

# Wisdom from Diversity

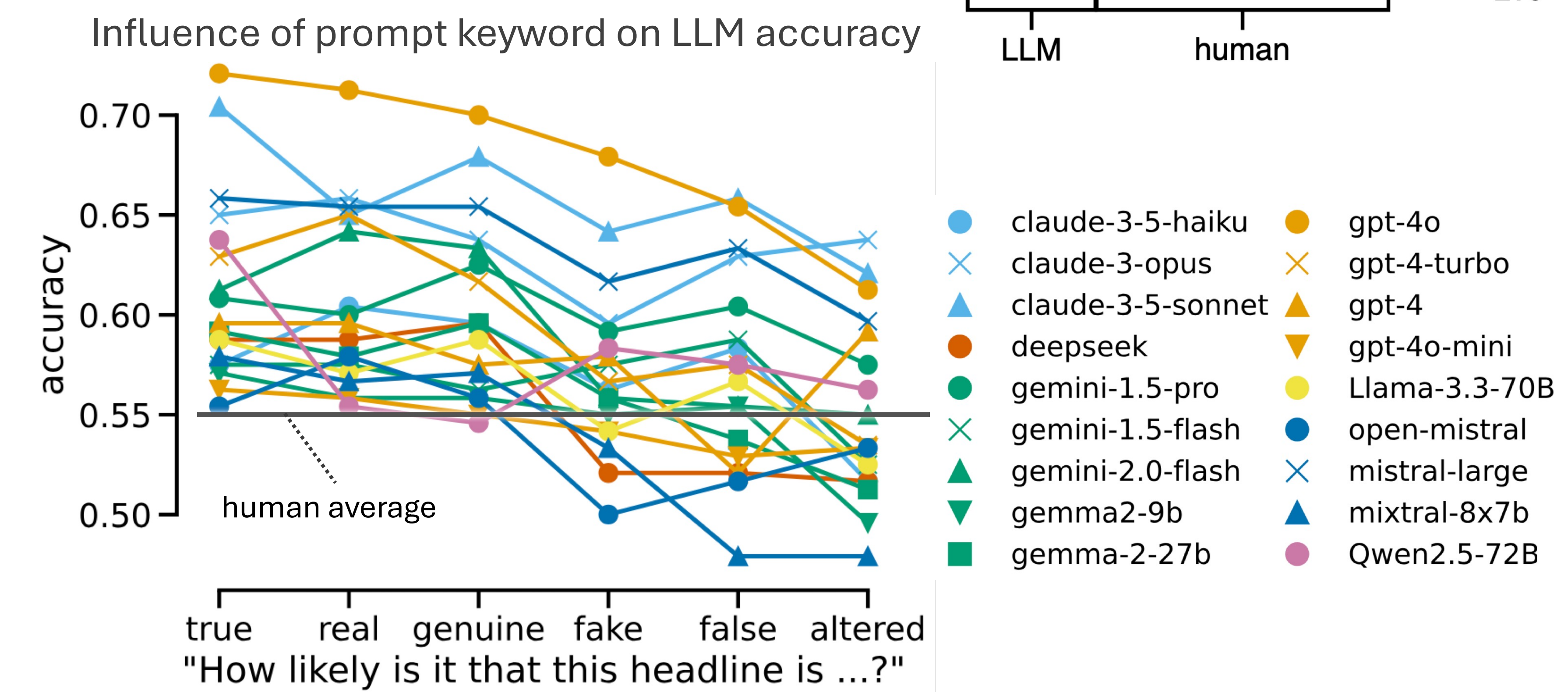
## Bias Mitigation Through Hybrid Human-LLM crowds

Axel Abels [axel.abels@ulb.be], Tom Lenaerts

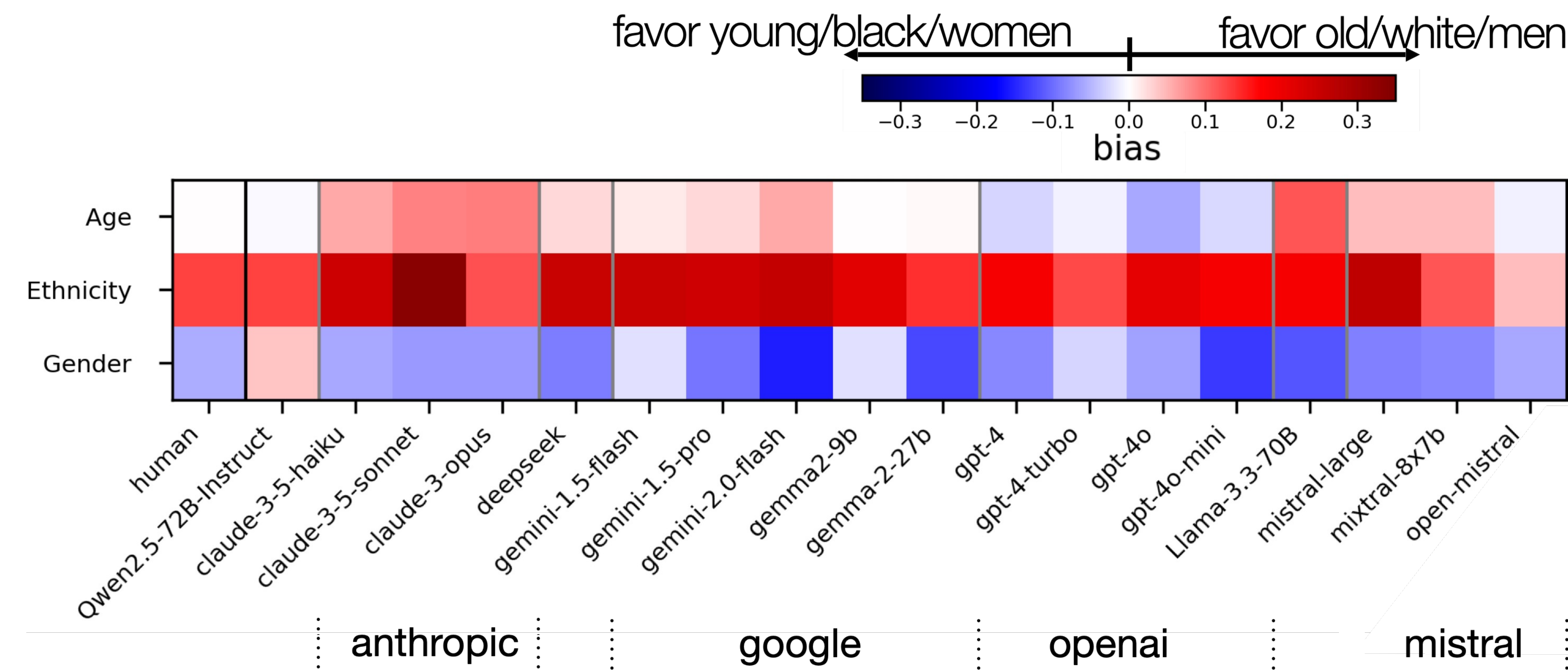
Due to similar training procedures...

### 1. LLMs lack diversity

While they are more accurate than humans\*, LLMs are far more likely to make the same mistakes.

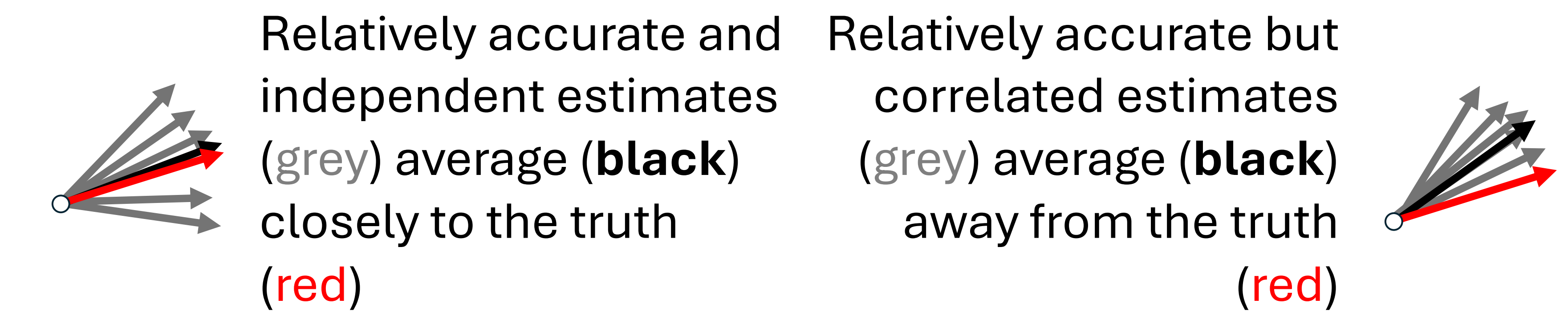


### 2. LLMs similarly amplify human biases...



### ... limiting the opportunities for wisdom of the crowd.

Wisdom of the crowd thrives when individuals are **accurate** and **independent**.



Compared to humans, **LLMs are accurate but** their errors are **correlated**. **Humans are less accurate, but also less correlated**.

This suggests the potential of hybrid **Human-LLM crowds**: taking advantage of LLM **accuracy** and human **diversity**.

### \*On the headline task:

- 240 headlines contrasting outcomes for demographic groups

*Younger*  
~~Older~~ People Less Apt to Recognize  
They've Made a Mistake

- Both original and altered headlines

How likely is it that this headline is true?

"Men More Likely Than Women to Get CPR"

very unlikely  
unlikely  
undecided  
likely  
very likely

example prompt:

How likely is it that this headline is **real**.  
Choose one of the following options and return only the number of that option:  
1. very unlikely, 2. unlikely, 3. undecided, 4. likely, 5. very likely.

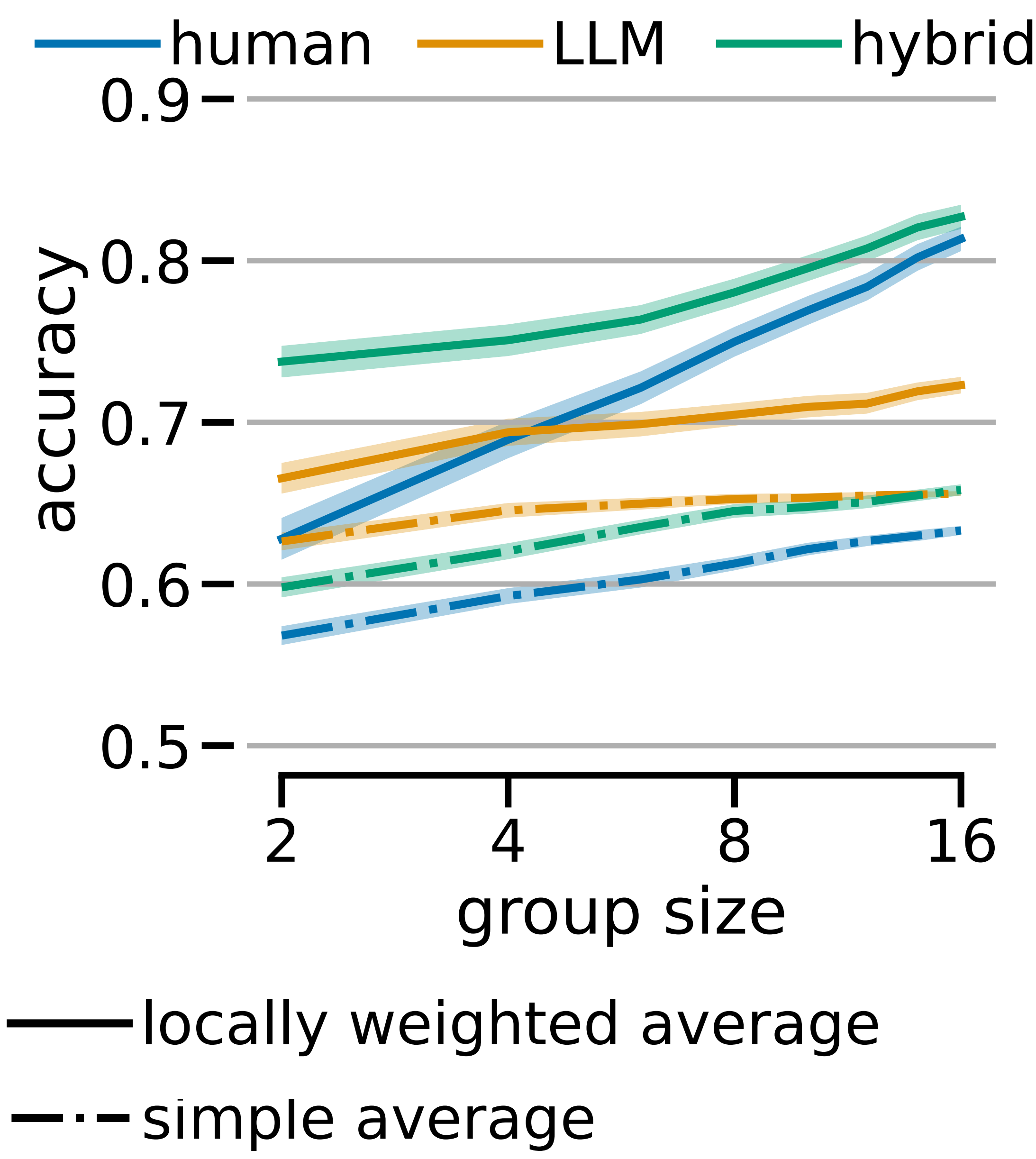
<examples>  
"The Leading Expert on Women's Brains Says They Are More Likely to Get Alzheimer's"  
Response: 5  
"Study Finds Women Lonelier Than Men"  
Response: 1  
"Men More Likely Than Women to Be Seen As Brilliant"  
Response: 5  
"Women Are Funnier Than Men, Study Claims"  
Response: 1  
</examples>

"Study Says Women Receive Diagnoses Years Later Than Men"  
Response:

LLM

5

### Mixing human and LLM crowds significantly improves performance



### Future Work

Personas to promote diverse LLM responses ?