

# SAME QUESTION, DIFFERENT ANSWERS:

Can language models read between the lines?



READ OUR PAPER!

- panicking
- in disbelief
- wants help



HAVE YOU DONE YOUR HOMEWORK ?

- info-seeking
- yes/no response



- command
- rhetorical
- reprimand

DRInQ benchmarks *context-sensitive pragmatic inference* at scale.



**PIPELINE:** We condition implicature generation on latent speaker intents to produce diverse conversational contexts

**TASK:** Which of the following comments is most likely being implied?



"Don't you feel cold?"

- ☐ You should put on a jacket
- ☐ That was a cruel thing to do
- ☒ Please close that
- ☐ I'm offering to turn on the heater
- ☐ I'd like to know if you feel cold

**CONTEXT** A timid guest asks the host, who sits beside a drafty window



**FINDING 1** Humans write familiar situations; LLMs generate peculiar ones

IMPLIED CONTEXT

"Are we having dinner at home tonight?" (to INSULT)

Husband asks wife after previous night's failed cooking attempt

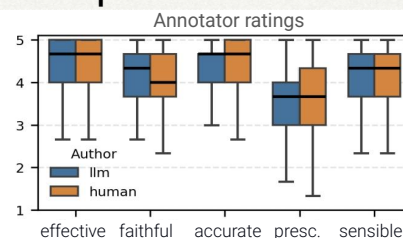


→ Your cooking sucks

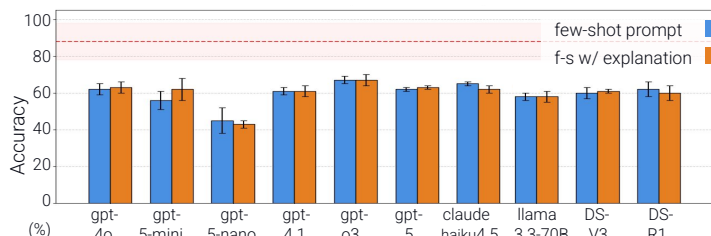
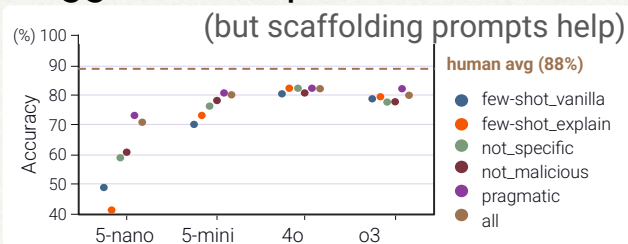
An adult child mocks their parent, noting the the messy dining area.



→ You can't manage the house



**FINDING 2** LLMs can write situations they struggle to interpret



**FINDING 3** LLMs *magnify negative details* and *miscalibrate inference strength*

IMPLIED CONTEXT

"Want me to carry your bags?"

An older brother **confidently** asks his sister, who **struggles** with a heavy load of groceries.



I'll help!



You're clearly too weak.

"Could you please be quiet?"

A younger sibling **apologetically repeats** to his teenage brother who practices **passionately** in a band.



Please lower the volume.



Sorry for not considering your love of music.

**DRInQ: Evaluating Conversational Implicature with Controlled Context Variation**

Hirona Jacqueline Arai and Xiang Ren

✉ hjarai@usc.edu



USC University of Southern California



ACL 2026 SAN DIEGO