KAERI 한국원자력연구원 Korea Atomic Energy Research Institute

HANYANG UNIVERSITY 한양 1939

Applied AI Section

# Local AI Models for the Reactor Operator Fundamentals Examination

## Comparison of full fine-tuning and LoRA adaptation

Isak Hwang, Yoonpyo Lee, Hogeon Seo

**Hanyang University · KAERI · UIUC · UST**

KNS 2026 Autumn Meeting | Gyeongju, October 28–30

# 1 Outline


| | | |
|---|---|---|
| **01** | Research question and comparison | 03–04 |
| **02** | Data and training | 05–10 |
| **03** | Answer extraction and evaluation | 11–12 |
| **04** | Results and interpretation | 13–16 |
| **05** | Conclusions and future work | 17–19 |

# 1 Research motivation and question

## Do medical and legal exam results extend to nuclear fundamentals?

### Medicine [1]

Flan-PaLM: 67.6% on MedQA
USMLE-style questions

### Law [2]

GPT-4: passing score in UBE evaluation
Uniform Bar Examination materials

## This study: evaluate local models on nuclear fundamentals

### NRC GFE

Question + options

→

### Local AI model

Compare training configurations

→

### Score vs. key

Compare with 80%

**Why NRC GFE? Public questions, official keys, and an 80% benchmark**

[1] Singhal et al. Nature 620, 172–180 (2023). doi:10.1038/s41586-023-06291-2

[2] Katz et al. Phil. Trans. R. Soc. A 382, 20230254 (2024). doi:10.1098/rsta.2023.0254

# 1 Models and adaptation strategies

**SFT / CPT: learning objective. Full fine-tuning / LoRA: weight updates.**

### Full fine-tuning · 4B

Update all model weights

### LoRA · 31B

Freeze base weights; train adapters

| Model | Configuration | Training setup |
|---|---|---|
| Gemma 3 4B-IT | SFT only | SFT with full weight updates |
| Gemma 3 4B-IT | CPT + SFT | Full-example text training, then SFT |
| Gemma 4 31B | Base | Public instruction-tuned baseline |
| Gemma 4 31B | SFT with LoRA | Freeze base weights; train adapters |

31B LoRA denotes SFT performed with LoRA throughout the results.

# 2 Data sources and classification

**NRC: questions and keys. DOE: references for training rationales.**

**NRC question bank**

PWR 2,140 + BWR 2,149

→

**Training candidates**

4,289 items → overlap filtering

**NRC March exams**

2015–2021 + **official keys**

→

**Evaluation QA**

**700** items: 350 PWR + 350 BWR

## DOE handbooks → reference passages for training rationales

Track reactor type, subject, image, year, and question ID

NRC: U.S. Nuclear Regulatory Commission. DOE: U.S. Department of Energy.

# 2 Training data and overlap removal

**4,289 candidates → 3,577 training examples**

**−711 matching entries** **−1 failed rationale**

**4,289**
NRC question-bank items

→

**3,578**
After exact-match filtering

→

**3,577**
Final training examples

**Normalize case + remove non-alphanumeric characters**

**PWR 1,797 + BWR 1,780**

Text matching reduces exact overlap; paraphrases and pretraining exposure are not fully controlled.

# 2 Examination design and workflow

**Train with rationales. Test without answers or retrieval.**

## TRAINING

**3,577 examples**

Question + rationale + answer

→

**Model adaptation**

4B FFT / 31B LoRA

→

**Adapted model**

Ready for evaluation

## EVALUATION

**700 QA items**

Question + options + image

→

**Model response**

**No keys** or external retrieval

→

**Extract and score**

Compare with **official key**

Aggregate accuracy and pass/fail for an individual examination are different measures.

# 2 Rationale generation and supervised learning

**Teacher uses answers and references. Evaluation excludes both.**

**Teacher input**

Question + **official answer**
Relevant DOE passages

**Gemini 3 Flash**

Generate rationale + final option

Training pair

**Student input**

Question + options + image

**SFT target**

Rationale + final answer

**This SFT setup computes loss on rationale and answer tokens only**

No quantitative assessment of rationale validity is reported in this presentation.

# 2 4B full fine-tuning

**SFT only vs. CPT followed by SFT**

## Text length: 2,048 tokens

## 8 examples per update

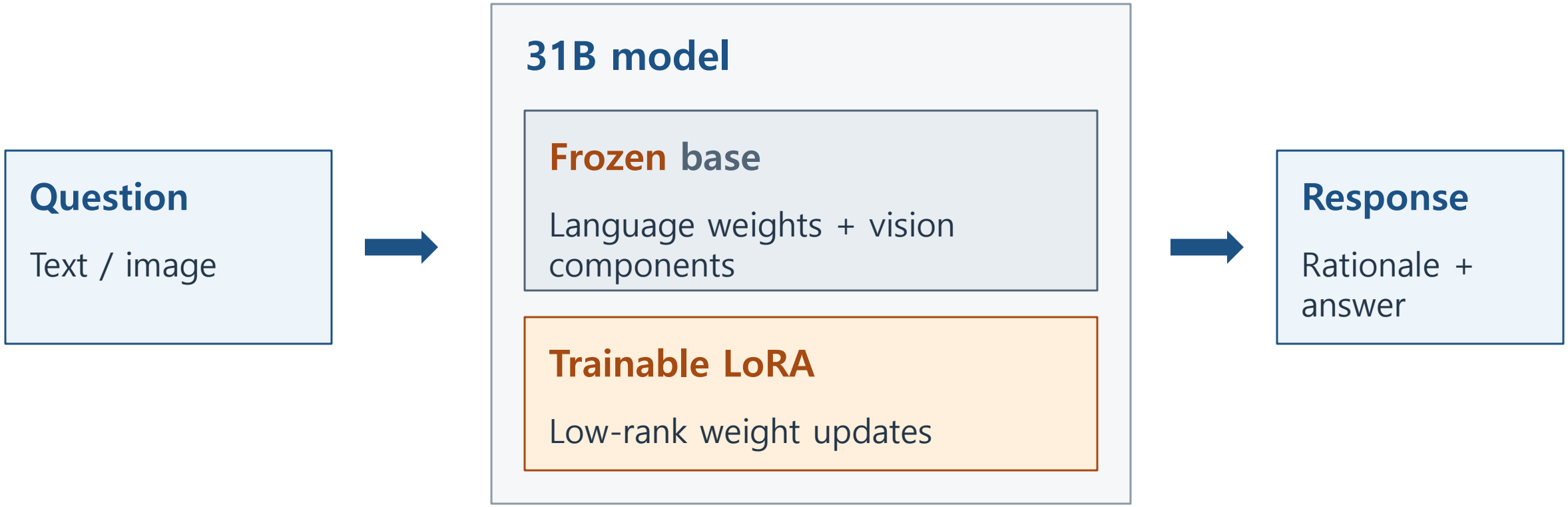
## 3 SFT training passes

| Comparison | SFT only | CPT + SFT |
|---|---|---|
| Sequence | SFT: 3 passes | CPT: 2 passes, then SFT: 3 |
| Learning target | **Rationale + final answer** | Full example, then response |
| Data used | Question–response pairs | Same examples: question + response |
| Shared setup | Update all model weights | Same SFT examples and settings |

CPT: continued pretraining on full examples. SFT: supervised fine-tuning on target responses.

# 2 31B supervised fine-tuning with LoRA

**SFT defines the target response. LoRA defines the trainable weights.**




**SFT on question–response pairs; 3 training passes**

Adapters are merged into the base model for inference.

# 3 Answer extraction and scoring

**Extract the final option, then compare with the key.**

**Model response**

Read final-choice expression

**Extract option**

Fallback pattern if needed

**Compare extracted option with the key**

## Illustrative example

**“Final answer: B”**

Extract B → key B → correct

**Extracted answer ≠ key: wrong**

**No readable option: invalid**

Proceedings §2.1: rule-based extraction of the final selected option. Detailed parser behavior requires code verification.

# 3 Accuracy and invalid responses

**Accuracy = correct / valid responses**

## Illustration: 10 responses




**6 correct** **3 wrong** **1 invalid**

### Reported accuracy

6 / (6 + 3) = 66.7%

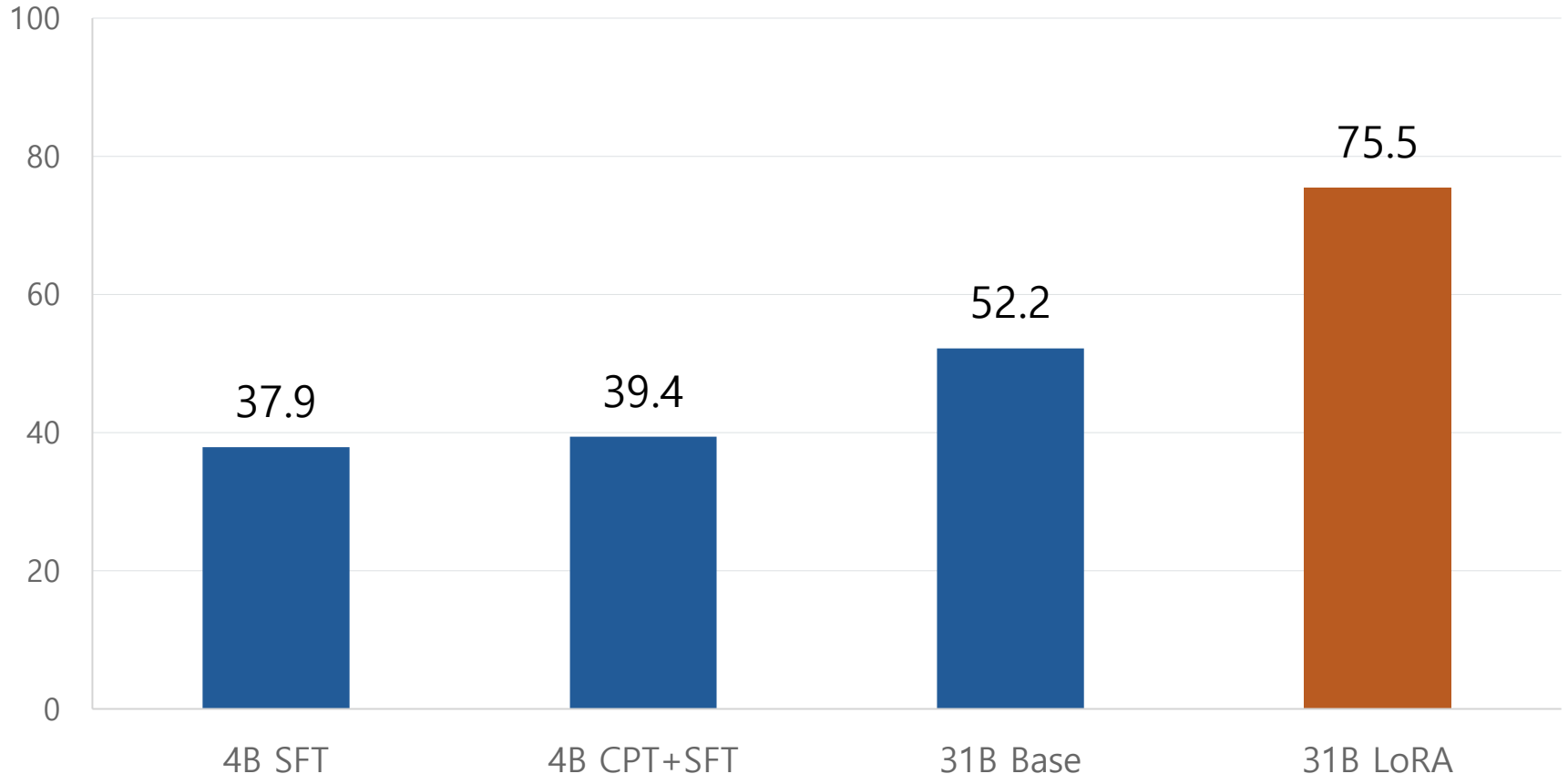
### All-item accuracy

6 / 10 = 60.0%

700 is the evaluation-set construction size. This presentation retains the proceedings scoring convention.

# 4 Results across four configurations

**Highest observed mean: 75.5%, below the 80% benchmark.**




## Observed mean

# 75.5%

**31B LoRA**
**4.5 pp below 80%**

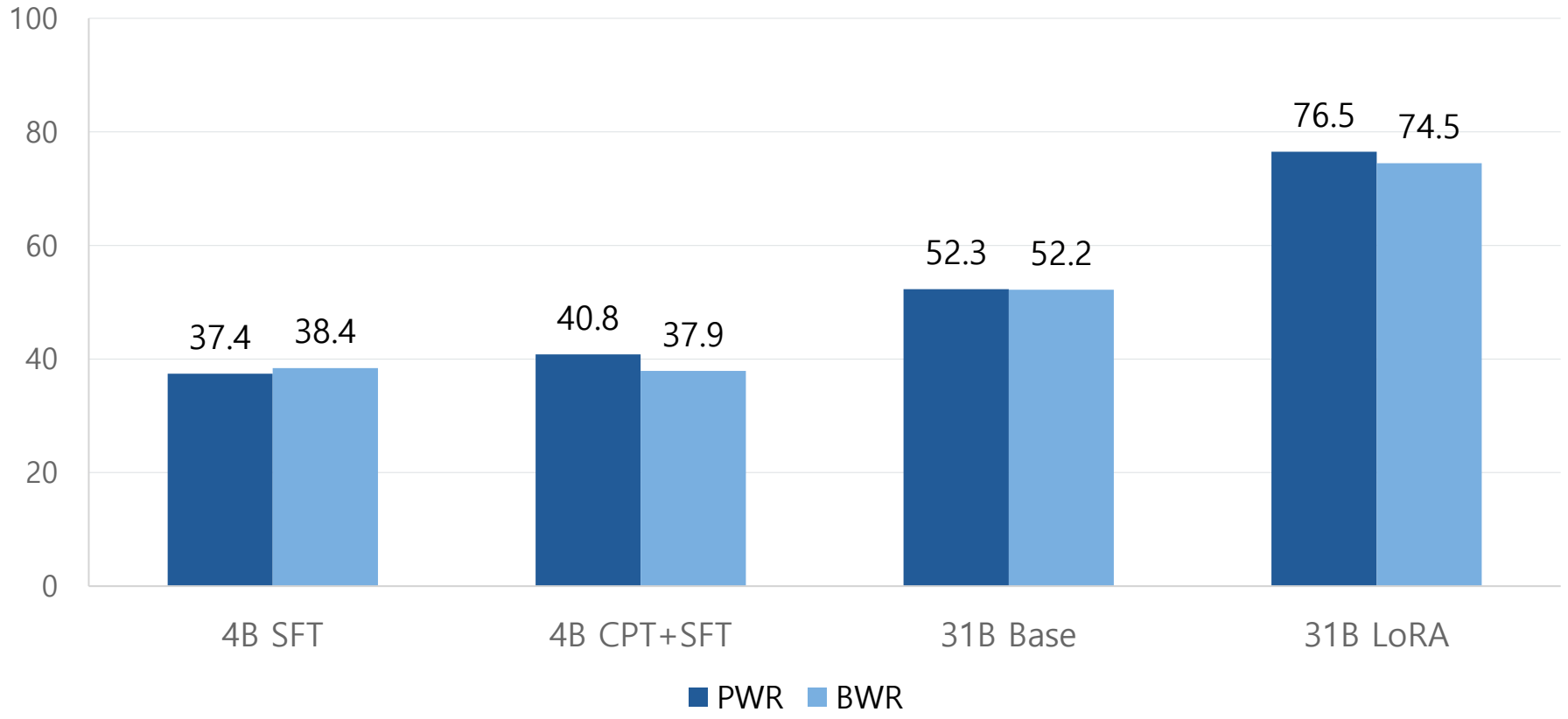
Means as reported in the proceedings

Source: KNS Autumn 2026 proceedings, Table I. Invalid responses excluded; reported values retained.

# 4 Accuracy by reactor type

**31B LoRA has the highest observed accuracy in both types.**

## Accuracy over valid responses (%)




### 31B LoRA

**PWR 76.5%**
**BWR 74.5%**

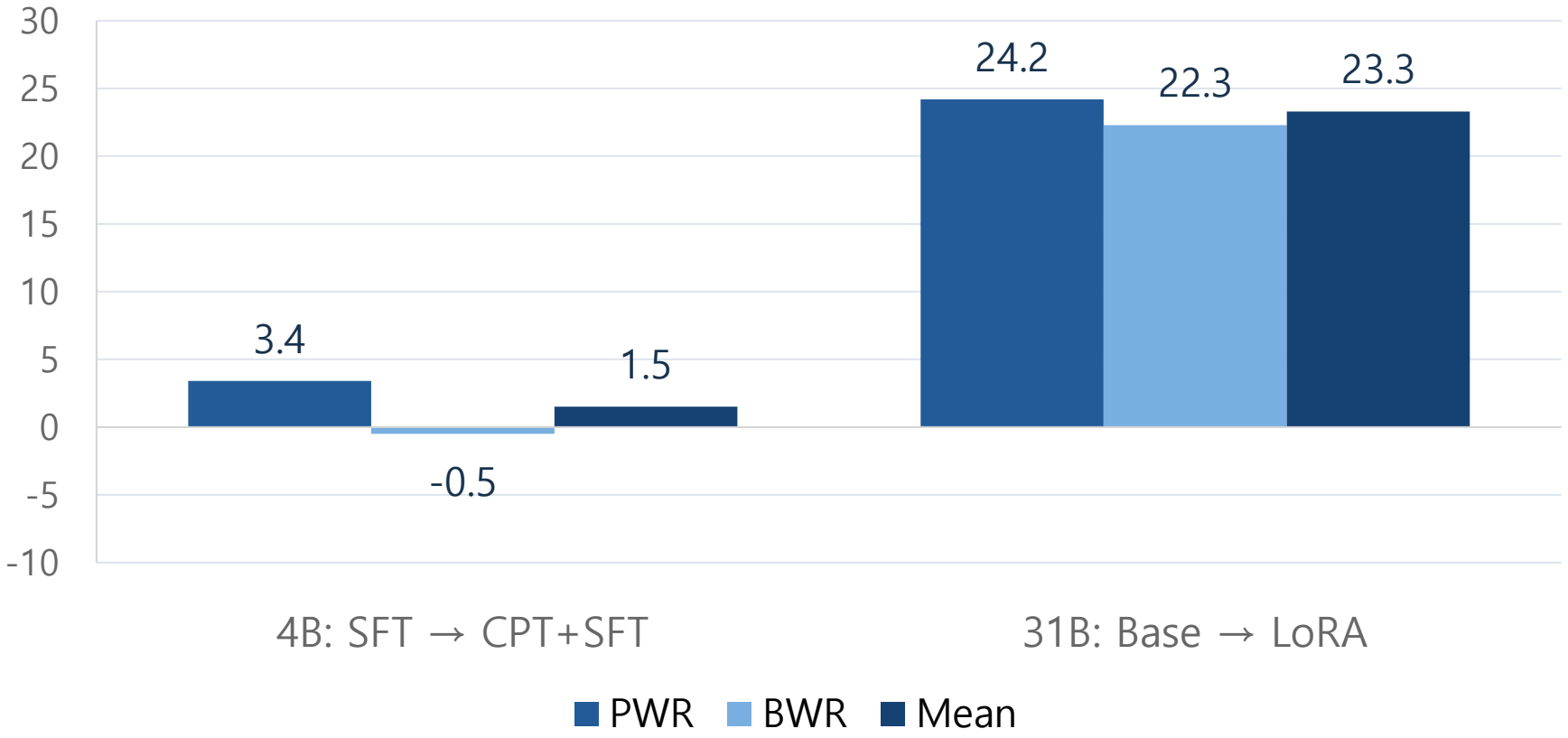
Both reactor types remain below **80%**.

Source: KNS Autumn 2026 proceedings, Table I. Accuracy over valid responses only.

# 4 Accuracy gains within each model

**31B: higher scores in both types. 4B: mixed CPT effects.**

## Change in accuracy (percentage points)




**31B: +23.3 pp**

Mean gain with LoRA

**4B: +1.5 pp**

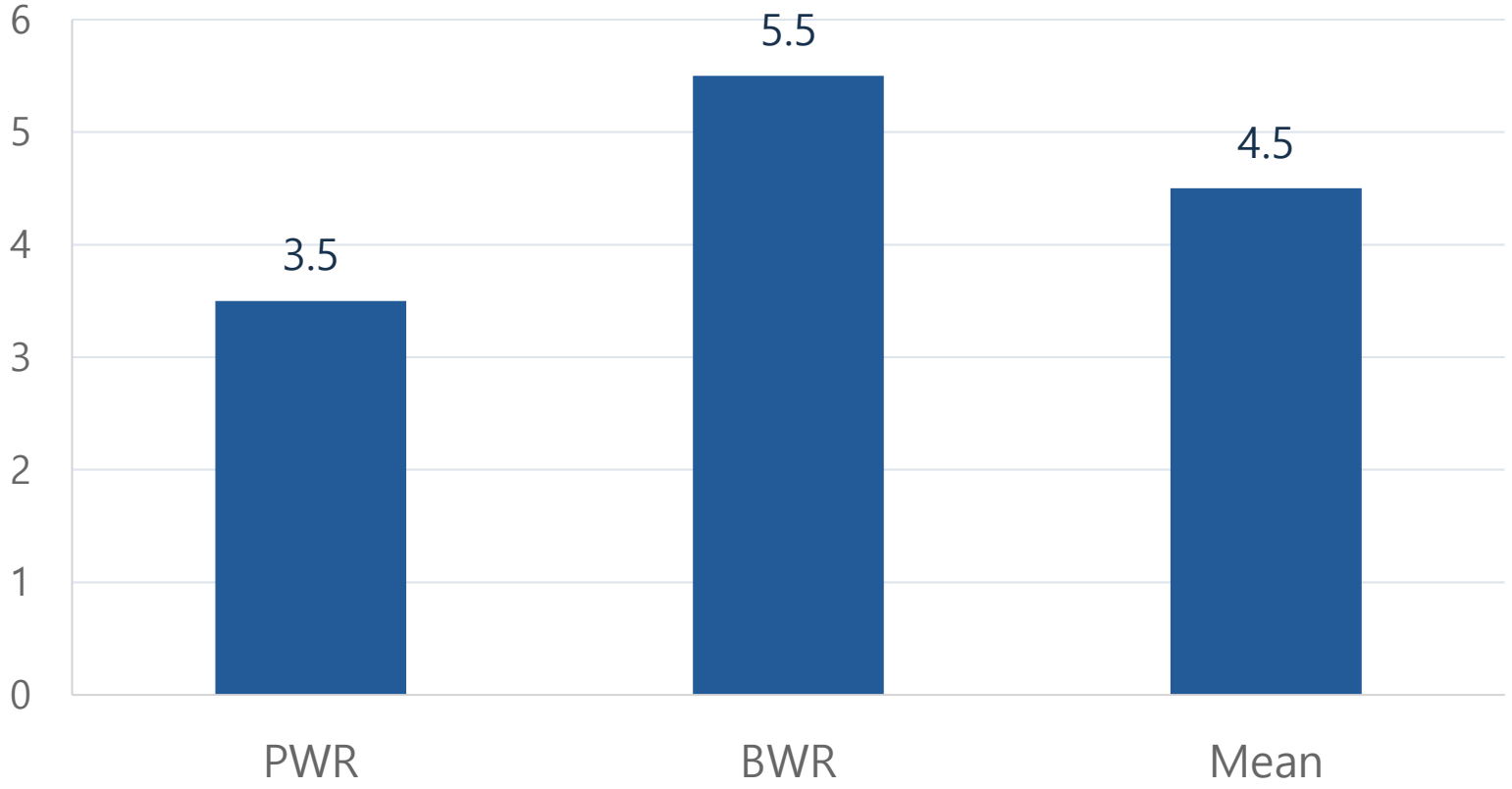
PWR improves
BWR declines slightly

Source: KNS Autumn 2026 proceedings, Table I. Accuracy over valid responses only.

# 4 Distance to the 80% benchmark

**Mean gap to 80%: 4.5 percentage points**

## Gap to 80% (percentage points)




## Current accuracy

| | |
|---|---|
| **PWR** | **76.5%** |
| **BWR** | **74.5%** |
| **Mean** | **75.5%** |

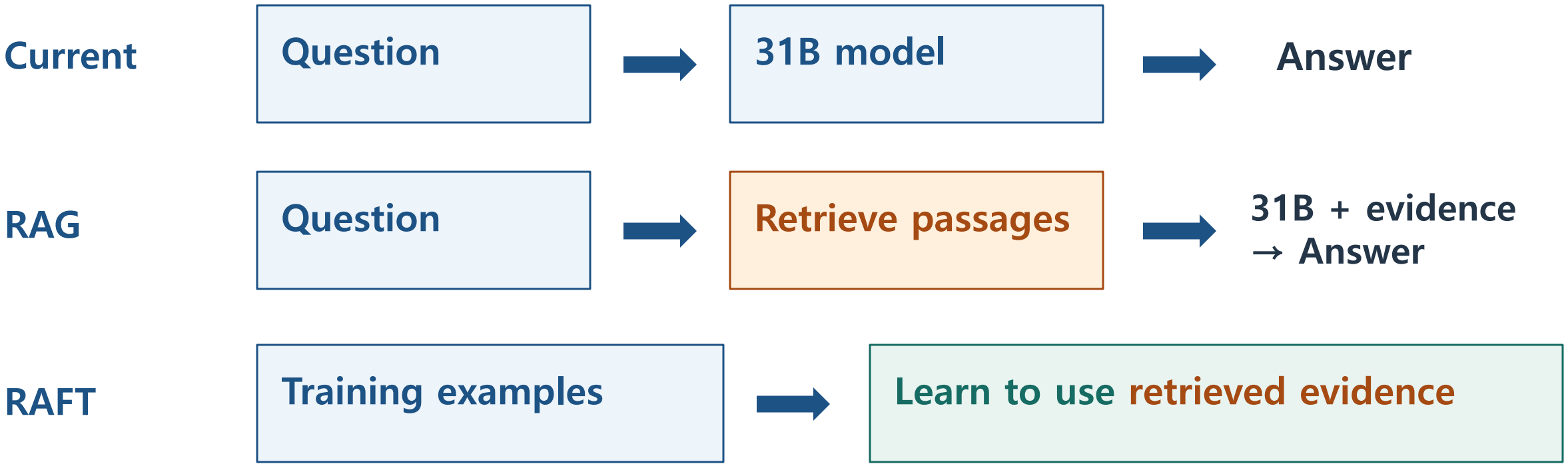
Per-exam passing and rationale validity require separate checks.

Source: KNS Autumn 2026 proceedings, Table I. Accuracy over valid responses only.

# 5 Planned follow-up with the 31B model

**Planned comparisons: data revision, retrieval, and evidence-use training.**

## PLANNED COMPARISONS




**Improve data first. Evaluate closed-book and retrieval settings separately.**

Planned work. Report closed-book and retrieval-enabled conditions separately.

# 5 Conclusions

**31B LoRA: 75.5% mean. All four means remain below 80%.**

| | |
|---|---|
| **Objective** | Find a local AI configuration that reaches **80%**. |
| **Finding** | 31B LoRA: **75.5%** mean. All four means <**80%**. |
| **Boundary** | Valid-response accuracy, not operational competence. |
| **Next step** | 31B follow-up: better data, RAG, and RAFT. |

NRC-based evaluation QA: 700 items. Reported scores follow the proceedings.

# Q&A