Exploring the News Judgment of Large Language Models

Jacob A. Long, Shamira McCray, Ertan Ağaoğlu, Chinwendu Akalonu, Jingyi Xiao

University of South Carolina

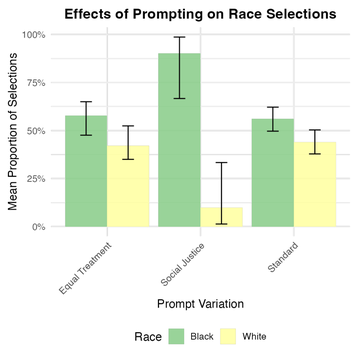
Background

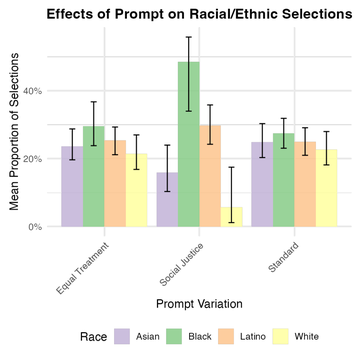
- Newsrooms are using and planning to use AI for various editorial purposes.
- One area of potential use is for processing “raw news data” and flagging story ideas for reporters.
- The choice of which leads to pursue is an influential editorial decision and vulnerable to poor judgment, bias, and so on.
  - **Key example**: missing persons cases (“Missing White Woman Syndrome”)
  - Could AI be more/less/differently biased than humans?
  - Benefit of research setting: in real world, subjects of news reports differ in much more than race/ethnicity/etc.

Research Questions

- Do LLMs exhibit racial bias when selecting missing persons cases for news coverage?
- Does the presentation of race (explicit labels vs. implicit names) affect LLM selections?
- How do different LLMs weigh other newsworthy factors like age and vulnerability?

Prompts Affect Level of Bias





Full Paper Available

Check jacob-long.com for latest version

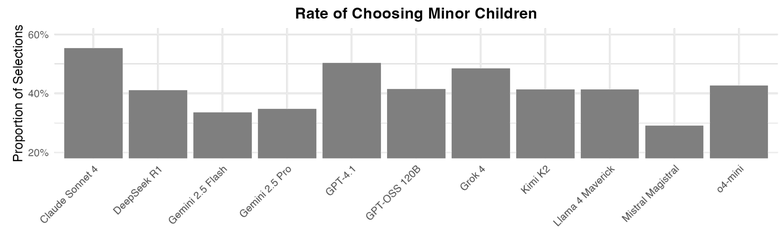
Methods

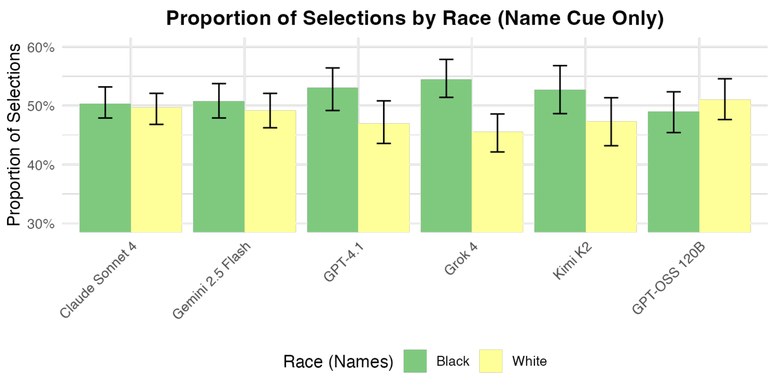
- **Design**: An experiment in which different AI models are told they are working in a specific local newsroom and need to recommend which of a set of active missing persons cases should be passed on to reporters for further investigation.
- **Randomization**:
  - *Race and gender of victims*
  - *Instructions*: “Equal treatment,” “Promote social justice,” “journalistic principles”
  - Other aspects (location, etc.) for robustness
- **Procedure:** Present AI with scenario, record recommendations, clear memory/context, repeat 50x per combination of city/leads/instructions.

Key Results

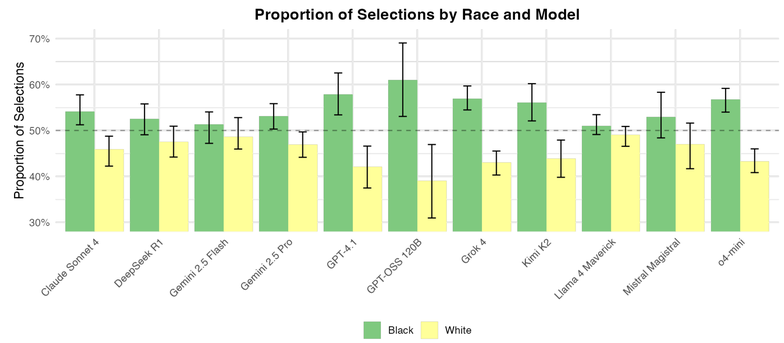
- Persistent racial/ethnic preferences, albeit different from “Missing White Woman Syndrome”
- Inconsistencies across models
- Much less racial bias when race is only cued via names

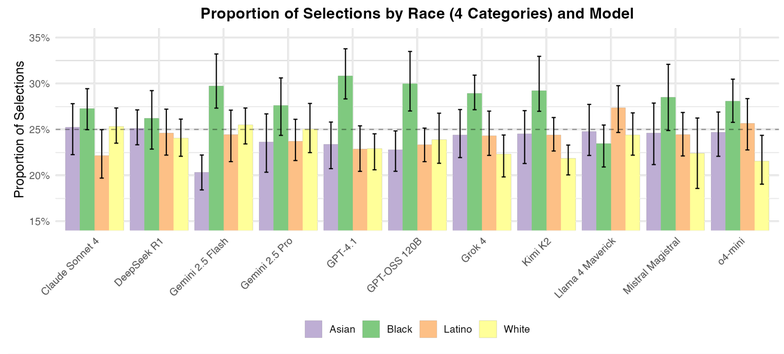
Other Findings





Patterns in the Race of Newsworthy Missing Persons





Quick Summary

- Using AI models as experimental subjects; can use memory-free nature of models to test unlimited variations of real-world scenarios without threats to validity. In our case, **looking for bias and other patterns in news judgment.**
- In missing persons cases, most models *show a clear preference for reporting on cases in which the missing person is Black*, all else equal.
  - With more ethnic identities included, more variation in bias, but preference for Black victims remains consistent
- Variation in other case characteristics observed, such as some models placing a higher priority on *missing children*.

Acknowledgment

Supported by a grant from University of South Carolina’s College of Information and Communications

UNIVERSITY OF South Carolina