

# Monitor-Generate-Verify (MGV):

Formalising Metacognitive Theory for Language Model Reasoning



# Monitor-Generate

# Generate-Verify

# Metalearning

# Metareasoning

#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

socius labs

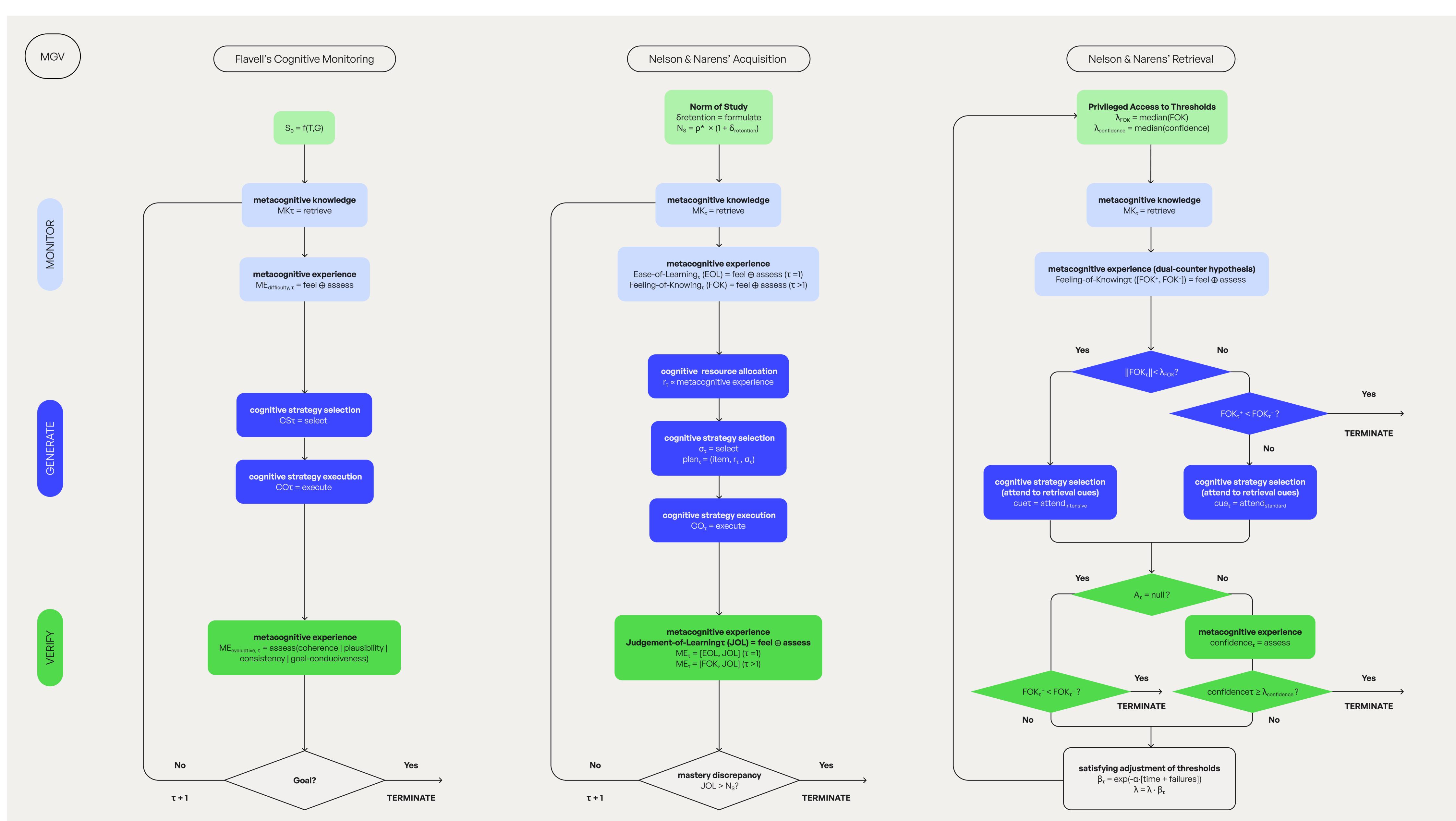


THE LONDON SCHOOL  
OF ECONOMICS AND  
POLITICAL SCIENCE ■

## TL;DR

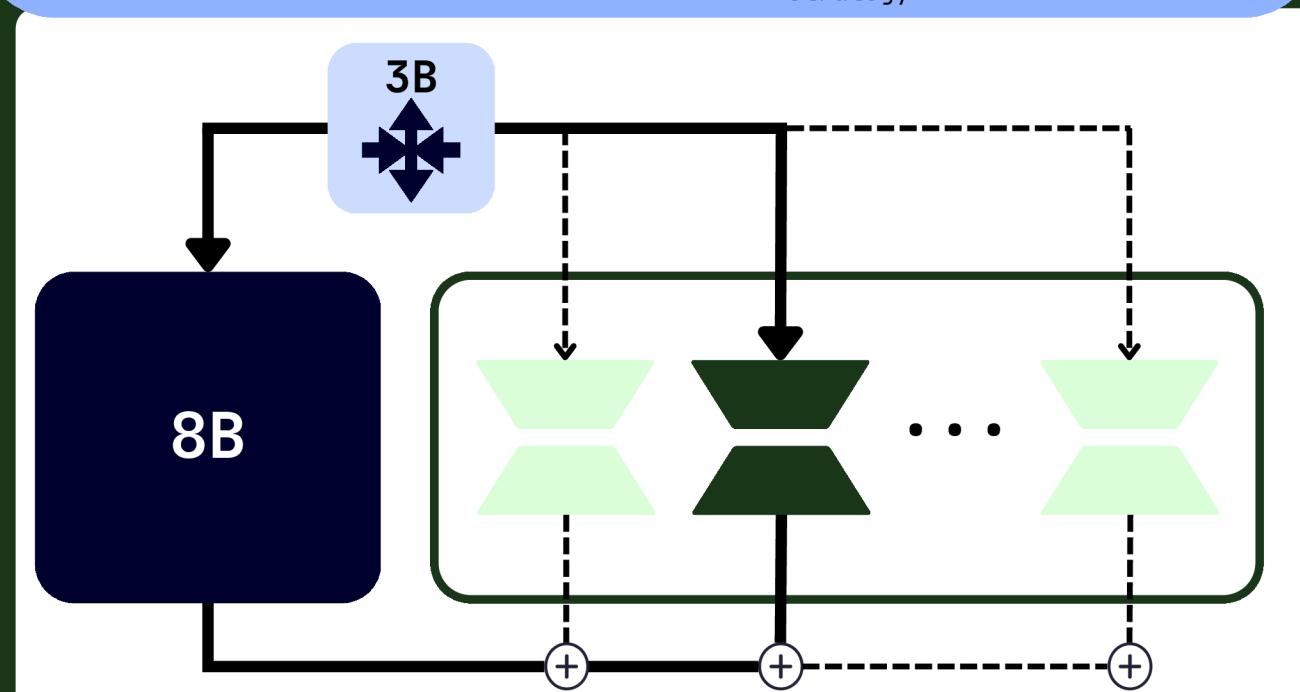
**Chain-of-Thought tells models to reason step by step.  
But it doesn't tell them *how* to reason – or *when* to stop.  
Before the chain, there are meta-decisions: *which strategy? How much computation? When to terminate?***

**Monitor-Generate-Verify (MGV) formalises metacognitive theories to address these questions.**



## Pilot Experiments

LLM Modular Framework  
(multi-LORA as  $MK_{\text{strategy}}$ )



Small models (0.5B) can reliably Plan, while Execution demands substantially larger models (Qin et al., 2025)

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)

## Zero-shot Reasoning

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

**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

## Limitations & Future Direction

**Resource-rational analysis:** Understanding human cognition as the optimal use of limited computational resources (Lieder and Griffiths, 2020)

$$a^\star = \arg \max_a \int u(o) \cdot p(o|a) do,$$

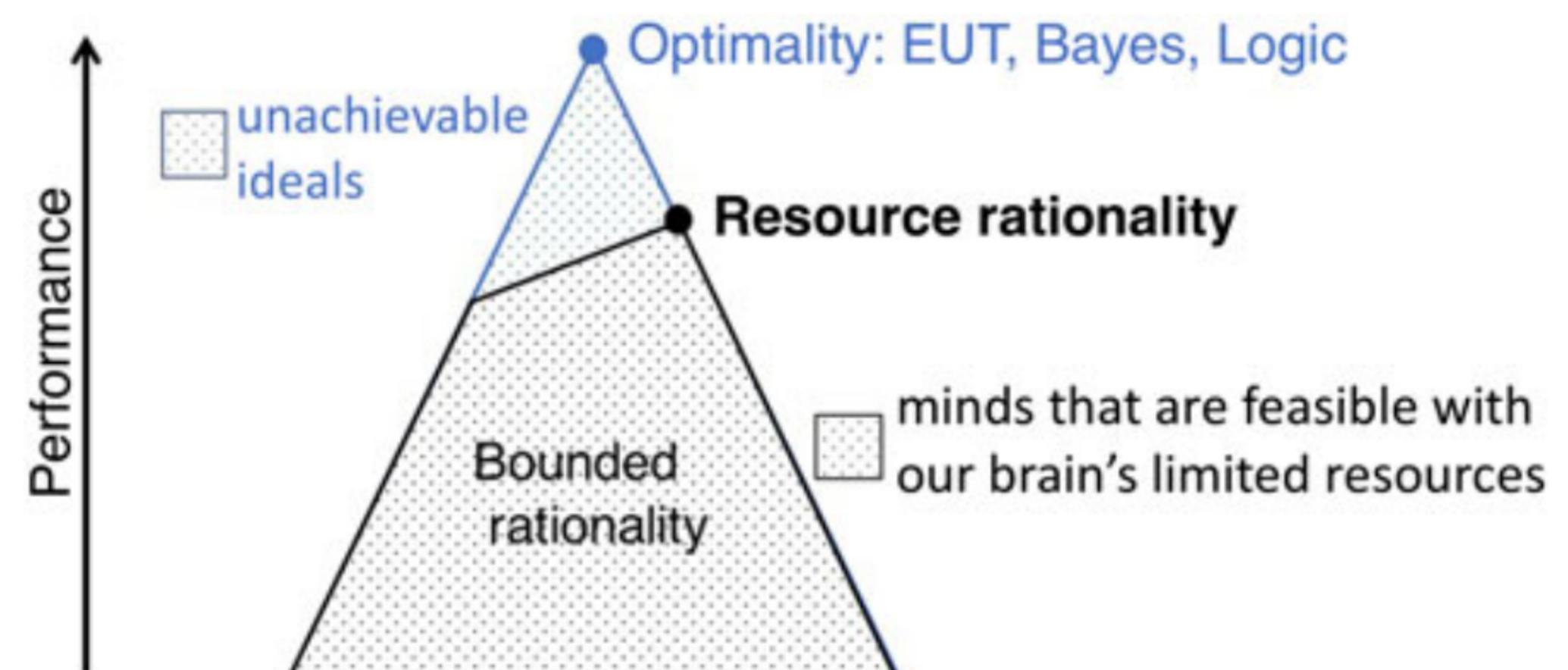


Figure 1. Resource rationality and its relationship to optimality and Tversky and Kahneman's concept of bounded rationality (Lieder and Griffiths, 2020)