

Responsible NLP Checklist

Paper title: *Self-Awareness before Action: Mitigating Logical Inertia via Proactive Cognitive Awareness*

Authors: *Fulong Fan, Peilin Liu, Liu FengZhe, Shuyan Yang, Gang Yan*

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?

This work is a fundamental study on LLM reasoning. It does not involve human subjects, sensitive personal data, or applications that could lead to foreseeable negative social impacts or safety risks.

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 publicly available benchmark datasets HotpotQA, StrategyQA, Big-Bench Hard. These datasets have been widely used in the NLP community and were pre-processed by the original creators to ensure anonymity and remove offensive content. We further manually inspected a random sample of the data and found no personally identifying information 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?

The dataset statistics are reported in Section 4.1

C. Did you run computational experiments?

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

See Section 4.1 for experimental setup

- 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?

As suggested by reviewers, we report multi-run statistics (mean and std) in Section 4.2

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.?

The full instructions and criteria provided to the human evaluators are detailed in Section 4.1

- 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)?

Volunteer graduate students were recruited as evaluators after providing informed consent. To safeguard their welfare and ensure ethical treatment, individual workload was capped at 2 hours.

- 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)?

All human evaluators were informed about the purpose of the study and how their ratings would be used, and they provided explicit consent before the evaluation.

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

This study was determined to be exempt from IRB approval by our institution as the evaluation involves assessing model outputs rather than collecting sensitive personal data from human subjects.

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?

AI assistants were used for grammar checking, language polishing, and code debugging during the research process.