Before you <think>, monitor:

Implementing Flavell's metacognitive framework in LLMs

Monitor-Generate   # Generate-Verify   # Monitor-Generate-Verify   # Metacognition   #LLM-Modulo Framework

Nick Oh
socius labs
London, UK
nick.sh.oh@socius.org

Fernand R. Gobet*
Centre for Philosophy of Natural and Social Science, London School of Economics
London, UK
f.gobet@lse.ac.uk

***ACKNOWLEDGEMENT:** This experimental implementation builds on the *Monitor-Generate-Verify* framework co-developed with Fernand Gobet.

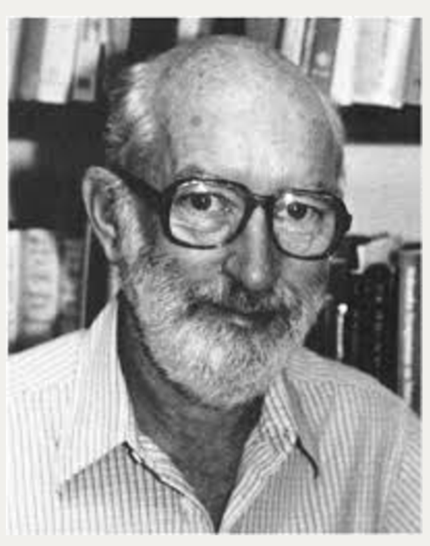
TL;DR

Monitor-Generate methods (e.g., SELF-DISCOVER) ***plan*** **but can't *verify*; Generate-Verify methods** (e.g., SELF-REFINE) ***refine*** **but start *blind.***

We chain them via Flavell's (1979) metacognitive theory.

GSM8K: 75.4% vs 68.4% accuracy, 1.3 vs 2.0 attempts. Better initial solutions need less verification.

Background



John H. Flavell
[source: Awards for Distinguished Scientific Contributions: John H. Flavell. American Psychologist, 40(3), 291-295]

Metacognition and Cognitive Monitoring

A New Area of Cognitive–Developmental Inquiry

JOHN H. FLAVELL *Stanford University*

Monitor-Generate-Verify: Formalising Metacognitive Theory for Language Model Reasoning (*Oh and Gobet, 2025*)

Algorithm 1 Flavell's Metacognitive Regulation

Initialise: $\mathcal{S}_0 \leftarrow f(\mathcal{T}, \mathcal{G})$; $\tau \leftarrow 0$
while $\mathcal{S}_\tau$ = ACTIVE **do**
  // MONITOR: Retrieve knowledge & assess experience
  $\mathcal{MK}_\tau \leftarrow$ **if** $\tau = 0$ **then** retrieve$(\mathcal{MK}, \mathcal{T}, \mathcal{G})$
    else $\mathcal{MK}_{\tau-1} \cup$ retrieve$(\mathcal{MK}, \mathcal{ME}_{\tau-1})$
  $\mathcal{ME}^{\tau}_{\text{difficulty}} \leftarrow$ feel$(\mathcal{T}, \text{Outcomes}_{\tau-1}) \oplus$ assess$(\mathcal{T}, \mathcal{MK}_\tau)$
  // GENERATE: Select & execute cognitive strategy
  $\mathcal{CS}_\tau \leftarrow$ select$(s \in \mathcal{MK}_{\text{Strategy}} \mid \mathcal{ME}^{\tau}_{\text{difficulty}}, \mathcal{MK}_\tau, \mathcal{T}, \mathcal{G})$
  $\mathcal{CO}_\tau \leftarrow$ execute$(\mathcal{CS}_\tau, \mathcal{T}, \mathcal{G})$
  // VERIFY: Evaluate progress & update knowledge
  $\mathcal{ME}^{\tau}_{\text{evaluative}} \leftarrow$ assess$(\mathcal{CO}_\tau, \mathcal{MK}_\tau)$
  $\mathcal{MS}_\tau \leftarrow$ select$(s \in \mathcal{MK}^{\text{meta}}_{\text{Strategy}} \mid \mathcal{ME}^{\tau}_{\text{evaluative}})$
  $\mathcal{MO}_\tau \leftarrow$ execute$(\mathcal{MS}_\tau, \mathcal{CO}_\tau, \mathcal{MK}_\tau, \mathcal{G})$
  $\mathcal{MK} \leftarrow$ update$(\mathcal{MK}, \Phi_\tau)$ where $\Phi_\tau = (\mathcal{ME}_\tau, \text{Strategy}_\tau, \text{Outcome}_\tau)$
  $\mathcal{S}_{\tau+1} \leftarrow$ **if** goal_achieved$(\mathcal{CO}_\tau, \mathcal{G})$ **then** TERMINATE **else** ACTIVE
  $\tau \leftarrow \tau + 1$
end while

Algorithm 1 Flavell's Model of Cognitive Monitoring

Require: task $\mathcal{T}$, model M, strategies $\mathcal{MK}$, prompts $\mathcal{P} = \{p_{monitor}, p_{strategy}, p_{execute}, p_{verify}\}$
Initialize $\mathcal{S}_0$ = ACTIVE, $\tau = 0$, $\mathcal{ME}^{-1}_{\text{evaluative}} = \emptyset$
while $\mathcal{S}_\tau$ = ACTIVE and $\tau < T$ **do**
  // Monitor: Assess task difficulty
  $\mathcal{ME}^{\tau}_{\text{difficulty}}, features_\tau = M(p_{mon} \| \mathcal{T} \| \mathcal{ME}^{\tau-1}_{\text{evaluative}})$ ▷ Assess task difficulty

  // Generate: Apply cognitive strategy
  $strategy_\tau = M(p_{str} \| features_\tau \| \mathcal{ME}^{\tau}_{\text{difficulty}} \| \mathcal{MK})$ ▷ Choose approach
  $solution_\tau = M(p_{exe} \| \mathcal{T} \| strategy_\tau)$ ▷ Execute strategy with adaptive params

  // Verify: Evaluate performance
  $\mathcal{ME}^{\tau}_{\text{evaluative}} = M(p_{ver} \| \mathcal{T} \| solution_\tau)$ ▷ Evaluate output quality

  $\mathcal{S}_{\tau+1}$ = **if** mean$(\mathcal{ME}^{\tau}_{\text{evaluative}}) \geq 0.85$ **then** TERMINATE **else** ACTIVE
  $\tau = \tau + 1$
end while
return $y_{\arg\max_i \text{mean}(\mathcal{ME}^{i}_{\text{evaluative}})}$

MGV framework (zero-shot)

MK_{Agent}	MK_{Task}	$MK_{Strategy}$
Knowledge about cognitive agents' characteristics and capabilities	Knowledge about cognitive situation assessment	Knowledge concerning the effectiveness of cognitive strategies and metacognitive strategies

MK $MK_{Strategy}$ as a pre-established repertoire of 20 domain-specific strategies (Didolkar et al., 2024)

Flavellian Metacognitive Knowledge

```
Monitor
Problem: {problem}
{if t > 0:
evaluation_scores:
coherence: {score}
plausibility: {score}
consistency: {score}
goal_conduciveness: {score}}

Before you <think>, analyse this problem WITHOUT solving it.

Output format:
<monitor>
Task_Features: List 2-3 keywords describing the math concepts needed
Difficulty (0-1): Assess the challenge level
{if t > 0: Recalibrate your assessment based on the
evaluation_scores in previous_cycle. Lower scores
suggest higher actual difficulty than previously assessed.}
</monitor>
```

```
Generate - Strategy Selection
Problem: {problem}
Task features: {task_features}
Difficulty: {difficulty:.3f}
Available strategies: {comma_separated_strategy_list}

Select ONE strategy from the list above that best fits this problem.

Output format:
Selected Strategy: [exact strategy name from list]
```

```
Generate - Strategy Execution
Problem: {problem}
Strategy: {strategy_type}

Output format:
<think>
Work through the problem step by step using the {strategy_type} approach.
Show all calculations and reasoning.
{if t > 0:
previous_cycle:
difficulty: {difficulty}
task_features: {task_features}
strategy_used: {strategy_type}
reasoning: [full solution text]
evaluation_scores:
coherence: {score}
plausibility: {score}
consistency: {score}Your paragraph text
goal_conduciveness: {score}
evaluation: [full evaluation text]}
</think>

<answer>
Output only the final numerical answer (no units or text).
</answer>
```

GENERATE	VERIFY
Token budget scales as 400 + 400*$ME_{difficulty}$ and temperature as 0.3 + 0.2*$ME_{difficulty}$	TERMINATE when mean evaluation ≥ 0.85 or T = 3

Adaptive Resource Allocation

```
Verify
Problem: {problem}
Solution: {solution}

Evaluate this solution on these dimensions (0-1 scale):

1. COHERENCE: Do the steps logically connect?
2. PLAUSIBILITY: Is the approach reasonable?
3. CONSISTENCY: Are calculations correct?
4. GOAL-CONDUCIVENESS: Does it answer the question?

Output format:
<evaluate>
Coherence: X.X
Plausibility: X.X
Consistency: X.X
Goal-conduciveness: X.X

Evaluation: [Provide a 2-3 sentence analysis explaining the scores.
Identify specific errors or strengths. For low scores, indicate
what went wrong (e.g., "arithmetic error in step 3", "misunderstood
the question", "skipped crucial reasoning"). For high scores, note
what worked well. Be specific and actionable.]
</evaluate>
```

MONITOR → GENERATE → VERIFY

Experimental Setup

`meta-llama/Llama-3.1-8B-Instruct (+NVIDIA H100 SXM)`

Self-Verification (Weng et al., 2022): Generate-*k*-Verify (w/ majority voting)
SELF-REFINE (Madaan et al., 2023): Iterating Generate-Verify *k* times
MGV (Flavell): Iterating Monitor-Generate-Verify *k* times

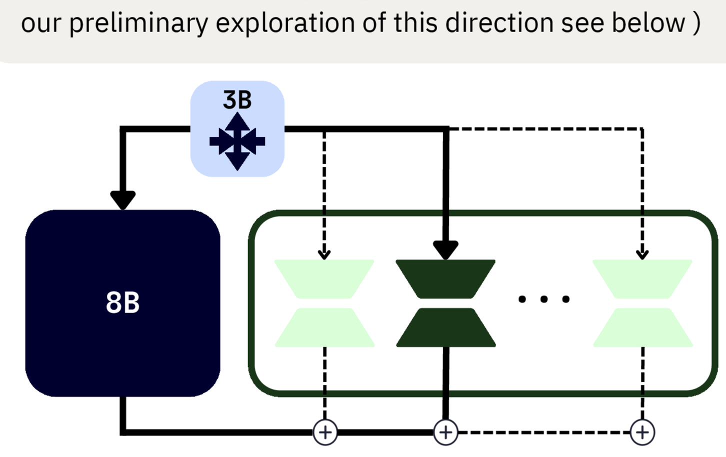
Results

Method	Accuracy	Avg Time (s)	Avg Attempts
Self-Verification	442/659 (67.07%)	7.52	1.2
Self-REFINE	451/659 (68.44%)	6.98	2.0
MGV (Flavell)	**497/659 (75.42%)**	9.60	1.3

Future Work (+ pilot experiment)

LLM Modular Framework

- **Small models (0.5B) can reliably *Plan*, while *Execution* demands substantially larger models** (Qin et al., 2025) (for our preliminary exploration of this direction see below)



multi-LoRA as $MK_{Strategy}$

Branched curriculum (2 epochs shared arithmetic → 3 specialized LoRAs @ 1 epoch each) with 3B routing outperformed 4-epoch GRPO by 1.4pp on GSM8K (8B models)

Model Intrinsic *MK*

- Following V-STaR's approach (Hosseini et al., 2024), future work could **train the Monitor through iterative self-supervised learning**: generate diverse solutions per problem, contrast successful versus failed attempts using DPO/GRPO, and progressively refine both difficulty assessment and strategy selection capabilities — building model-intrinsic metacognitive knowledge through bootstrapped preference learning.

Metacognitive Space within LLM

- **Implicit confidence measures dervied from token likelihoods exhibit greater metacognitive sensitivity** than explicitly prompted confidence (Xiong et al., 2023)
- Some evidence of **subjectivity** (e.g., confidence and certainty) **corresponding to linearly seperable directions** in representations (Zou et al., 2023; Liu et al., 2023)