

Responsible NLP Checklist

Paper title: *Modeling Multi-Dimensional Cognitive States in Large Language Models under Cognitive Crowding*

Authors: *Lin Zhong, Siyu Zhu, Zizhen Yuan, Jinhao Cui, Xinyang Zhao, Lingzhi Wang, Hao Chen, Qing Liao*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- N/A 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?

Discussed in Section 8. We discuss potential risks including over-interpretation of cognitive dimensions in LLMs and misuse of benchmark scores as proxies for human cognition.

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?

Discussed in Section 3 and Appendix A. All data were sourced from publicly available text and manually reviewed to remove personally identifying information and 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?

Reported in Section 3 and Appendix A, including dataset size, label distributions, and train/test splits.

C. Did you run computational experiments?

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

Discussed in Section 5 and Appendix D, including model architectures, training procedures, and hyperparameter settings.

- 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?

Reported in Section 5, Appendix D and Appendix E, including the average performance of multiple runs and the results on each subset.

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.?

Provided in Appendix F, including task instructions and examples.

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)?

Discussed in Section 3, including recruitment details and compensation scheme.

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)?

Annotators were informed about data usage and consented prior to participation.

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

The study involved minimal-risk annotation of non-sensitive text data and did not require institutional ethics approval.

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 did not use any AI assistants in our research, coding, or writing.