RL didn't teach your model to think

Does Reinforcement Learning Really Incentivize Reasoning Capacity in LLMs Beyond the Base Model?

Yang Yue, Zhiqi Chen, Rui Lu, Andrew Zhao, Zhaokai Wang, Shiji Song, Gao Huang. LeapLab, Tsinghua University and SJTU.

arXiv 2504.13837 · NeurIPS 2025 Oral · Best Paper Runner Up · presented by S Akash, IIT Patna

QUICK FACT Reported to have received a perfect review score at NeurIPS 2025.

S Akash

GPU computing and efficient ML inference. EEE, IIT Patna.

CERN	Vendor agnostic GPU code generation for TMVA SOFIE.
IIT PATNA	SUMMIR, in collaboration with Microsoft (ECIR 2026).
LLVM	Developers' Meeting poster: heterogeneous dispatch.
vLLM	Enabled DeepSeek-V2-Lite-Chat FP8 in tpu-inference.
AUTOSTEP	Founding engineer (YC S25). Agent infrastructure.
F.INC	Alumnus of Founders, Inc.
ON STAGE	TOKEN2049, HK Web3 Festival, ETHSingapore (top five).

FIND ME

GITHUB LINKEDIN

WEBSITE

QUICK FACT I train models with PPO myself, so tonight's result applies to my own work.

Agenda

TONIGHT · ABOUT THIRTY MINUTES PLUS QUESTIONS

Part 1 **Context**: how models are post trained 4′

Part 2 **The metric**: measuring a capability boundary 5′

Part 3 **The result**: the crossover 7′

Part 4 **The mechanism**: sharpening, not expanding 4′

Part 5 **The theory**: why the leash exists 5′

Part 6 **The debate**: rebuttals and replications 3′

Part 7 **Practice**: what to do differently 2′

Part 8 **Epilogue**: the boundary since April 2′

Companion reading for Part 5 and Part 8: The Invisible Leash (arXiv 2507.14843) · DeepSeekMath 5.2.2 (arXiv 2402.03300) · FunSearch (Nature, 2023).

QUICK FACT Section 5 is the part most talks skip: the mathematics behind the result.

A quick poll

BEFORE WE BEGIN · A SHOW OF HANDS

> *We just watched DPO teach models taste. Now the harder question: does RL fine tuning on top teach them **new** reasoning, beyond what the base model could already do?*

True | **False**

NOTE YOUR ANSWER. WE RETURN TO IT AT THE END.

QUICK FACT Through most of 2025 this claim was close to consensus across the field.

Where RLVR sits in the training pipeline

PART 1 · CONTEXT · SIXTY SECONDS

Stage 1

Pretraining

Next token prediction over web scale corpora. Builds **raw capability**.

Stage 2

SFT

Supervised fine tuning on curated demonstrations. **Shapes format and skill.**

Stage 3

DPO / RLHF

Preference optimization. **Style and judgment**. The previous talk.

Stage 4, tonight

RLVR

Reinforcement learning against **verifiable correctness**. Tonight's subject.

The 2025 reasoning boom rests on stage 4: replace the human judge with a program, a math checker or a unit test suite, and let the model chase it.

QUICK FACT Learning from human preferences predates GPT-1: Christiano et al., NeurIPS 2017.

RLVR: reward only what a verifier can check

PART 1 · CONTEXT · THE OBJECTIVE

$$\max_{\theta} \; \mathbb{E}_{x\sim\mathcal{D},\, y\sim\pi_\theta(\cdot|x)}\big[\, r(x,y)\,\big], \qquad r \in \{0,1\}$$

In words: choose model parameters that maximize the expected reward over problems x and the model's own sampled answers y, where the reward is one if the verifier accepts the answer and zero otherwise.

Sample answers from the current policy, score each with a binary verifier (answer match for math, unit tests for code), reinforce what passed. No reward model, no human in the loop.

WHICH ONE HAS A VERIFIABLE REWARD?

A poem no programmatic verifier

3847×42 deterministic check ✓

An apology preference territory

QUICK FACT GRPO, the algorithm behind DeepSeek-R1, deletes PPO's critic network entirely. Its estimator appears in Part 5.

Move 37: when RL genuinely invented

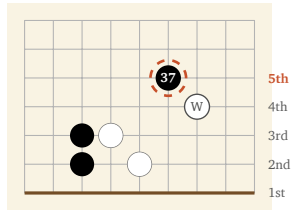
PART 1 · THE BELIEF · 2016

Game 2, AlphaGo vs Lee Sedol, March 10, 2016. AlphaGo (Black) plays a **shoulder hit on the fifth line**, a move professional doctrine treated as simply wrong.

Commentators assumed a mistake. Lee Sedol needed over twelve minutes to answer. The move anchored the win.

1 in 10,000

DEEPMIND'S ESTIMATE THAT A HUMAN PROFESSIONAL WOULD PLAY THE MOVE



GAME 2, MOVE 37 (SCHEMATIC): BLACK'S DIAGONAL SHOULDER HIT ON THE FIFTH LINE, OVER WHITE'S FOURTH LINE STONE

QUICK FACT DeepMind's Game 2 commentary with Fan Hui is titled simply "Invention".

The assumption: RLVR is language's Move 37

PART 1 · THE BELIEF · 2025

> *Train against verifiable rewards long enough, and the model will discover reasoning strategies its base never had, the way self play discovered Move 37.*

It is a natural reading of the o1 and R1 results: sudden "aha" moments, self correction, long chains of thought that emerge during RL.

This paper tests the assumption directly. Keep two regime differences in mind, because Part 5 will cash them in:

AlphaGo's regime

An exact simulator, an adversary, and **self play that keeps generating new states** to learn from.

RLVR fine tuning

A fixed prompt set, and samples drawn **from a policy that starts at the base model** and stays close to it.

Amplifier, or teacher?

PART 2 · THE QUESTION

Operationally: **is there any problem the RL trained model can solve that its own base model cannot, under any sampling budget?** In the authors' words:

> ***From the abstract, Yue et al., arXiv 2504.13837 (v5)***
>
> as the evaluation metric. While RLVR improves sampling efficiency towards correct paths, we surprisingly find that current training *rarely* elicit fundamentally new reasoning patterns. We observe that while RLVR-trained models outperform their base models at smaller values of k (*e.g.*, k=1), base models achieve higher pass@k score when k is large. Moreover, we observe that

Two hypotheses, one testable split: an **amplifier** makes existing ability cheaper to reach; a **teacher** adds ability that was not there. The whole talk lives in that difference.

QUICK FACT Across five arXiv revisions the central claim was softened from "RLVR" to "current RLVR methods". Science in public.

Three domains, six algorithms, ten benchmarks

PART 2 · EXPERIMENTAL DESIGN

Table 1, Yue et al. 2025: experimental setup

Task	Start Model	RL Framework	RL Algorithm(s)	Benchmark(s)
Mathematics	LLaMA-3.1-8B Qwen2.5-7B/14B/32B-Base Qwen2.5-Math-7B	SimpleRLZoo Oat-Zero DAPO	GRPO	GSM8K, MATH500 Minerva, Olympiad AIME24, AMC23
Code Generation	Qwen2.5-7B-Instruct DeepSeek-R1-Distill-Qwen-14B	Code-R1 DeepCoder	GRPO	LiveCodeBench HumanEval+
Visual Reasoning	Qwen2.5-VL-7B	EasyR1	GRPO	MathVista MathVision
Deep Analysis	Qwen2.5-7B-Base Qwen2.5-7B-Instruct DeepSeek-R1-Distill-Qwen-7B	VeRL	PPO, GRPO Reinforce++ RLOO, ReMax, DAPO	Omni-Math-Rule MATH500

Math uses **zero RL** (RL applied directly to the base model) so nothing else confounds the comparison. Code and visual reasoning start from instruct models, following community convention. Sampling: temperature 0.6, top p 0.95, up to 16,384 tokens, no few shot prompts for either side.

QUICK FACT The evaluated RLVR checkpoints are the community's own releases, not reruns.

pass@k measures the boundary, not the average

PART 2 · THE METRIC

pass@1

Success in a single attempt. Average case behavior. **What a user experiences.** The number in every launch post.

pass@256, pass@1024

Whether **any** of k samples passes the verifier. **Whether the capability exists at all.** The reasoning boundary.

The exam metaphor: pass@1 is exam day. pass@1024 locks the student in a room with 1,024 attempts and asks whether the knowledge exists *anywhere* in them. If a valid solution never appears in 1,024 independent tries, the capability is absent for practical purposes.

Note what this is *not*: best of N ranks candidates with a reward model and measures selection. pass@k at large k measures raw **coverage** of the solution space.

Estimating pass@k without bias

PART 2 · THE METRIC, PRECISELY

$$\widehat{\text{pass@}k} = \mathbb{E}_x\left[1 - \binom{n - c_x}{k} \Big/ \binom{n}{k}\right]$$

In words: for each problem, draw n samples and count c correct. Take the chance that a random group of k of them contains at least one correct answer. Average over problems.

Worked example

$n = 16$ samples, $c = 2$ correct, $k = 8$:

$$1 - \binom{14}{8} \Big/ \binom{16}{8} = 1 - \tfrac{3003}{12870} \approx 0.767$$

That is pass@8 ≈ **76.7%**.

The naive plug in $1 - (1 - c/n)^k$ is biased at small n; this form (Chen et al., 2021, the Codex paper) is exact. Here n runs up to 2,048.

Make a prediction

INTERLUDE · THIRTY SECONDS WITH YOUR NEIGHBOUR

Model A, after RLVR: pass@1 $= 55\%$.
Model B, its untouched base: pass@1 $= 30\%$.

At $k = 256$, which one covers **more** problems?

A	**B**	**Tie**
the RL model	the base	

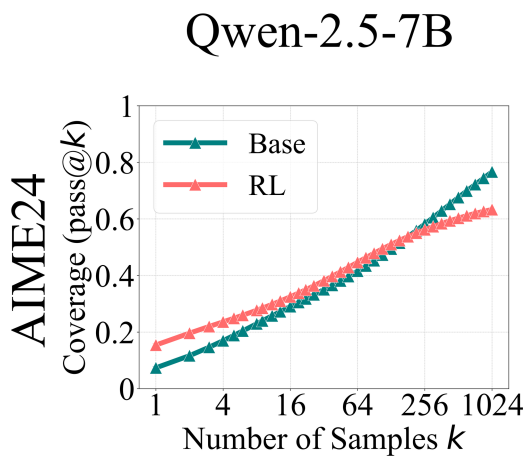
(Numbers illustrative; the pattern is the paper's.) Argue for a side: "A, RL made it strictly better" versus "B, RL concentrated probability and lost diversity".

The result: the curves cross

PART 3 · THE RESULT · AIME24, QWEN2.5-7B

- **RL leads at $k = 1$**: the number every launch post cites.
- **The base overtakes near $k \approx 150$ to 256**, and is still climbing at 1,024.
- At $k = 1024$: base 0.76 versus RL 0.63.

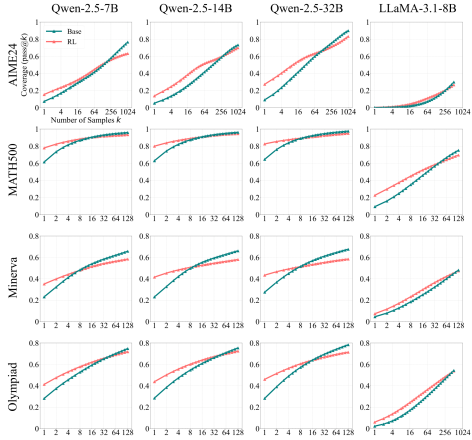
> *The base model was never less capable. It was less **concentrated**.*



YUE ET AL. 2025, FIG. 2, AIME24 PANEL. ZERO RL SETTING, SIMPLERLZOO.

Not a fluke: every family, every scale

PART 3 · REPLICATION

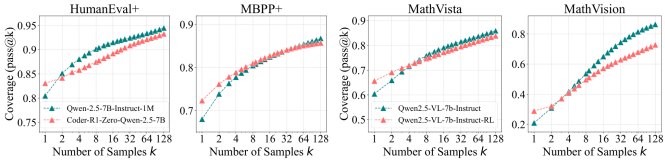


YUE ET AL. 2025, FIG. 2: QWEN2.5 7B / 14B / 32B AND LLAMA-3.1-8B

- Sixteen panels, one shape: coral leads left, teal leads right.
- Minerva, 32B: base ahead by about **9 points** at $k = 128$.
- Stronger recipes shift the crossover right, not away: Oat-Zero starts 30 points ahead, and is still overtaken.

Code and vision tell the same story

PART 3 · GENERALITY



YUE ET AL. 2025, FIG. 4: CODE (LEFT), VISUAL REASONING (RIGHT)

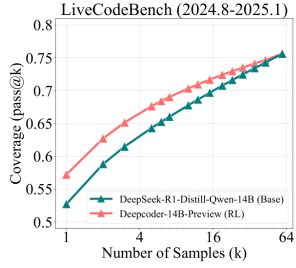


FIG. 3: LIVECODEBENCH, DEEPCODER-14B VS BASE

Code is the clean domain: unit tests cannot be lucky guessed.

QUICK FACT LiveCodeBench uses problems published after training cutoffs, blocking memorization.

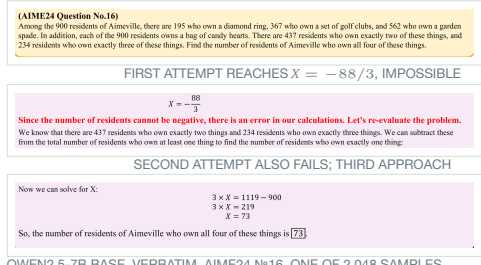
Are the base model's rare wins genuine?

PART 3 · VALIDITY · THE LUCKY GUESS OBJECTION

The objection. With 1,024 tries, a right answer may be luck.

The audit. The authors hand read every chain behind correct answers to the hardest problems (accuracy below 5%): GSM8K **24 of 25** valid, AIME24 **5 of 6**. Code, where guessing is impossible, agrees.

On the right: an unedited sample. Wrong equation, **explicit self correction**, a second failure, then a correct third attempt.



QWEN2.5-7B BASE, VERBATIM. AIME24 №16, ONE OF 2,048 SAMPLES.

The solvable sets nest

PART 3 · AIME24 AT K = 1024, MATH500 AT K = 128

0.0%

AIME24 PROBLEMS SOLVED BY THE RL MODEL THAT ITS BASE COULD NOT SOLVE

*The RL model's solvable set nests **inside** the base's.*

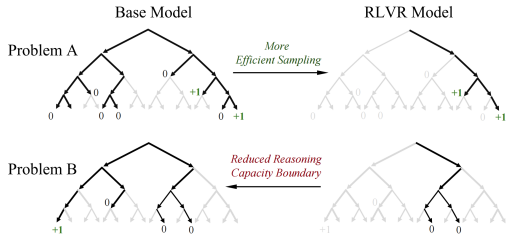
BASE	RL (SimpleRLZoo)	AIME24	MATH500
✓	✓	63.3%	92.4%
✓	×	13.3%	3.6%
×	✓	**0.0%**	**1.0%**
×	×	23.3%	3.0%

THE BASE EVENTUALLY CRACKS 76.7% = 23 OF 30 AIME24 PROBLEMS. RL LOSES 13.3% AND GAINS NOTHING. TABLE 2.

QUICK FACT AIME has 30 problems a year; every answer is an integer from 0 to 999.

Sharpening, not expanding

PART 4 · MECHANISM



YUE ET AL. 2025, FIG. 1: EFFICIENT SAMPLING (A); PATHS PRUNED (B)

Sampling efficiency

Rises

COVERAGE (BOUNDARY)

Flat, or narrows

Output entropy

Falls

QUICK FACT Raising the RL model's temperature does not restore the boundary (their Fig. 18).

In the authors' words

PART 4 · EVIDENCE · DISTRIBUTIONAL

Summary of section 4.1 (pages 8 and 9)

Summary. Combining the above analyses, we arrive at three key observations. First, problems solved by the RLVR model are also solvable by the base model; the observed improvement in average scores stems from more efficient sampling on these already solvable problems, rather than learning to solve new problems. Second, after RLVR training, the model often exhibits narrower reasoning coverage compared to its base model. Third, all the reasoning paths exploited by the RLVR model are already present in the sampling distribution of the base model. These findings indicate that RLVR does not introduce fundamentally new reasoning capabilities and that the reasoning capacity of the trained model remains bounded by that of its base model.

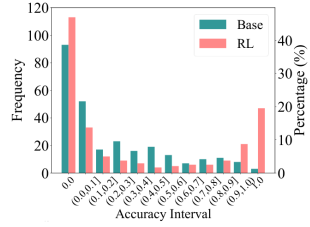


FIG. 5: ACCURACY HISTOGRAM, MINERVA, BEFORE AND AFTER RLVR

After RLVR, mass grows at accuracy near 1 **and at exactly 0**.

Observation 1

Everything the RL model solves, the base solves.

Observation 2

Score gains come from efficiency, not new problems.

Observation 3

RL's reasoning paths already exist in the base.

The base model recognizes RL's solutions

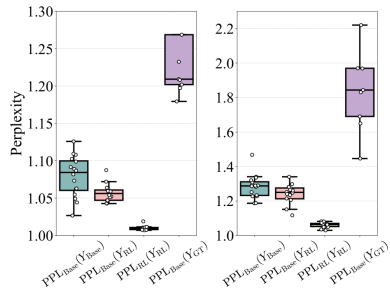
PART 4 · EVIDENCE · PERPLEXITY

$$\mathrm{PPL}(Y \mid x) = \exp\left(-\frac{1}{T}\sum_t \log P(y_t \mid x, y_{<t})\right)$$

In words: perplexity is the exponential of the average surprise per token. Low perplexity means the model finds the text familiar.

Score the RL model's winning solutions *under the base*. They land in the low tail of the base's own outputs: the base already expected them.

The control works too: **OpenAI o1's solutions do surprise the base**. Genuinely foreign reasoning is detectable. RL's is not foreign.



YUE ET AL. 2025, FIG. 6: PERPLEXITY OF RESPONSES UNDER THE BASE.

RIGHT PANEL: o1 RESPONSES STAY SURPRISING.

QUICK FACT Appendix C.4: the base's perplexity on RL outputs keeps falling during training.

Why the policy cannot leave its own support

PART 5 · THEORY · THE GRADIENT ARGUMENT

$$\nabla_\theta J(\theta) = \mathbb{E}_{y \sim \pi_\theta(\cdot|x)} \left[r(x,y)\, \nabla_\theta \log \pi_\theta(y \mid x) \right]$$

In words: the update direction is an average over the model's own sampled answers of the reward times the direction that makes that answer more likely.

1. The average runs over the policy's **own samples** (REINFORCE, Williams 1992; PPO and GRPO inherit it).
2. Probability zero means **never sampled**: no term, no gradient, no update.
3. Rewards **reweight** what sampling reaches. They cannot mint probability for the unreachable.

$$\pi_{\theta_0}(y \mid x) = 0 \implies \pi_{\theta_t}(y \mid x) = 0$$

In words: what starts at zero probability stays at zero, at every step.

In practice "zero" means astronomically small: unreachable in any realistic budget. The KL penalty toward the base ties tighter still.

The Invisible Leash: the argument, formalized

PART 5 · COMPANION PAPER · WU ET AL., ARXIV 2507.14843

Theorem C.1: support preservation under RLVR

Theorem C.1 (Support Preservation under RLVR). *Let $\pi_\theta(y \mid x)$ be the RLVR-trained distribution obtained via standard on-policy gradient updates on verifiable rewards R. Then for all $x \in \mathcal{X}$,*

$$\mathrm{supp}(\pi_\theta(\cdot \mid x)) \subseteq \mathrm{supp}(q(\cdot \mid x)).$$

In particular, if $q(y^* \mid x) = 0$ for some correct solution y^*, then RLVR cannot discover y^*.

Corollary C.2: asymptotic sampling upper bound

Corollary C.2 (Asymptotic Sampling Upper Bound). *Let $pass@k_p(x)$ be the probability that at least one out of k samples $y_i \sim p(\cdot \mid x)$ is correct, i.e. $pass@k_p(x) = 1 - (\Pr_{y\sim p}[R(x,y) = 0])^k$. Under the conditions of Thm. C.1 and the sampling independence, we have*

$$\limsup_{k\to\infty} pass@k_{\pi_\theta}(x) \;\le\; \limsup_{k\to\infty} pass@k_q(x).$$

The crossover, restated as an inequality: as k grows, RLVR's pass@k is **bounded above by the base's**.

The same paper proves an **entropy and reward trade off**: sharpening toward rewarded outputs shrinks exploration. Precision is bought with diversity.

Its title ends in a question mark: the bound assumes sampling from the policy itself. Inject probability mass **from outside**, a teacher for instance, and the leash can loosen. That is Part 6.

QUICK FACT The authors span Stanford, UW and NVIDIA. The question mark is doing honest work.

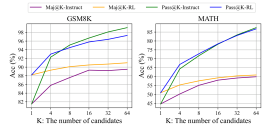
DeepSeekMath saw it first, in February 2024

PART 5 · PRIOR EVIDENCE · FROM THE TEAM THAT INVENTED GRPO

DeepSeekMath 5.2.2: "Why does RL work?"

5.2.2. Why RL Works?

In this paper, we conduct reinforcement learning based on a subset of instruction tuning data, and it achieves significant performance enhancement upon the instruction tuning model. To further explain why reinforcement learning works. We evaluate the Pass@K and Maj@K accuracy of the Instruct and RL models on two benchmarks. As shown in Figure 7, RL enhances Maj@K's performance but not Pass@K. These findings indicate that RL enhances the model's overall performance by rendering the output distribution more robust, in other words, **it seems that the improvement is attributed to boosting the correct response from TopK rather than the enhancement of fundamental capabilities.** Similarly, (Wang et al., 2023a) identified a **misalignment problem** in reasoning tasks within the SFT model, showing that the reasoning performance of SFT models can be improved through a series of preference alignment strategies (Song et al., 2023; Wang et al., 2023a; Yuan et al., 2023b).



DEEPSEEKMATH FIG. 7: RL IMPROVES MAJ@K BUT NOT PASS@K, ON GSM8K AND MATH

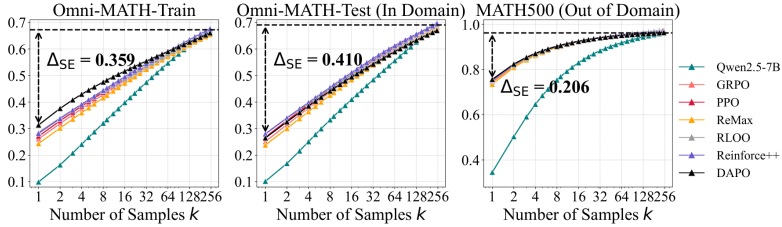
$$\hat{A}_i = \frac{r_i - \text{mean}(r_1, \ldots, r_G)}{\text{std}(r_1, \ldots, r_G)}$$

In words: score each answer by how far its reward sits above the group average, in units of the group spread.

QUICK FACT Eleven months separate this buried observation from the R1 moment it powered.

Six algorithms, one ceiling

PART 6 · ROBUSTNESS · “YOU USED THE WRONG ALGORITHM”



YUE ET AL. 2025, FIG. 8 TOP: PPO, GRPO, REINFORCE++, RLOO, REMAX, DAPO. VERL REIMPLEMENTATIONS ON OMNI-MATH-RULE.

$$\Delta_{\text{SE}} = \text{pass@}256\ (\text{base}) - \text{pass@}1\ (\text{RL})$$

In words: the coverage the base reaches in 256 tries, minus what RL delivers in one try. Lower is better.

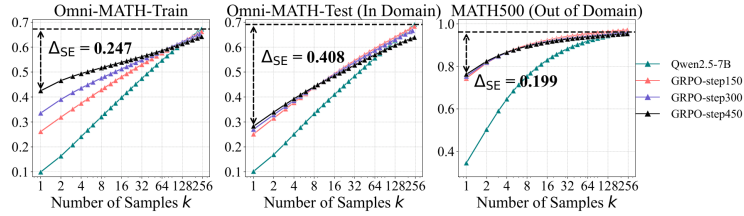
It stays **above 40 points** in domain for every algorithm.

The limitation is the objective itself.

QUICK FACT The spread across all six algorithms is about one point: GRPO 43.9 versus RLOO 42.6.

Training longer narrows the boundary

PART 6 · ROBUSTNESS · “JUST TRAIN LONGER”



YUE ET AL. 2025, FIG. 8 BOTTOM: GRPO AT 150, 300 AND 450 STEPS

From step 150 to 450, training pass@1 climbs **26.1 to 42.5** while pass@256 **declines**. The metric being optimized improves; the boundary being assumed contracts. More rollouts per prompt (8 to 32) softens but does not reverse this. Adding a KL penalty makes it worse.

QUICK FACT Dashboards cannot show the narrowing unless they log pass@k at large k.

What could push a model past its base?

INTERLUDE · ASK THE ROOM

More data? **A larger model?** **Distillation?**

The theory just told us what to look for: the leash binds learning from the policy's *own* samples. What operation places probability mass on sequences the model would never sample by itself?

QUICK FACT The answer is also the paper's control experiment, which is what makes the whole methodology trustworthy.

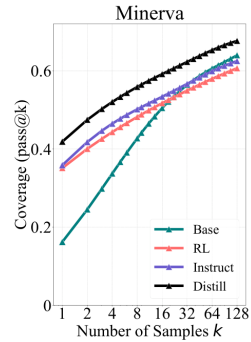
Distillation crosses the line RL cannot

PART 6 · THE CONTROL CONDITION

DeepSeek-R1-Distill-Qwen-7B exceeds its base at **every** k: the whole curve shifts upward.

A teacher's traces are data from **outside the student's policy**. Supervised training on them places probability on token sequences the student would never sample. That is precisely the operation the RLVR gradient cannot perform.

> *A reward cannot reach a path never taken.*
> *A teacher can put the path there.*



YUE ET AL. 2025, FIG. 7: MINERVA. BASE, RL, INSTRUCT, DISTILLED

QUICK FACT Decisive control: pass@k can register expansion, and distillation shows it.

Three serious rebuttals, and a softened claim

PART 6 · OPEN QUESTIONS, 2025 TO 2026

“Too few steps”

ProRL (2505.24864): thousands of steps, KL control, reference resets. Reports wins at all k where the base fails. Replication ongoing.

STATUS: CONTESTED

“Wrong ruler”

CoT-pass@k (2506.14245): answer only grading over credits guesses; grade the reasoning too and RLVR wins at every k. Counter: the audit and the code results.

STATUS: UNRESOLVED

“A Qwen artifact”

Spurious Rewards (2506.10947): **random** rewards lift Qwen Math about 21 points on MATH500 (true: 29). LLaMA and OLMo gain nothing. Much evidence is Qwen centric.

CUTS BOTH WAYS

QUICK FACT v1 said “RLVR”; v5 says “current RLVR methods”. Also: random rewards helped Qwen. Sit with that.

Three operational consequences

PART 7 · PRACTICE

Budget vs boundary

With a verifier and a test time budget, **base plus wide sampling** rivals the RL model on coverage. 256 chains is a real GPU bill: choose by economics, not defaults.

Capability needs a teacher

To add abilities: a stronger base, or distillation from a stronger teacher. Use RLVR afterwards for what it demonstrably does: compress the boundary into pass@1.

Audit benchmarks

pass@1 conceals ceilings. When a release claims RL "taught" reasoning, ask for the **pass@k curve**, and check whether the evidence is Qwen only.

The authors' own forward agenda: stronger exploration, deliberate data curation, process rewards, and agent interaction over many turns. Routes to an RL that might genuinely expand.

QUICK FACT 256 chains at 16,384 tokens is about four million tokens per problem.

One map of post training

PART 7 · SYNTHESIS · RECONCILING WITH TALK 1

DPO

shapes **preferences**: which existing behavior surfaces

RLVR

buys **reliability**: one shot access to existing capability

Pretraining, distillation

create **capability**: the support everything else reweights

Complementary operations on the same distribution. Two of them reweight it. Only one of them extends it.

QUICK FACT This also settles tonight's friendly feud: the DPO talk and this one were describing two different reweighting operators.

Epilogue: the boundary since April

PART 8 · 2025 TO 2026 · SO HOW DID MODELS GET SO GOOD?

FunSearch, AlphaEvolve

New mathematics from an LLM inside an **evolutionary search loop**: a 48 multiplication algorithm for 4×4 matrix products, first progress in 56 years, plus improved bounds on open problems.

ESCAPE: NOVELTY STORED OUTSIDE THE WEIGHTS

IMO 2025 gold

Gemini Deep Think (officially graded) and an OpenAI experimental model reach **gold level** on the 2025 Olympiad.

ESCAPE: NEW BASES, PROCESS REWARDS, PARALLEL SEARCH

A verified new bound

GPT-5 Pro improved a constant in an open convex optimization problem; humans checked it.

ESCAPE: WIDE SAMPLING PLUS EXPERT VERIFICATION

The ceiling was real. The answer was not more GRPO. The field changed the game.

CLOSING

If RL only concentrates what pretraining created, where will new machine reasoning come from?

And the mirror question: how much of human practice is also sharpening rather than invention?

READING LIST, ARXIV

the paper 2504.13837
The Invisible Leash 2507.14843 · DeepSeekMath 5.2.2 2402.03300 · ProRL 2505.24864
CoT-pass@k 2506.14245 · Spurious Rewards 2506.10947 · FunSearch, Nature 2023

QUICK FACT DeepSeekMath quietly published tonight's finding eleven months before it became a NeurIPS award winner. Read section five of papers.

Same claim. Vote again.

“RL fine tuning teaches models new reasoning, beyond their base.”

True or false, thirty minutes later?

`github.com/Akasxh` · `/in/s-akash-project-gia` · `akasxh.github.io`

GITHUB LINKEDIN

SLIDES AND READING LIST: SCAN EITHER CODE, OR ASK ME.

QUICK FACT Watching the recording? Tell us in the comments which side you are on now.