

Responsible NLP Checklist

Paper title: *LightReasoner: Can Small Language Models Teach Large Language Models Reasoning?*

Authors: *Jingyuan Wang, Yankai Chen, Zhonghang Li, Chao Huang*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- the authors indicated that the question does not apply to their work
- the authors did not respond to the checkbox question

For background on the checklist and guidance provided to the authors, see the [Responsible NLP Checklist](#) page at ACL Rolling Review.

A. Questions mandatory for all submissions.

- A1. Did you describe the limitations of your work?

This paper has a Limitations section.

- A2. Did you discuss any potential risks of your work?

Our work does not present any foreseeable societal, ethical, or safety risks beyond those commonly associated with general LLM research.

B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

- B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?

We used only publicly available math reasoning datasets that contain no personally identifiable or offensive content.

- B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?

We report dataset statistics and usage details in Section 3 and Appendix H.

C. Did you run computational experiments?

- C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?

We describe the experimental setup, hyperparameter settings, and training configurations in Section 3 and Appendix H.

- C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?

We report descriptive statistics, including accuracy improvements and efficiency metrics, in Section 3 and Appendix H.

The Responsible NLP Checklist used at ACL Rolling Review is adopted from NAACL 2022, with the addition of ACL 2023 question on AI writing assistance and further refinements based on ARR practice. ACL 2026 used a subset of ARR checklist form.

D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?

D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

We did not use human subjects or annotators; all experiments rely solely on publicly available datasets.

D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

No human participants or annotators were involved; therefore, no recruitment or payment was conducted.

D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?

We did not collect or use any data involving human participants; all datasets are publicly released and consented for research use.

D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

We did not conduct any human-subject or data-collection studies; all datasets used are publicly released for research purposes.

E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?

E1. If you used AI assistants, did you include information about their use?

We disclose our use of AI assistants for writing and phrasing assistance in Appendix B.