

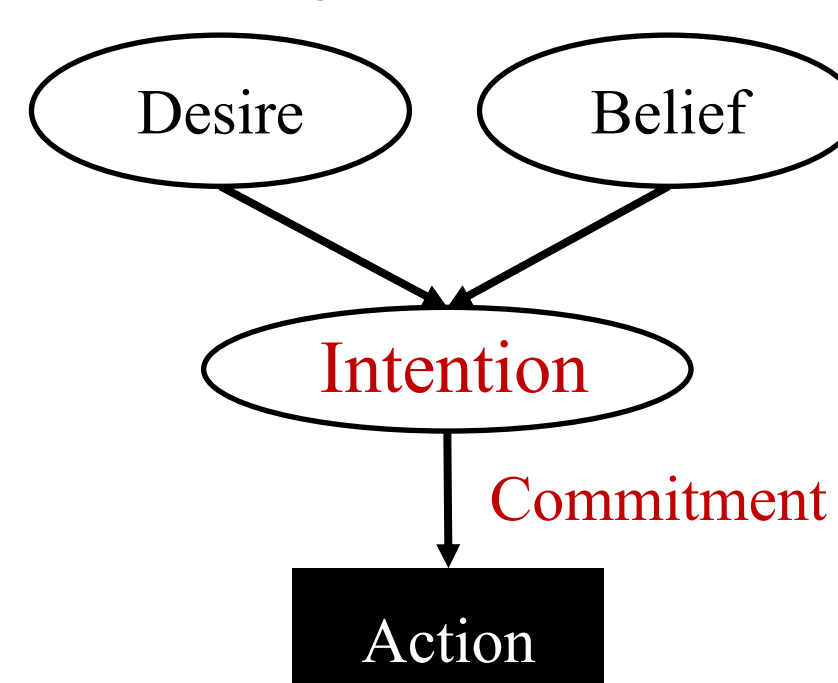
Motivation

How to act with conflicting desires



The problem of conflicting desires.

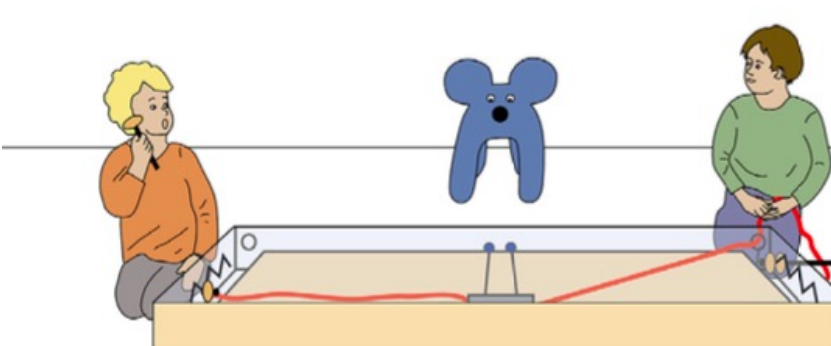
Intention theory (e.g., Bratman, 1987)



Commitment to intentions for coherent actions over time.

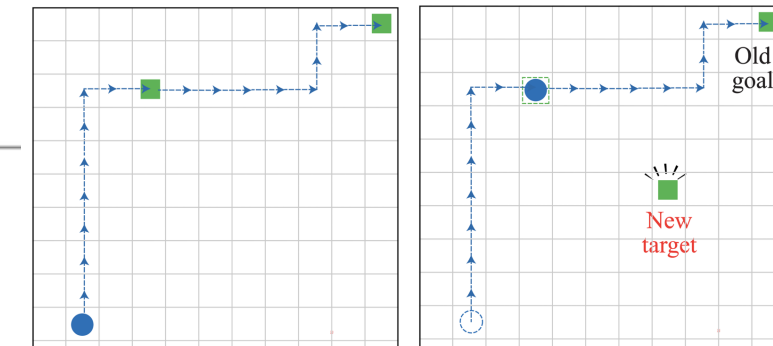
Visual-grounded commitment

Joint-commitment



- Children commit to joint goals for cooperation
- Earliest sign of uniquely human (Tomasello, 2019)

Self-commitment



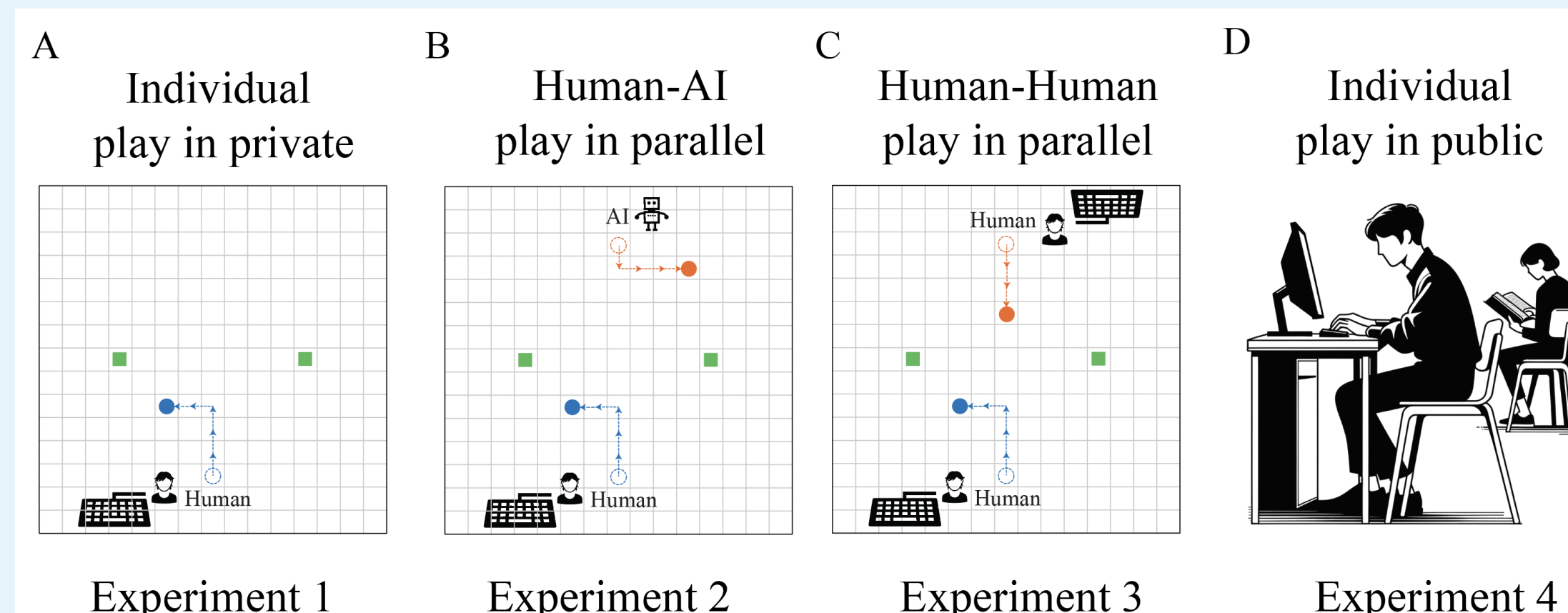
- Adults commit to intentions in individual decision-making.
- Qualitatively different from AI (e.g., RL agent)

Questions:

Self-commitment has a social origin?
Can social contexts impact the strength, content, and timing of self-commitment in individual tasks?

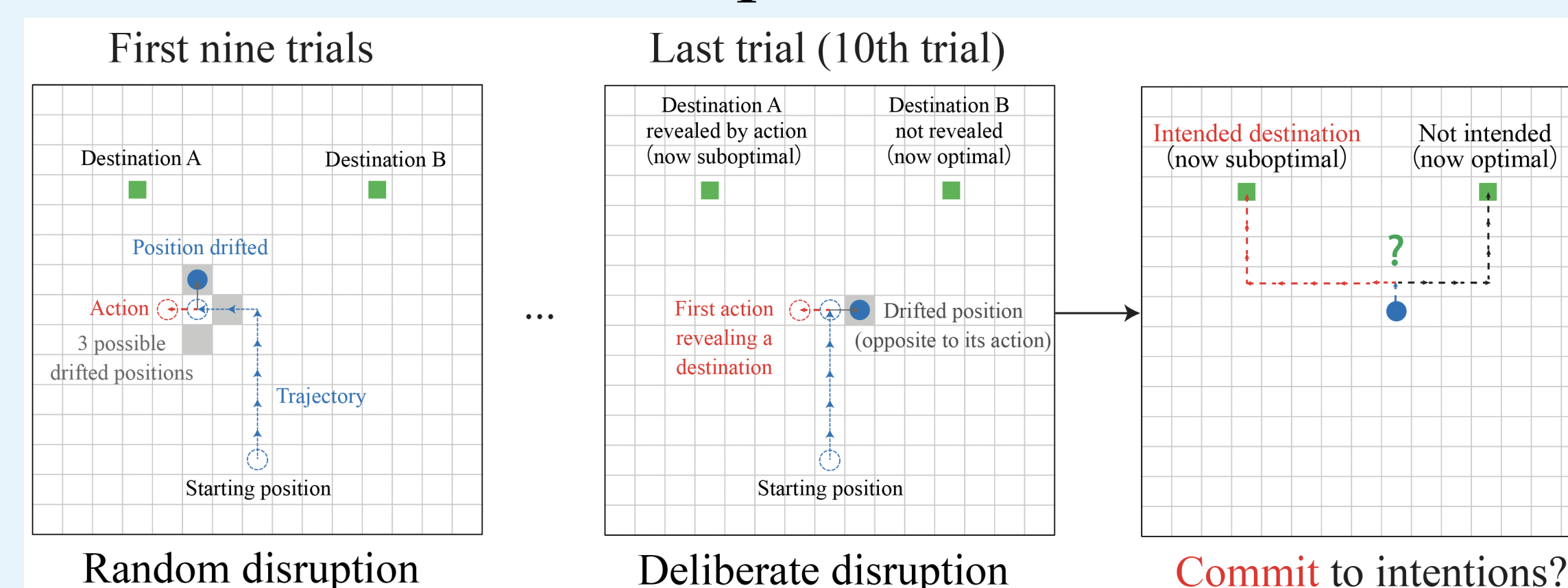
Design

Manipulations of social contexts

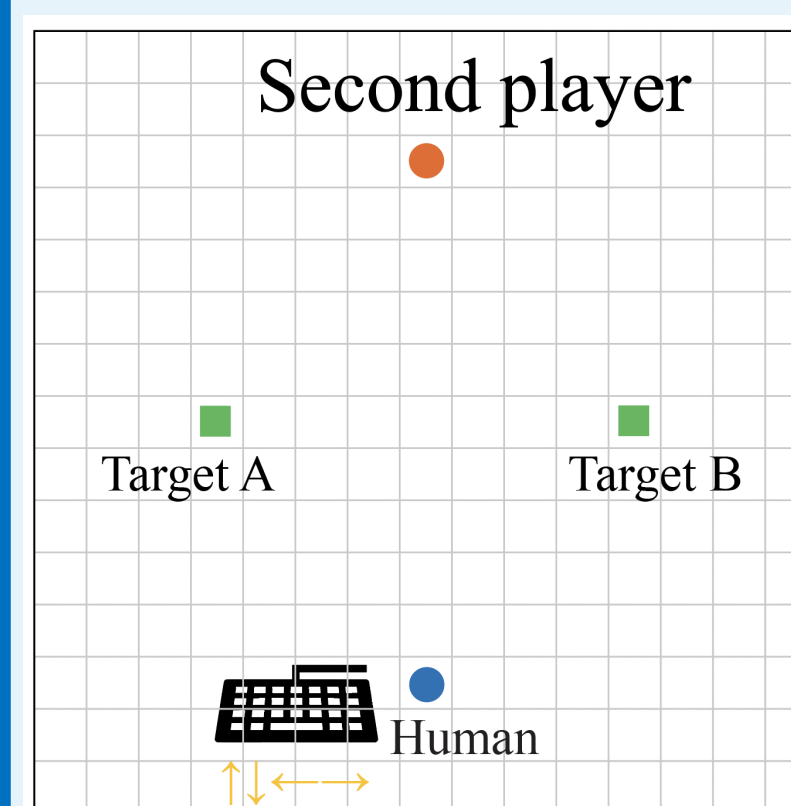


Social contexts do not change the individual nature of the task.

Task in Experiment 1 & 4

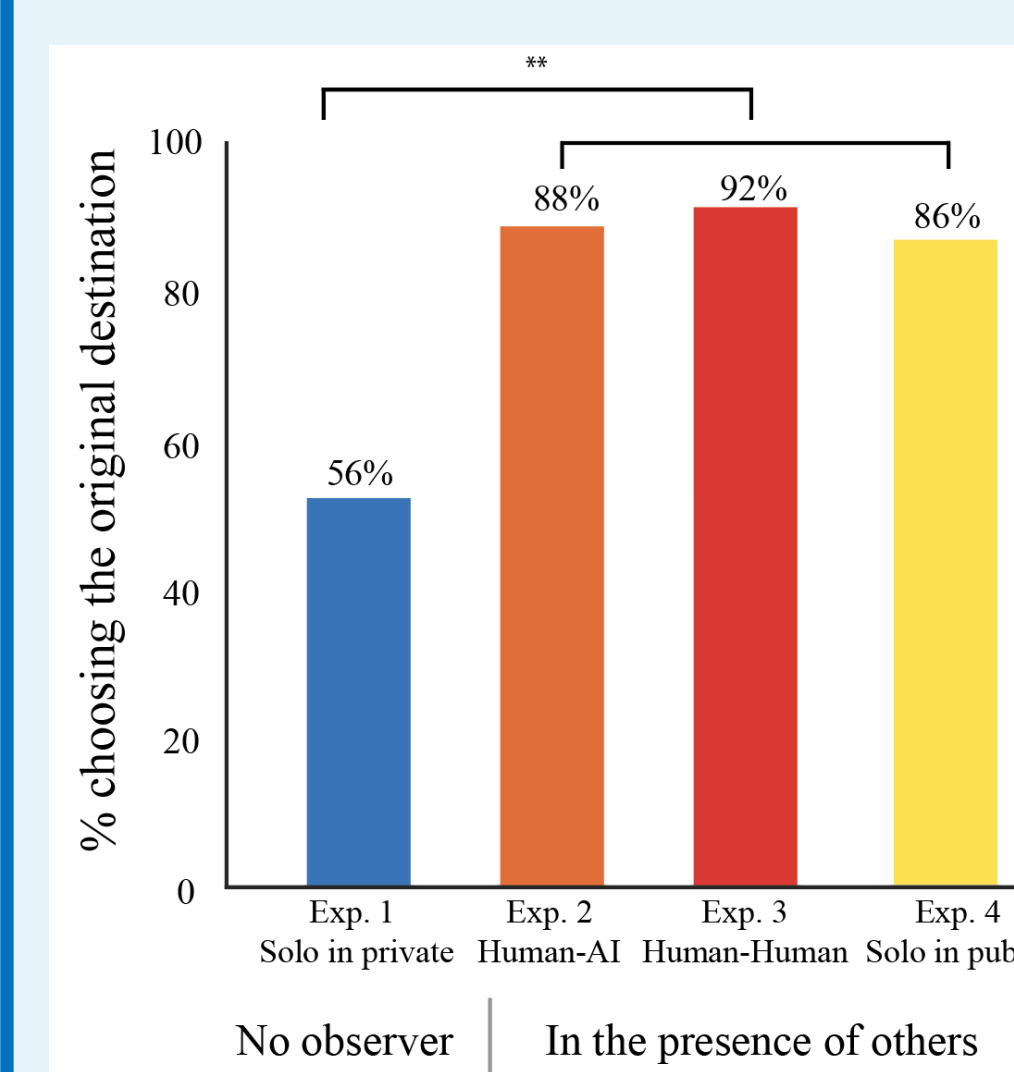


Task in Experiment 2 & 3

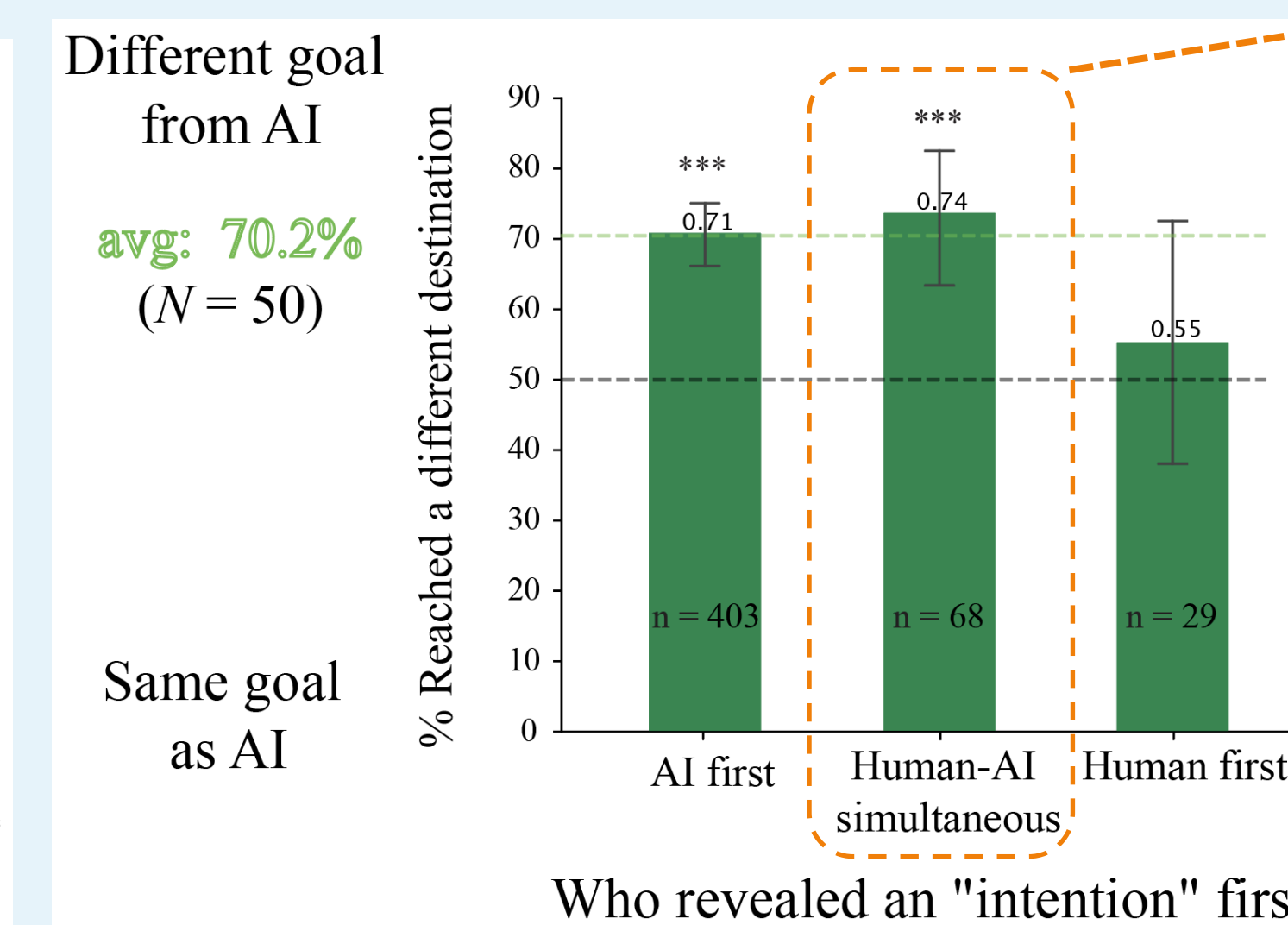


- Same as Exp. 1, except another player will present in the game.
- Participants were explicitly informed **no** need to consider other's actions.
- Second player was a reinforcement learning agent, who optimize it's own utilities (Exp. 2) or a human player (Exp. 3)
- Both introduced as a human player.

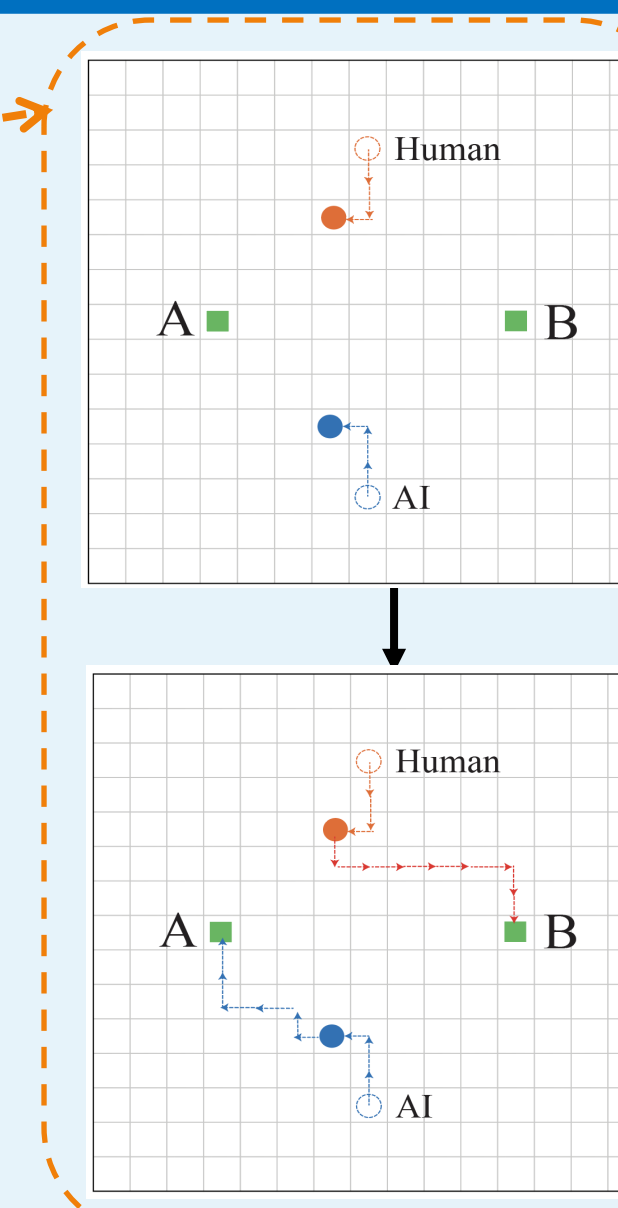
Results: Adults across different social contexts



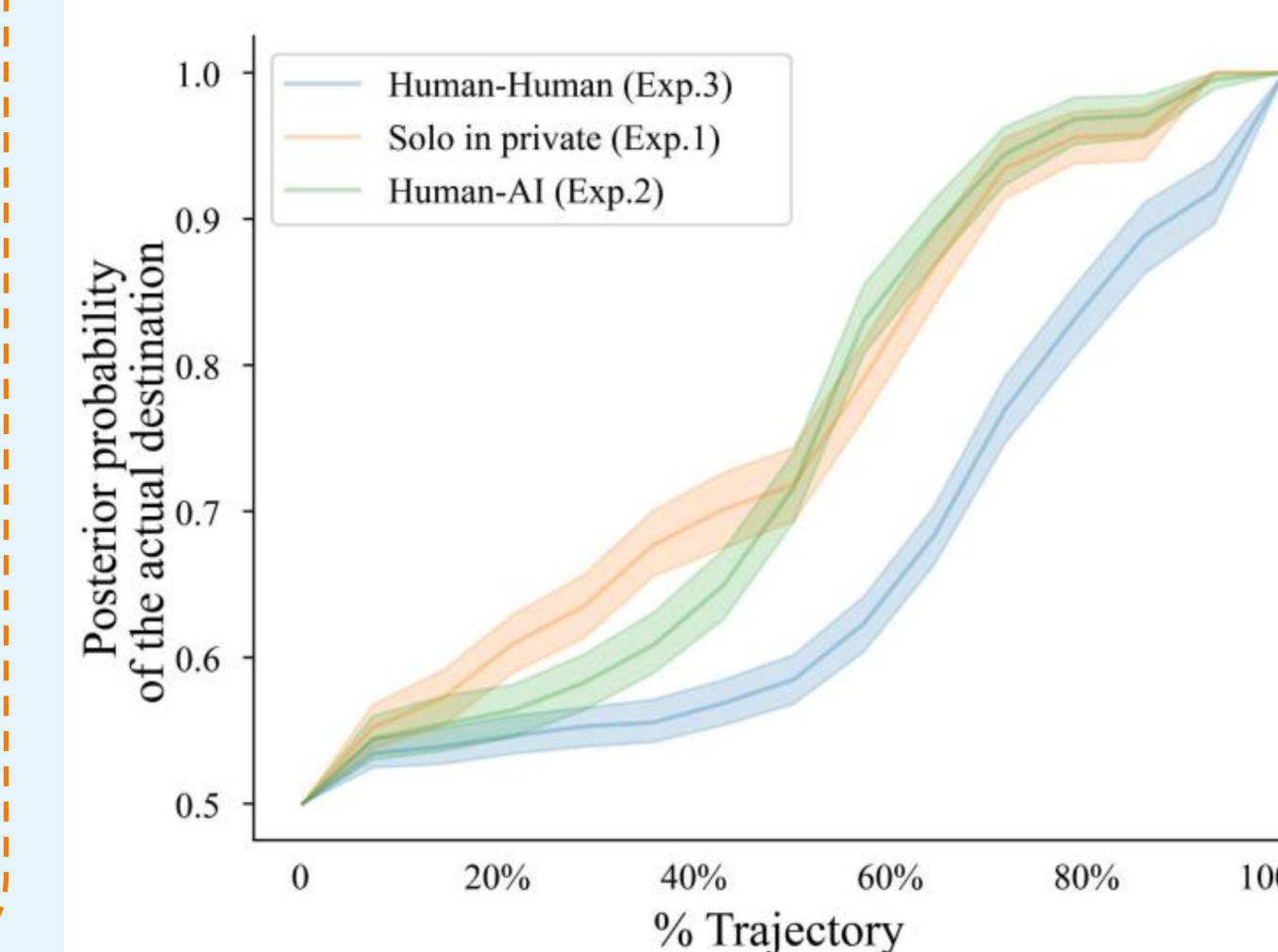
Strength of commitment increased



Preferred a different goal, as long as they were not the first to reveal an intention.



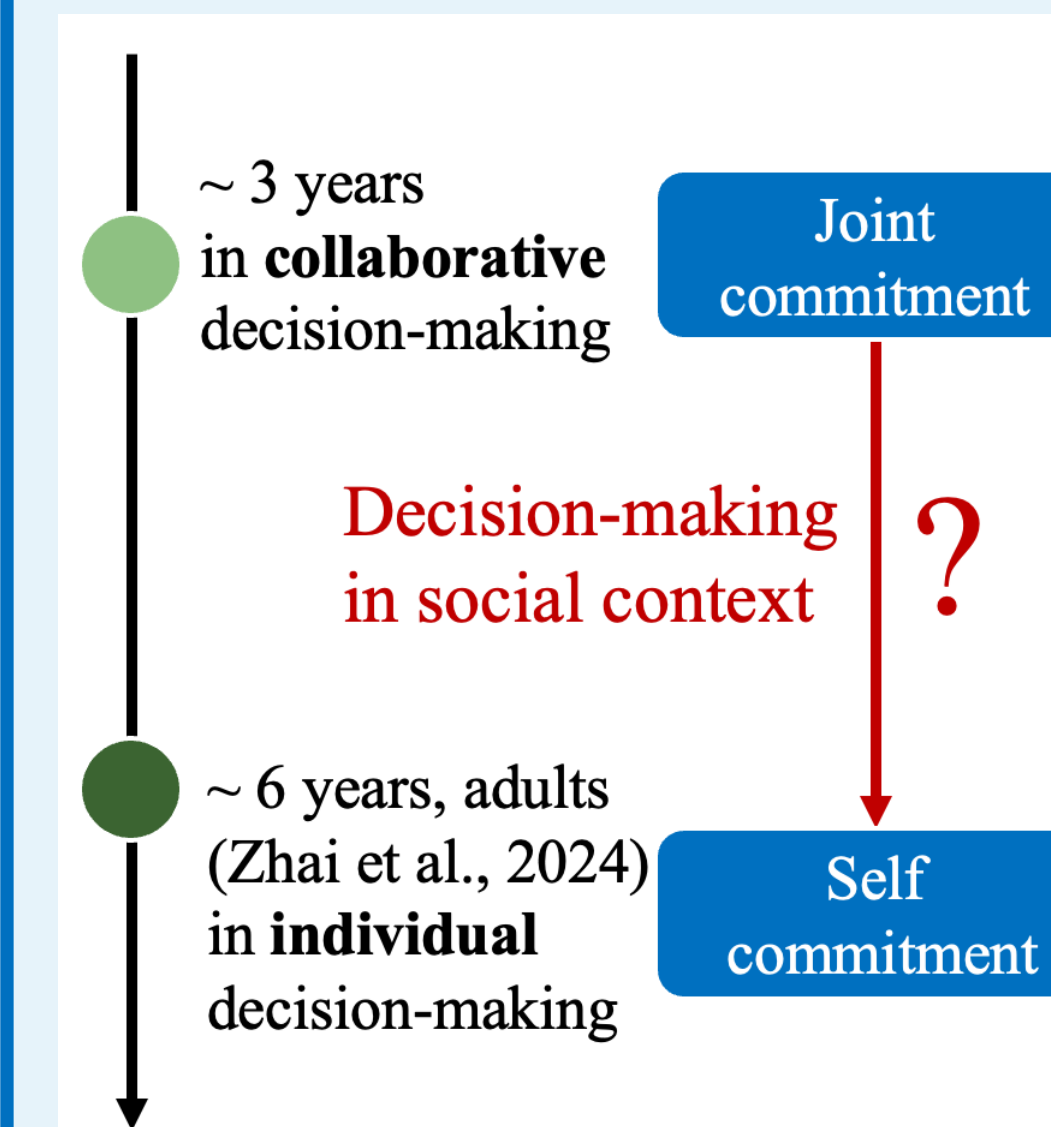
Bayesian goal inference



Delayed commitment revelation in dual play.

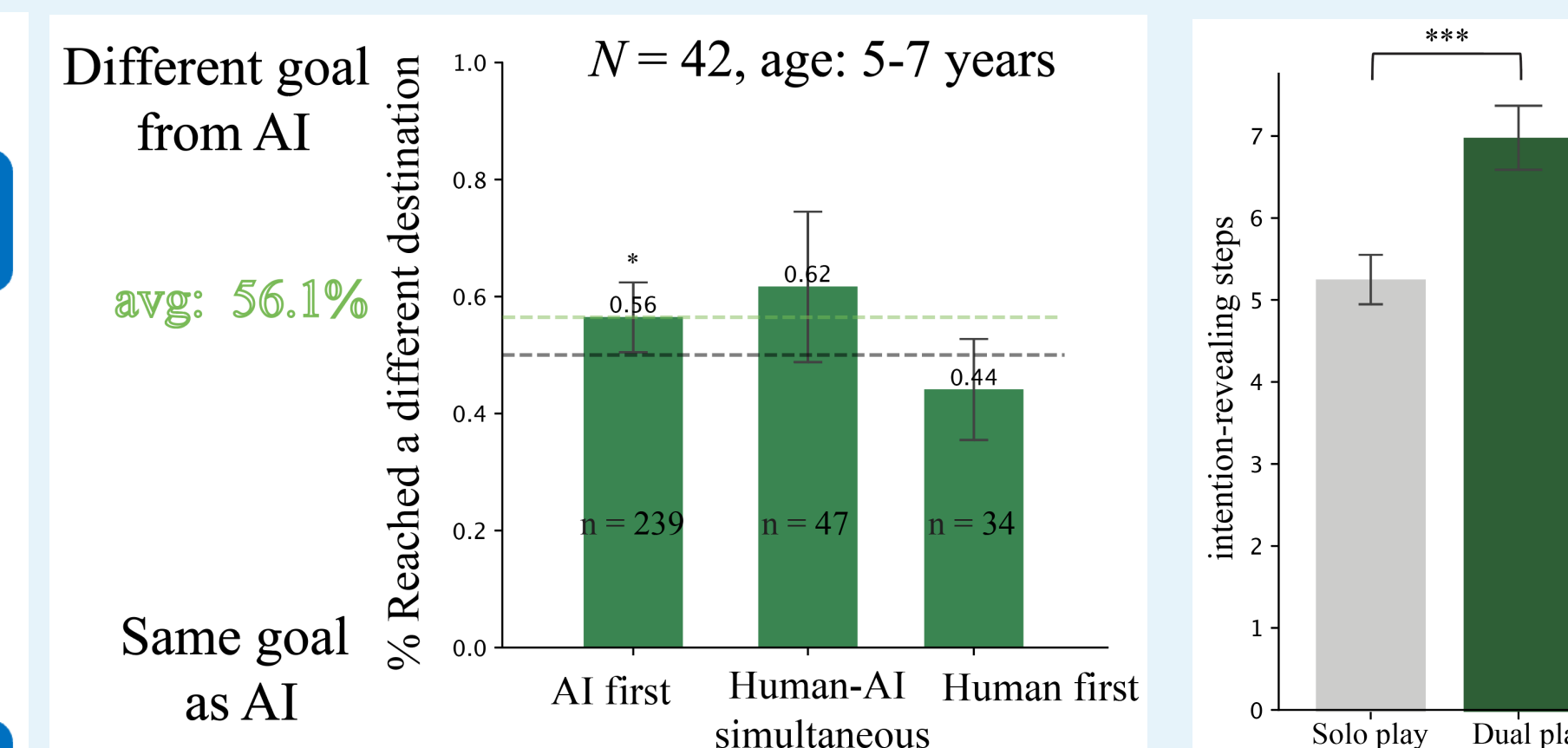
Ongoing studies: Children-AI

Developmental Hypothesis



Emerge before 6 years?

Within-subj design: solo + dual



Children aged 5-7 years showed spontaneous intention coordination; they took longer to decide on a goal in social contexts and preferred a different goal when others signaled commitment first.

Conclusions & future work

- Even in individual tasks, self-commitment is deeply intertwined with social context, influencing how people manage their goals and interactions with others.
- What if humans were told they are playing with AI? What if there exist both collaborative and individual goals?
- What are the critical cognitive processes being socialized/internalized that underlie the emergence of self-commitment?



Email: chengshaozhe@ucla.edu
 Web: chengshaozhe.github.io