Adaptive Apprenticeship: Small Average Gain, Inconsistent Lineages

Abraham October 2026

Abstract

Can a stronger teacher improve a specialized student by adapting its teaching to the student's mistakes? Three paired training lineages received two rounds of static or feedback-guided curricula, with matched numbers of problems and weight updates. Final complete-answer accuracy was 172/384 (44.79%) for static teaching and 180/384 (46.88%) for adaptive teaching. The mean paired difference was +2.08 percentage points, but only one of three lineages improved. This exploratory result does not establish a reliable teaching advantage, absolute improvement over the starting checkpoint, or recursive self-improvement.

Question and design

The student was Qwen2.5-7B-Instruct with previously trained 768-update parent adapters. These were not untouched base models. Each of three parent lineages was branched into static and adaptive teaching conditions. Both received 144 new problems and 48 replay problems per round, with separate answer and explanation supervision: 384 additional updates per round, or 768 across two rounds. Learning persisted in adapter weights rather than additional evaluation-time lesson prompts.

The teacher could use practice feedback to choose both problems and explanations in the adaptive arm. Static teaching did not receive that student-specific feedback. Problem counts and training updates were matched; actual tokens and execution costs were not assumed equal. This tests the combined teaching policy, not explanation quality in isolation.

The primary endpoint required the entire answer: every safe route, its cost, and the selected minimum. Cheapest-route correctness was secondary. Each final adapter answered the same fixed 128 problems: 96 familiar-size and 32 larger structural-transfer tasks.

Primary results

Training lineage	Static complete answer	Adaptive complete answer	Adaptive minus static
11	46/128 (35.94%)	62/128 (48.44%)	+12.50 pp
23	61/128 (47.66%)	59/128 (46.09%)	-1.56 pp
47	65/128 (50.78%)	59/128 (46.09%)	-4.69 pp
Pooled descriptive accuracy	172/384 (44.79%)	180/384 (46.88%)	+2.08 pp

The experimental replication unit is the paired training lineage, not each question. More evaluation questions do not create more independent training replications. The large positive difference in lineage 11 was partly offset by two negative differences. No population significance, superiority, or equivalence claim is made.

Cheapest-route accuracy moved in the opposite direction: 287/384 (74.74%) for static and 281/384 (73.18%) for adaptive. Adaptive teaching produced 101 answers with the correct minimum but an incomplete or incorrect full answer. Reporting only route selection would obscure the primary failure mode.

Improvement over time

The same 48-question monitor was evaluated at the starting parent and after each round. Values below are means across the three lineages. Update counts refer to additional training after the already-trained parent.

Condition	Starting parent: 0 additional updates	Round one: 384	Round two: 768
Static	51.4%	52.8%	49.3%
Adaptive	51.4%	51.4%	59.7%

Monitor improvement is not the final endpoint. The starting checkpoints were not measured on the primary 128-question final pool, so final absolute improvement over those checkpoints cannot be computed. The monitor and final pools must not be substituted for each other or selected after seeing which looks better.

Transfer, efficiency and audit

Larger-task adaptive-minus-static differences were +12.50, -9.38 and +9.38 points across the three lineages. Familiar-size differences were +12.50, +1.04 and -9.38 points. These mixed results do not establish consistent structural transfer.

The complete study contained 4,608 training updates and 2,208 completed local generations. An independent recorded-data replay corroborated the complete graph, primary scores, derived analysis, and unchanged result artifacts. It accepted the study with explicit historical limitations; this is an audit of recorded evidence, not an external scientific replication or peer review.

Interruptions required audited recovery. A saved 384-update adapter and completed evaluations were preserved, then only unfinished evaluation was continued; that adapter was not blindly retrained. Two historical generation attempts have unknown outcomes and remain counted separately, yielding 2,210 physical attempts. Teacher and reviewer overhead, failed work and unknown usage are not free. No deployment break-even, wall-time advantage or efficiency superiority is established here.

Limitations and next experiment

Three related training lineages, one synthetic task family, one teaching pipeline and two rounds. This is exploratory specialization, not general intelligence.

Historical execution was heterogeneous. Resource-tail coverage and durability gaps remain disclosed; recovery does not erase them.

No starting-parent measurement on the primary final pool. Within-monitor curves cannot supply that missing comparison.

Feedback changed problem selection and explanations together. Their individual contributions are not identified.

No strong evidence that adaptive teaching reliably beats static teaching; negative lineages and secondary regressions remain part of the result.

A fresh follow-up is being prepared with granular practice-error feedback, common task guidance for both teachers, starting-parent measurement on a fresh final pool, and the same matched-dose static/adaptive comparison. It has not begun scientific training or answer generation. Its result will be reported separately rather than replacing this study.

Evidence record

Final numerical analysis SHA-256:
db3c3bbc76650ec659533a527ccc7d9b59d3d77b93ecbb5a1ff16624d759053a.

The complete graph and analysis replay passed with historical limitations, recorded 1 October 2026. Public snapshots omit raw task answers, credentials, personal filesystem paths and identifying device details. A companion essay remains for the author to supply.