

Responsible NLP Checklist

Paper title: *Rhetorical Questions in LLM Representations: A Linear Probing Study*

Authors: *Louie Hong Yao, Vishesh Anand, Yuan Zhuang, Tianyu Jiang*

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?

No. This work is a descriptive representational analysis of existing models and datasets and does not introduce new models, deployments, or capabilities that would create additional 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?

No. This work uses existing, publicly released social media datasets with anonymized text and established usage practices, and does not involve new data collection or the handling of personally identifying information.

- 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?

Yes. Dataset sizes, class distributions, and train/validation/test splits are reported in Section 3.1 (Datasets).

C. Did you run computational experiments?

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

Yes. The experimental setup, representation choices, probe definitions, PCA dimensionality, and evaluation metrics are described in Sections 3.23.4, with additional implementation and stability details provided in Appendices A and B.

- 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?

Yes. Descriptive statistics including layer-wise AUROC trends, confidence intervals for Spearman correlations, and summary alignment metrics are reported in Sections 46, with additional statistics provided in Appendix G.

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.?

This work does not involve human subjects or annotators, and relies exclusively on previously published datasets with existing annotations.

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)?

This work does not involve recruitment or payment of human participants, as it relies solely on previously published datasets with existing annotations.

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)?

This work uses previously published, publicly available datasets, and does not involve new data collection or direct interaction with data subjects.

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

This work does not involve new data collection or research with human subjects, and therefore did not require ethics review board 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?

AI assistants were used for limited language polishing and minor code assistance, without contributing to experimental design, data processing, analysis decisions, or scientific claims.