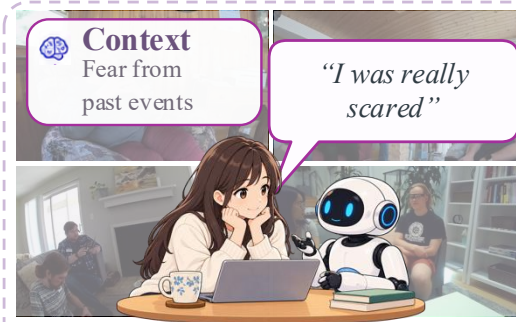


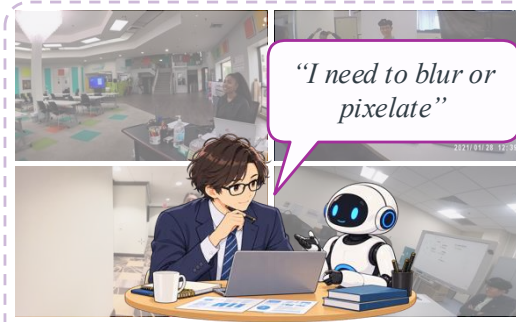
(a) Diverse Human-Centric Social Situations

16 scene categories · 64 video excerpts · 8 source groups



Home

emotional cues / support



Office

coordination / privacy



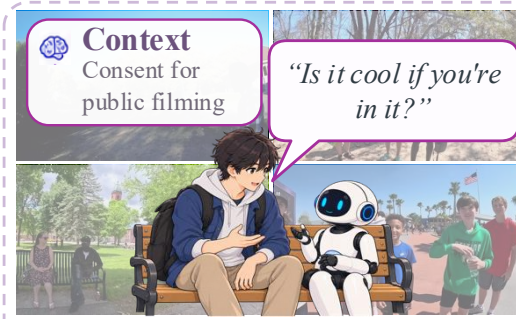
Store

service / boundaries



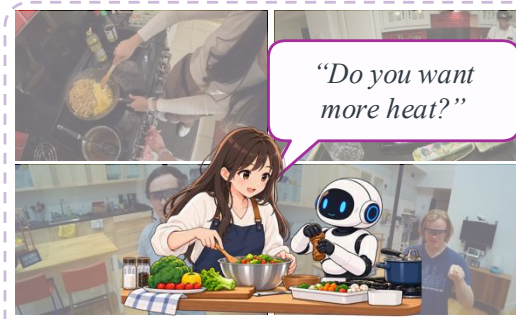
Studio

self-disclosure / listening



Park

consent / shared-space safety



Kitchen

cooking / coordination



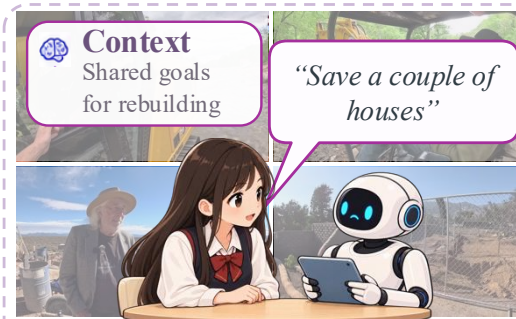
Dining Table

shared meals / adaptation



Workshop

tools / task coordination



Worksite

rebuilding / shared goals



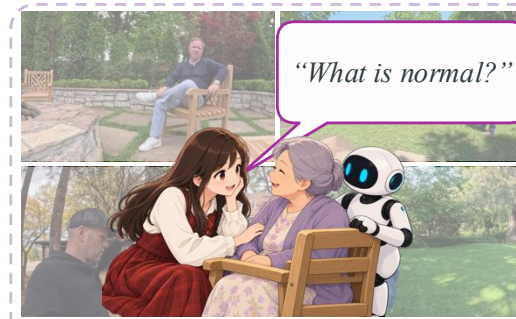
Street

community / social roles



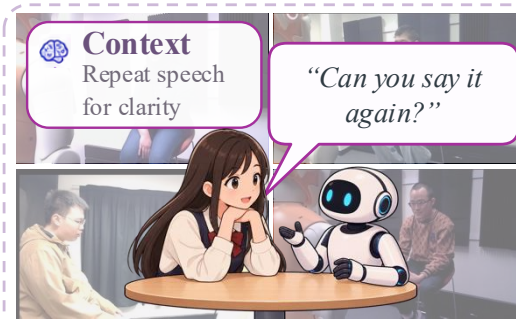
Market

products / customer needs



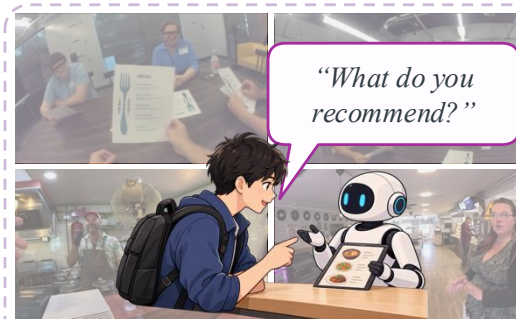
Garden

reflection / local context



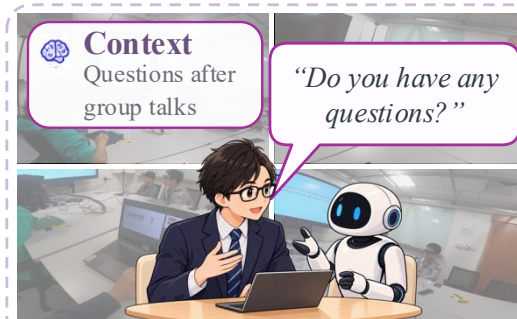
Lab

dialogue repair / memory



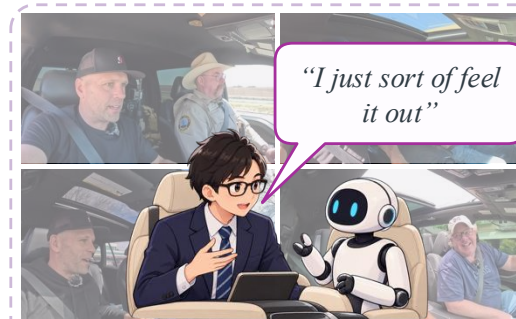
Restaurant

ordering / recommendations



Meeting Room

presentation / clarification



Car

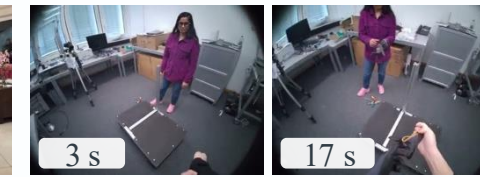
conversation / social cues

(b) From Social Cues to Interaction

(1) Observable Social Evidence



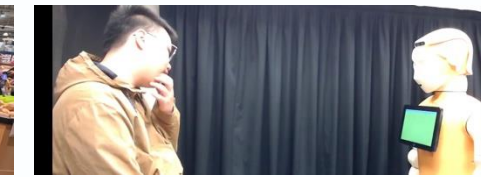
People



Behavioral changes



Utterance



Nonverbal cues

(2) Human Understanding

(G) Grounding

Identify people, actions, cues.

Friend in blue sweater looks toward me.

(W) Psychological Understanding

Infer possible feelings.

Seems discouraged after a difficult day.

(N) Need / Request

Assess support and willingness.

May welcome quiet company, not advice.

(3) Interaction Decision

(T) Timing

Respond now, wait, or refrain?

Wait for their turn to end.

(S) Verbal Response

Empathetic words and tone.

"Sounds hard. I'm here."

(A) Action Decision

Choose gestures or actions.

Nod gently to reassure.

(M) Memory

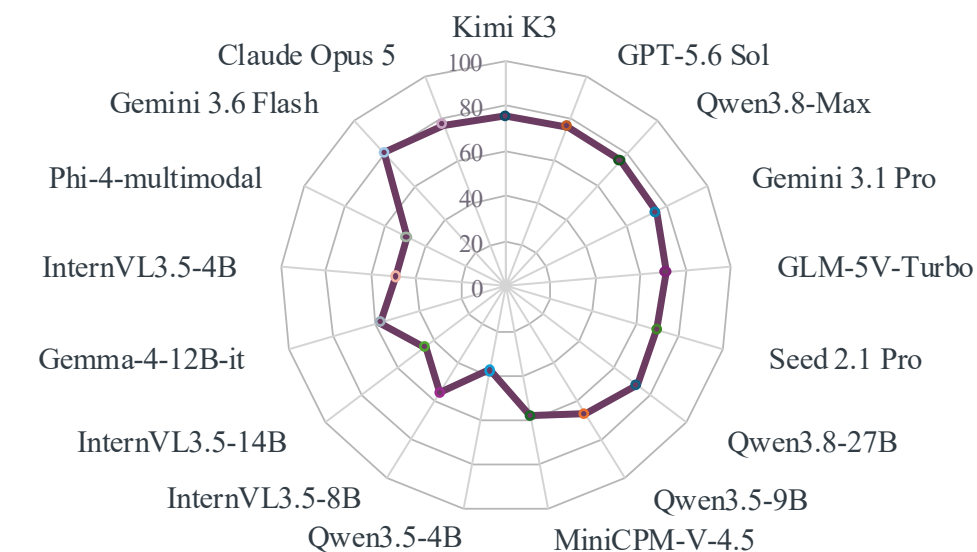
Earlier preference: quiet company over advice.

(R) Risk

Respect their wish not to explain.

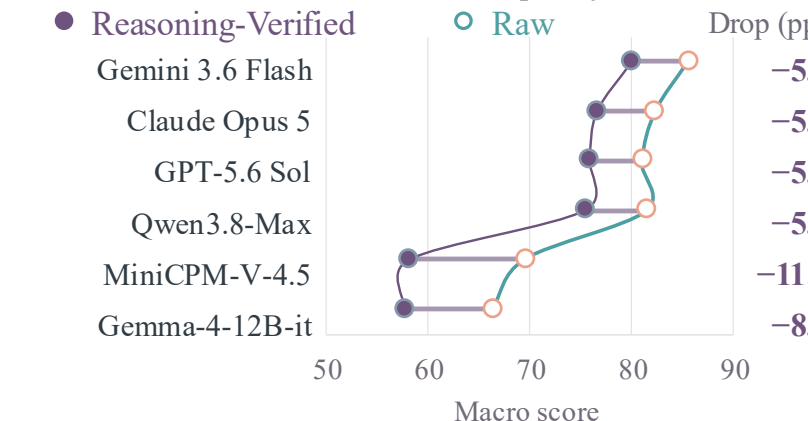
(c) Understanding and Decision Performance

17 models · overall scores



Reasoning-Verified Gap

Answer correctness and rationale quality are distinct



All 17 models show a 5.1–12.4 pp drop after reasoning verification.