University of Nottingham
UK | CHINA | MALAYSIA

A Test and Training Platform Utilizing a Raspberry Pi 5:

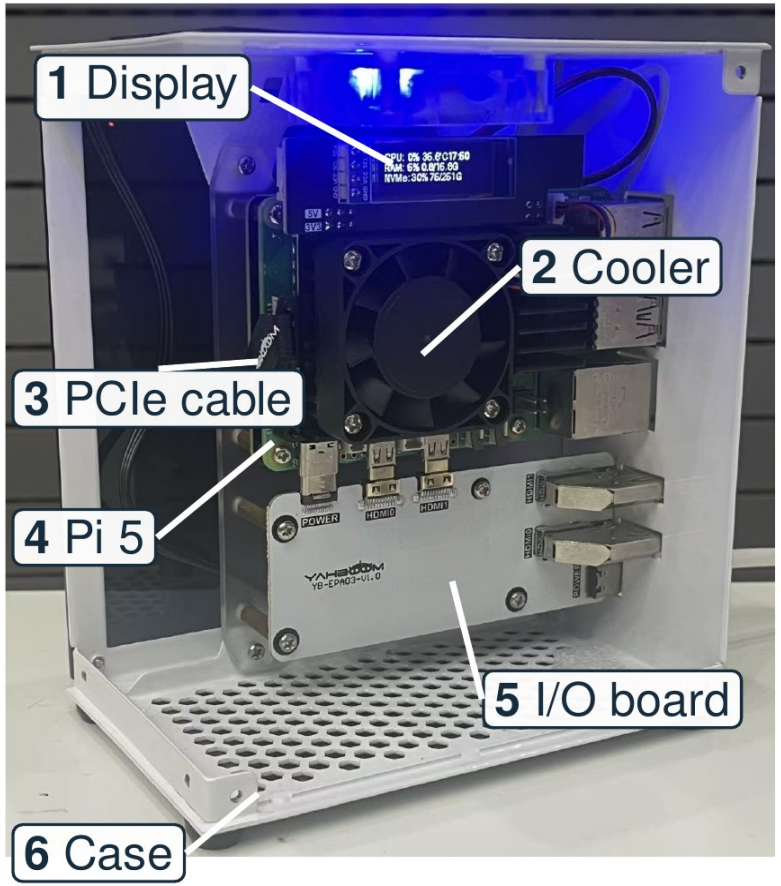
AI-Retrieval, Tools and a Compact LLM

Local inference for course practice
Design goals: less cloud reliance, more control over cost and data

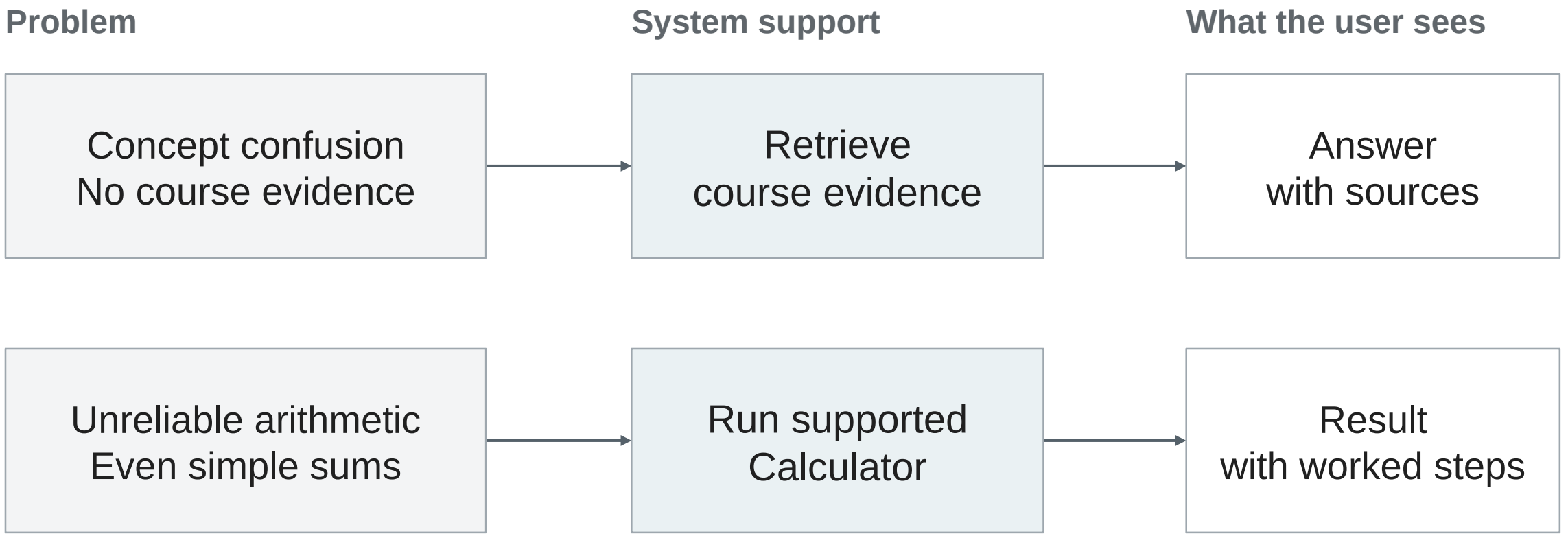
Siheng Ye, Xiaotong Li, Chang Qi,
Linfei Yin, Yaokun Zhang

Supervisor: Ian Pratt

University of Nottingham Ningbo China
2026 Summer FURP

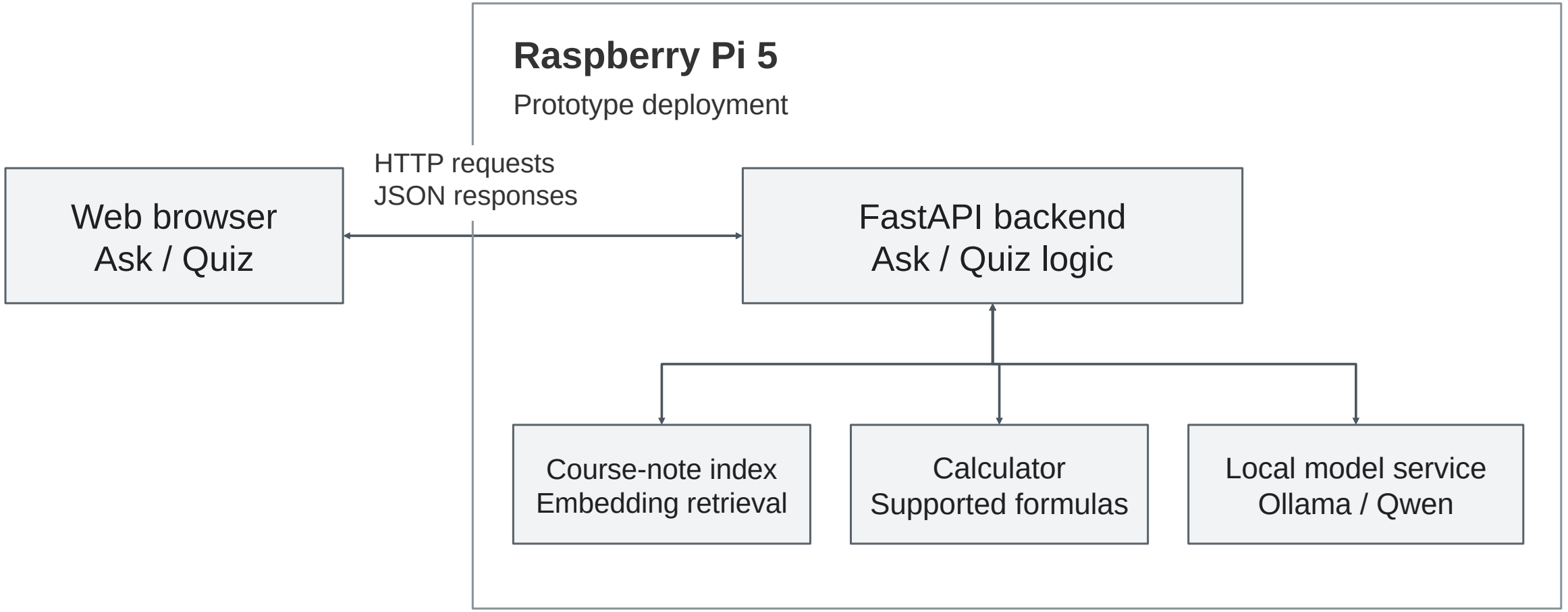


Engineering questions



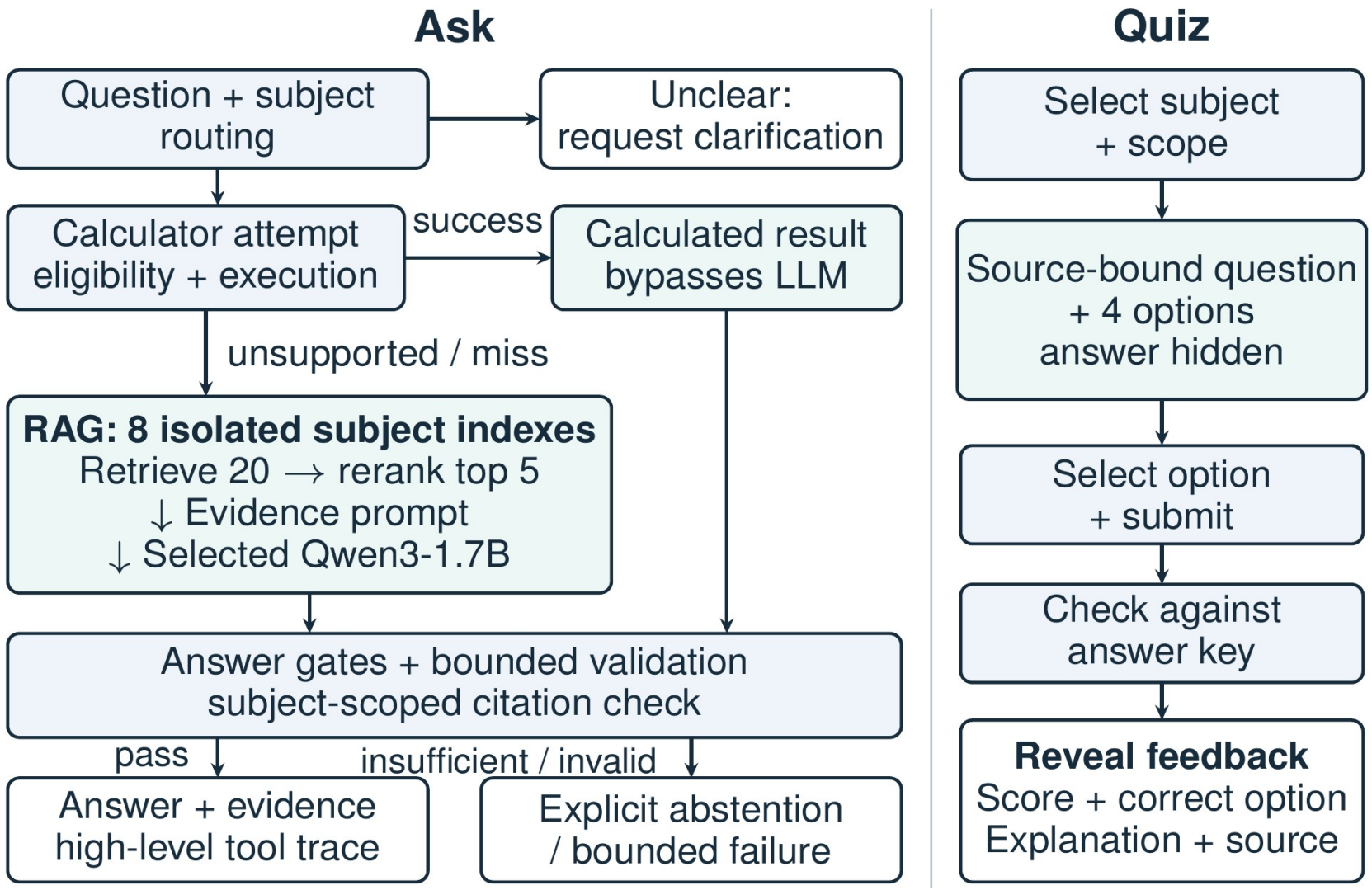
Design motivation: course concepts and calculation.
Formal component evaluations follow.

Web and Raspberry Pi



Prototype architecture. Pi quality results cover Base inference only.

System workflow



Ask and Quiz

@PE A balanced three-phase load draws an RMS line current of 12 A at an RMS line-to-line voltage of 415 V with a power factor of 0.85 lagging. Calculate the total real power and show the calculation steps.

Aug 27, 01:06 PM

PE answer ok · preview Aug 27, 01:06 PM

P = 7331.77 W.

Worked solution

1. Use RMS line quantities: $V_L = 415\,\mathrm{V}$, $I_L = 12\,\mathrm{A}$, and power factor $\cos\varphi = 0.85$.
2. For a balanced three-phase load, total real power is $P = \sqrt{3} V_L I_L \cos\varphi$.
3. First find total apparent power: $|S| = \sqrt{3} V_L I_L = \sqrt{3} \times 415 \times 12 \approx 8625.61\,\mathrm{VA}$.
4. Apply the power factor: $P = |S| \cos\varphi = \sqrt{3} \times 415 \times 12 \times 0.85 \approx 7331.77\,\mathrm{W}$.
5. Therefore, total real power is $P \approx 7331.77\,\mathrm{W} \approx 7.332\,\mathrm{kW}$. Calculations use unrounded values; displayed powers are rounded.

Question 1 of 1 · Score 1

Which course-note section contains this exact excerpt: "In a balanced star connection, line voltage is $\sqrt{3}$ times phase voltage and line current equals phase current."?

- A. Circuit Quantities — Charge, current, voltage, power and energy
- B. Ideal Transformer Model — Turns ratio and referred impedance
- C. Transmission and Distribution — Voltage level, current and losses
- D. Balanced Three-Phase Systems — Star and delta relationships — Correct

See Results | Exit Quiz

WHY THIS IS CORRECT

The excerpt is recorded under "Balanced Three-Phase Systems — Star and delta relationships" at the exact heading "Three-Phase and Power Systems — Balanced Three-Phase Systems — Star and delta relationships — Definition / Principle": In a balanced star connection, line voltage is $\sqrt{3}$ times phase voltage and line current equals phase current.

REFERENCE NOTE

Three-Phase and Power Systems · Balanced Three-Phase Systems — Star and delta relationships — Definition / Principle

Ask: a complete question and worked answer

Quiz: answer first, then see feedback

Supervised fine-tuning (SFT)

Setting	Initial SFT	Expanded SFT
Training records	801	2,400 of 4,000 candidates
Epochs	3	1
Learning rate	2×10^{-4}	1×10^{-4}
Effective batch (records per update)	4	16 (batch 4 × accumulation 4)

Shared LoRA: query/value projections, rank 8, alpha 16, dropout 0

AdamW optimizer. Sequence length: 2,048 tokens. Same parent Qwen3-1.7B.

Train LoRA in BF16, merge the adapter, then convert to Q4_K_M for testing.

The recipes change several settings. This does not isolate the effect of data volume.

Benchmark protocol

Question type	Count
Concept	28
Retrieval integration	28
Supported calculations	24
Unsupported calculations	8
Insufficient evidence	8

96 public questions across 8 EEE courses
Year 2 and Year 3 of our programme
7 configurations produce 672 responses

Fixed questions, corpus and scoring rubric.
Match prompts within SFT and device pairs.
One run per question and configuration.

Generation settings

Greedy decoding (T = 0), seed 20260826
Context 2,048, output limit 256 tokens
8 threads, all-layer GPU offload

Offline: BM25 selects 20 passages, reranked to 5.
The Web interface uses separate embedding retrieval.

Separate runs: 24 retrieval questions, 96 paired GPU/Pi Base answers,
and 8 pre-rendered timing prompts per device.

Answer scoring

Semantic scoring: add components, apply the cap, then divide by 10

Component	Points	Full-credit requirement
Core correctness	0–4	The main answer is correct
Coverage	0–3	All required information is present
Consistency	0–2	Reasoning and evidence agree
Relevance	0–1	The answer addresses the question

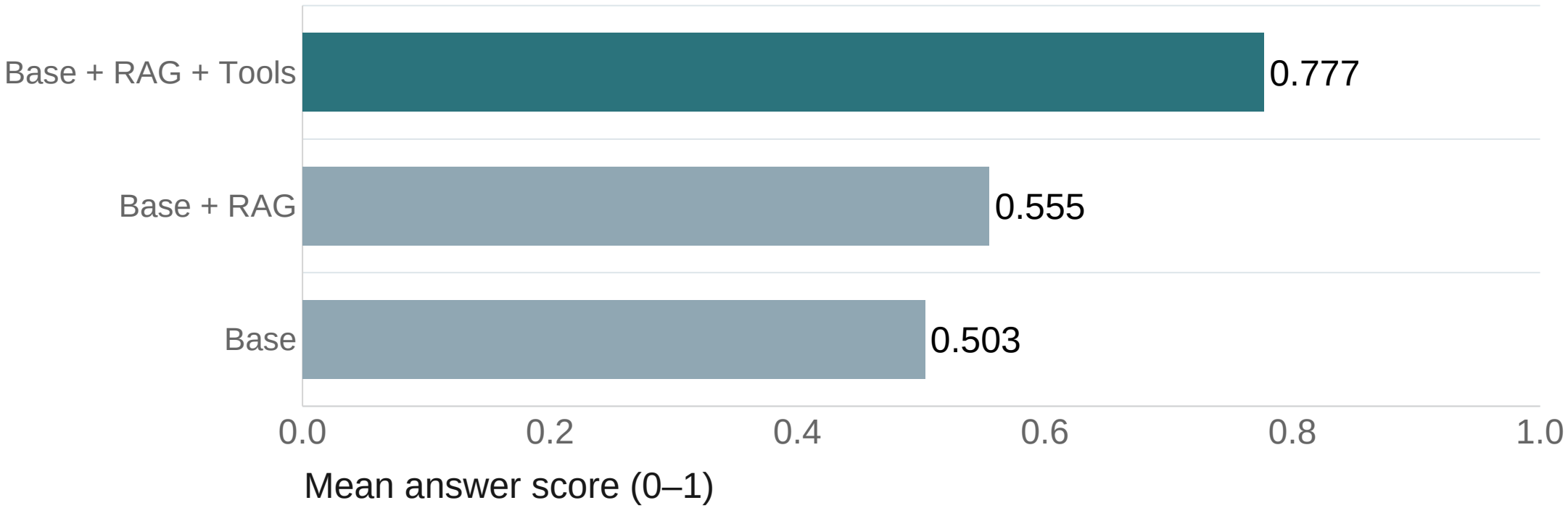
Core ≤ 1: total capped at 4/10. Core = 2: capped at 6/10.

Objective items: exact match or numerical tolerance, scored 0 or 1.

Fully correct semantic answer: core 4, coverage 3 and consistency 2.

ChatGPT supports practical first-pass content grading, followed by review and first-author audit.
No independent assessor. Mean scores and fully-correct counts are separate outcomes.

Overall answer quality



Base: model alone. RAG: course notes in the model's input.
Tools: Calculator for supported tasks, course notes for other questions.

96 questions on GPU. Best mean score: 0.777. Fully correct: 54/96.

Retrieval and calculation results

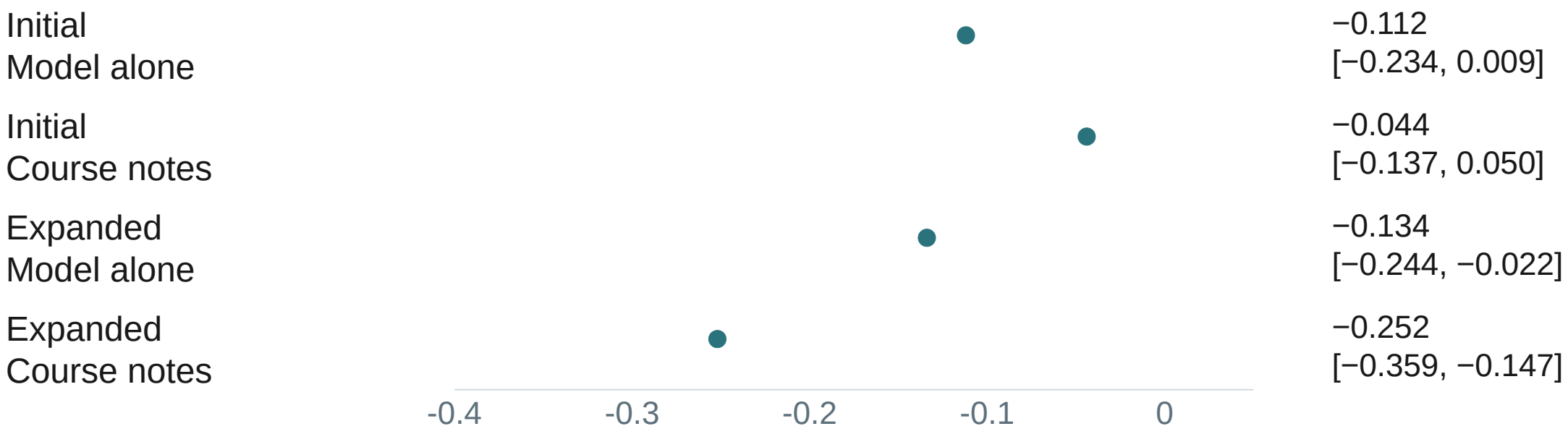
Test	Comparison	Result
Retrieval 24 separate questions	Base + RAG versus Base	Mean-score change +0.454 95% CI [0.350, 0.558]
Supported calculations 24 main-set questions	Calculator versus Base + RAG	24/24 fully correct versus 0/24

The two rows use different sets of 24 questions.

The calculation result covers preclassified supported tasks.
Routing on unseen questions and the separate Web retrieval path remain untested.

Fine-tuning results

SFT minus matched Base on 64 questions. Dots show means, lines show 95% CIs.



Initial intervals include zero. Expanded scores are lower in these tests.

Same 64 questions and matched prompts/passages. The subset excludes 24 supported calculations and 8 insufficient-evidence items.

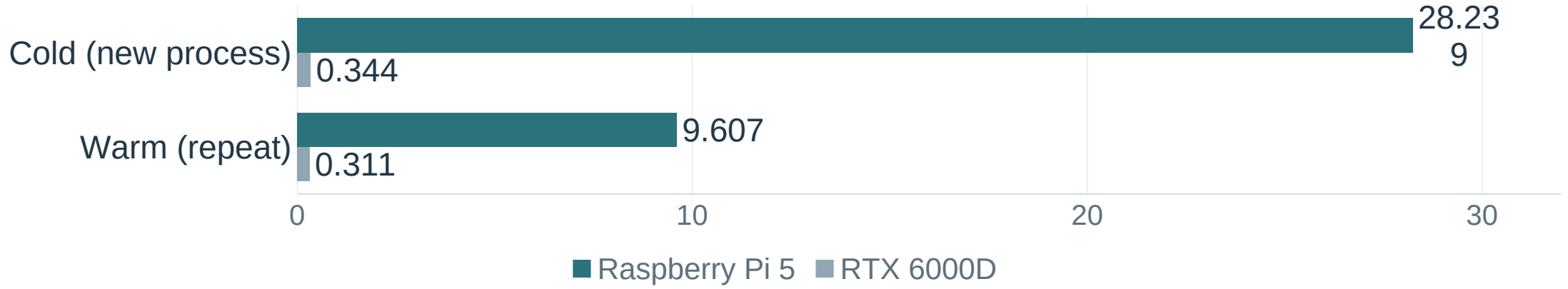
GPU and Raspberry Pi

Answer quality: 96 paired questions, Base model alone

Mean score (0–1)	RTX 6000D	Raspberry Pi 5
Same questions and rubric	0.503	0.498

Pi minus GPU: −0.005, 95% CI [−0.041, 0.032]. Equivalence remains unproven.

Median model-request time (s): 8 prompts/device, 1 cold and 2 warm runs



Same Base Q4_K_M model and matched prompts. Context 2,048, output cap 256, T = 0.

Timing excludes retrieval, Calculator and Web processing. Warm runs may reuse caches.

Conclusions and next steps

Main findings

Retrieval + Calculator helped

0.777 mean; 54/96 fully correct
(GPU, 96 questions).

Neither training setup showed a gain

64 questions; same inputs in each pair.

Pi ran Base, with slower responses

96 model-only questions completed.
Similar scores do not prove equivalence.

Limits and next steps

Finding a source is not enough

Base + RAG: files found 28/28;
main-answer errors 9/28.

Calculator supports fixed tasks

17 task types; new requests need testing.

Users and the full Pi app

No usability or learning study.
Pi quality covers Base inference only.

Next: independent user tests and quizzes targeting learning gaps.

Thank you