Zimu Lu, Houxing Ren, Aojun Zhou, Mingjie Zhan, and Hongsheng Li. Measuring multimodal mathematical reasoning with math-vision dataset. *Advances in Neural Information Processing Systems*, 37:95095–95169, 2024a.

Mengzhao Wang, Yanli Ji, Wangmeng Zuo, Peng Ye, and Chongjun Tu. Position rebinding cache reuse: Replay-free visual revisiting for interleaved multimodal reasoning. *arXiv preprint arXiv:2606.26631*, 2026b.

Zora Zhiruo Wang, Jiayuan Mao, Daniel Fried, and Graham Neubig. Agent workflow memory. *arXiv preprint arXiv:2409.07429*, 2024b.

Penghao Wu and Saining Xie. V*: Guided visual search as a core mechanism in multimodal llms. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13084–13094. IEEE, 2024.

Guowei Xu, Peng Jin, Ziang Wu, Hao Li, Yibing Song, Lichao Sun, and Li Yuan. Llava-cot: Let vision language models reason step-by-step. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2087–2098. IEEE, 2025.

Ling Yang, Zhaochen Yu, Tianjun Zhang, Shiyi Cao, Minkai Xu, Wentao Zhang, Joseph E Gonzalez, and Bin Cui. Buffer of thoughts: Thought-augmented reasoning with large language models. *Advances in Neural Information Processing Systems*, 37:113519–113544, 2024.

Yi Yang, Xiaoxuan He, Hongkun Pan, Xiyan Jiang, Yan Deng, Xingtao Yang, Haoyu Lu, Dacheng Yin, Fengyun Rao, Minfeng Zhu, et al. R1-onevision: Advancing generalized multimodal reasoning through cross-modal formalization. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2376–2385. IEEE, 2025.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. React: Synergizing reasoning and acting in language models. *arXiv preprint arXiv:2210.03629*, 2022.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. *Advances in neural information processing systems*, 36:11809–11822, 2023.

Bowen Ye, Yongchao Xu, Zhijian Li, Xiang Yin, Junkai Ma, and Wenzhao Li. Coevo-mem: Co-evolving retrieval policy and memory bank for llm agents. *arXiv preprint arXiv:2608.01739*, 2026.

Jianghao Yin, Qingbin Li, Kun Sun, Cheng Ding, Jie Wang, Qin Chen, Jie Zhou, Nan Wang, Changqing Li, Pei Wu, et al. Mimic human cognition, master multi-image reasoning: A meta-action framework for enhanced visual understanding. *arXiv preprint arXiv:2601.07298*, 2026.

Xinlei Yu, Chengming Xu, Zhangquan Chen, Bo Yin, Cheng Yang, Yongbo He, Yihao Hu, Jiangning Zhang, Cheng Tan, Xiaobin Hu, et al. Dual latent memory for visual multi-agent system. *arXiv preprint arXiv:2602.00471*, 2026a.

Xinlei Yu, Chengming Xu, Guibin Zhang, Zhangquan Chen, Yudong Zhang, Yongbo He, Peng-Tao Jiang, Jiangning Zhang, Xiaobin Hu, and Shuicheng Yan. Vismem: Latent vision memory unlocks potential of vision-language models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 31544–31555, 2026b.

Shuozhi Yuan, Jinqing Wang, Zihao Liu, Miaomiao Yuan, Haoran Peng, Jin Zhao, Bingwen Wang, and Haoyi Wang. Ted: Training-free experience distillation for multimodal reasoning. *arXiv preprint arXiv:2603.26778*, 2026.

Xiang Yue, Tianyu Zheng, Yuansheng Ni, Yubo Wang, Kai Zhang, Shengbang Tong, Yuxuan Sun, Botao Yu, Ge Zhang, Huan Sun, et al. Mmmu-pro: A more robust multi-discipline multimodal understanding benchmark. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 15134–15186, 2025.

Renrui Zhang, Dongzhi Jiang, Yichi Zhang, Haokun Lin, Ziyu Guo, Pengshuo Qiu, Aojun Zhou, Pan Lu, Kai-Wei Chang, Yu Qiao, et al. Mathverse: Does your multi-modal llm truly see the diagrams in visual math problems? In *European Conference on Computer Vision*, pp. 169–186. Springer, 2024.

Yunxiang Zhang, Yiheng Li, Ali Payani, and Lu Wang. Adamem: Test-time adaptive memory for language agents. *arXiv preprint arXiv:2606.05684*, 2026.

Zhuosheng Zhang, Aston Zhang, Mu Li, Hai Zhao, George Karypis, and Alex Smola. Multimodal chain-of-thought reasoning in language models. *arXiv preprint arXiv:2302.00923*, 2023.

Longtao Zheng, Rundong Wang, Xinrun Wang, and Bo An. Synapse: Trajectory-as-exemplar prompting with memory for computer control. In *International Conference on Learning Representations*, volume 2024, pp. 19036–19066, 2024.

Pei Zhou, Jay Pujara, Xiang Ren, Xinyun Chen, Heng-Tze Cheng, Quoc V Le, Denny Zhou, Swaroop Mishra, Huaixiu S Zheng, et al. Self-discover: Large language models self-compose reasoning structures. *Advances in Neural Information Processing Systems*, 37:126032–126058, 2024.

Ziqing Zhuang, Linhai Zhang, Jiasheng Si, Deyu Zhou, and Yulan He. Beyond meta-reasoning: Metacognitive consolidation for self-improving llm reasoning. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 23884–23913, 2026.

A Additional Baseline Context

A.1 Recent memory-related methods

Table 3 collects reported results from recent methods on benchmarks used in this paper. Each entry retains the backbone and evaluation variant named in its source. ViLoMem (Bo et al., 2026), which uses a persistent visual and logical semantic memory with Qwen3-VL-8B-Instruct, appears in Table 1. TRAM retains within-question reasoning in latent memory (Liu et al., 2026); TED distills teacher-guided experience from MathVerse into an in-context store (Yuan et al., 2026). PRCR reuses visual KV cache during reasoning (Wang et al., 2026b), while L^2-VMAS (Yu et al., 2026a) and VisMem (Yu et al., 2026b) use latent visual memory. MemQ (Liao et al., 2026), CoEvo-Mem (Ye et al., 2026), and DG-Mem (Wang et al., 2026a) study external memory construction or retrieval on MMMU-Pro.

Table 3: Reported results of additional memory-related methods on overlapping benchmarks.

Method	Backbone	Benchmark / metric	Score (%)
TRAM	Qwen3-VL-8B	MATH-Vision / MathVerse	49.67 / 60.57
TED	Qwen3-VL-8B	MATH-Vision, Mean@5	70.2
PRCR	Qwen3-VL-8B-Instruct	M3CoT	68.68
L^2-VMAS	Qwen3-VL-8B-Instruct	MuirBench	67.4
VisMem	Qwen2.5-VL-7B	MuirBench	69.8
MemQ	Gemma-4-E4B-it	MMMU-Pro, held-out	54.43 ± 0.88
CoEvo-Mem	Gemma-4-E4B-it	MMMU-Pro, held-out	55.49 ± 1.27
DG-Mem	Qwen3.5-27B	MMMU-Pro, held-out	73.6

Earlier reported system. `GPT-4V` reports 22.76 on MATH-Vision, 39.40 on MathVerse, and 62.60 on M3CoT (Yin et al., 2026).

B Limitations & Future Work

Our evaluation focuses on visual question answering and multimodal reasoning; extending the framework to interactive environments remains future work. Metacognitive seeds are currently represented in text, leaving richer representations that jointly encode visual and textual experience for future exploration.