Spontaneous coordination with self-commitment: How the presence of others alters the strength, goal and timing of commitment

Shaozhe Cheng*[1,2,5], Jingyin Zhu[1], Jifan Zhou*[1], Mowei Shen*[1], Tao Gao*[2,3,4]

[1]Department of Psychology and Behavioral Sciences, Zhejiang University

[2]Department of Communication, UCLA

[3]Department of Statistics, UCLA

[4]Department of Psychology, UCLA

[5]Department of Psychology and Neuroscience, Duke University

Author Note

The authors have no conflicts of interest to disclose.

This work was supported by UCLA startup funding to Tao Gao, and grants from National Natural Science Foundation of China (32071044 to Mowei Shen, 32371088 and 6233700 to Jifan Zhou). We thank Tamar Kushnir and Michael Tomasello for their helpful feedback. We also wish to acknowledge anonymous reviewers for helpful comments on previous drafts.

*Correspondence concerning this article should be addressed to

Shaozhe Cheng, Department of Psychology and Neuroscience, Duke University, 417 Chapel Dr, Durham, NC 27708 (present address). Email: chengshaozhe@ucla.edu or shaozhe.cheng@duke.edu

Jifan Zhou or Mowei Shen, Department of Psychology and Behavioral Sciences, Zhejiang University, Zijingang Campus, 866 Yuhangtang Road, Hangzhou, ZJ 310058, P.R. China. Email: jifanzhou@zju.edu.cn or mwshen@zju.edu.cn

Tao Gao, Department of Communication, UCLA, Los Angeles, CA 90095, USA. Email: taogao@ucla.edu

Abstract

Commitment is a paradoxical feature of human behavior, often seen as both an irrational bias and a virtue for achieving goals. This study investigates its social roots, revealing how social contexts shape the strength, content, and timing of self-commitment, even in individual tasks. Through a series of game-like experiments, participants pursued one of two equally desirable goals via sequential actions under varied social conditions: alone in a private room (Experiment 1), alongside an optimal reinforcement learning (RL) agent (Experiment 2) or another human (Experiment 3) on a shared display, or alone with a passive observer present (Experiment 4). Our results demonstrate that (1) all social contexts consistently heightened self-commitment, underscoring its sensitivity to the public nature of tasks; (2) in parallel-play settings (Experiments 2 and 3), participants spontaneously inferred others' intentions and avoided selecting the same goal, despite instructions that such avoidance was unnecessary, suggesting that theory-of-mind (ToM) inference of another agent is spontaneously evoked to bias goal selection; and (3) Bayesian ToM modeling indicated that participants delayed revealing their intentions in parallel-play settings but not in the observer condition, implying that spontaneous bargaining with a potential partner—rather than mere observation—prompts more cautious commitment formation. These findings illuminate that, even in individual tasks, self-commitment is deeply intertwined with social context, influencing how people manage their goals and interactions with others.

Keywords: commitment, intention, social context, self-presentation, decision-making, theory of mind, reputation

Introduction

Commitment presents a paradox in human life. Within the traditional contexts of behavioral economics and organizational psychology, commitment is often studied as an irrational bias, termed the escalation of commitment or sunk cost fallacy, which causes individuals to persist with decisions despite suboptimal outcomes (Staw, 1976; Arkes & Ayton, 1999). On the other hand, in everyday life, commitment is widely regarded as a virtue essential for achieving challenging goals, at both the individual and collective levels. The appreciation of commitment is well demonstrated by Winston Churchill's inspiring speech, in which he urged perseverance in the face of difficulty: "...never give in. Never, never, never, never—in nothing, great or small, large or petty…" Echoing this appreciation for commitment in real life, philosophers have long emphasized it as a defining characteristic of human rationality, rather than viewing it as a weakness of human nature. It has been argued that while animal rationality predominantly focuses on finding the optimal *means* to satisfy certain desires, the unique aspect of human rationality lies in their ability to form intentions and commit to them in the face of complex, even conflicting desires (Bratman, 1987; Searle, 2003). On this view, commitment is an essential characteristic of human intention, where forming an intention already entails a commitment to that intention (Bratman, 1987). Commitment, therefore, is not merely an assertion of intent but involves a plan for executing a sequence of actions aimed at fulfilling that intention. As a result, commitment to an intention entails “inertia” and “resistance to reconsideration” (Bratman, 1987)—even when environmental changes might suggest a better alternative. The ability to commit is central to human planning and decision-making. For example, when planning a summer vacation and considering several appealing destinations, one must ultimately commit to a single option in order to form a coherent plan and carry it out. Without commitment, even simple decisions can become paralyzing or unstable.

Previous studies, using introspection and subjective self-reports, have found that leveraging commitment as a high-level strategy can facilitate goal attainment (Gollwitzer & Brandstätter, 1997; Gollwitzer, 1999; Gollwitzer & Sheeran, 2006; Ajzen et al., 2009; Nenkov & Gollwitzer, 2012; Ajzen & Kruglanski, 2019). For example, when individuals are encouraged or explicitly instructed to make a commitment—such as signing a commitment statement or setting a New Year's resolution—they are more likely to achieve their intended future goals, like quitting smoking or losing weight (Gollwitzer, 1999).

Recent studies have begun to explore the phenomenon of spontaneous commitment in the context of rational decision-making, where both the goal value and the action cost are clearly defined, and human actions can be evaluated from a utility calculus perspective. One study found that when faced with two valuable moral goals, both adults and children (aged 4-6 years) tend to stick with their chosen goal even if it becomes costly later on (Chu & Schulz, 2022). Another study examines how commitment regulates conflicting desires in sequential decision-making tasks (Cheng et al., 2023). This study draws inspiration from the Buridan's ass paradox, a thought experiment hypothesis that an ass, lacking intention, might starve when faced with two equally desirable piles of hay due to indecision. The findings indicate that humans resolve conflicting desires by committing to an intention, making their actions qualitatively different from those of a purely desire-driven agent, which acts solely to maximize expected utilities. A subsequent study further demonstrated that 6-year-olds, but not 5-year-olds, spontaneously commit to a future goal, even if new opportunities emerge that make their initial choice suboptimal. Moreover, they found a positive correlation between children's self-commitment and their proactive attentional control (Zhai et al., 2024), suggesting that commitment may be a spontaneous process but simultaneously requires effortful self-regulation. Collectively, these studies support the view that human intentional actions transcend pure reward maximization models, emphasizing intention as a crucial intermediate mental state in understanding human decision-making across various domains (Molinaro & Collins, 2023).

The commitment nature of human intention has also sparked studies in the context of social decision-making. It has been suggested that one important function of commitment is to make one's future actions more predictable, thus facilitating social coordination (Bratman, 1987; Michael & Pacherie, 2015; Michael, 2022). Developmental and comparative findings show that, unlike chimpanzees, children as young as three can form joint commitments to regulate both others and their own behavior (Warneken et al., 2006; Graefenhain et al., 2009; Hamann et al., 2012; Duguid et al., 2014). In cooperative tasks where children and chimpanzees are motivated to play with a partner together, only humans attempt to re-engage their partners when collaborative activity is interrupted, thus being more successful in cooperation (Warneken et al., 2006). Moreover, human toddlers who engage in a prior collaborative activity are significantly more likely to continue assisting their partner even after receiving their share of rewards prematurely, compared to a baseline condition where no prior collaboration occurred (Hamann et al., 2012). These findings have led to the argument that the commitment to shared intention is critical for understanding what makes humans unique (Tomasello, 2014).

Research on the social implications of commitment has primarily focused on cooperative scenarios, where joint commitment to a shared intention is essential for coordination. Yet, in everyday life, individuals pursuing their own objectives still influence—and are influenced by—their social surroundings. Consider, for instance, the difference in efforts when completing a challenging lifting set at the gym alone versus in front of others, or the awkwardness when actresses on the red carpet discover they have chosen the same dress. This raises a critical question that bridges the largely separate lines of research on individual and joint commitment: how is commitment to an individual task influenced by the social context when non-cooperating others are present? Exploring this question will illuminate commitment—whether individual or joint—as a unique feature of human rationality, potentially shaped by the distinct human challenge of navigating complex social environments.

Regulating self-presentation in social contexts

Exploring the social nature of decision-making by examining how it changes under different social contexts has a long tradition. Early work on social facilitation demonstrated that the mere presence of an observer can enhance an individual's performance, often through increased arousal (Zajonc, 1965; Rajecki et al., 1977; Aiello & Douthitt, 2001). More recent studies, however, have explained boost performance with a more nuanced view of shared reality and shared attention (Higgins et al., 2021; Shteynberg, 2015) — when people believe they are collectively attending to the same task ("we are attending to task X"), they tend to devote more cognitive resources to the task and, as a result, improve their goal pursuit (Shteynberg & Galinsky, 2011).

In addition to enhancing arousal and cognitive resources, the presence of others can trigger a cascade of cognitive and social processes, primarily centered on the need to "perform" as an actor before an "audience." Goffman (1959) theorized that individuals engage in self-presentation during face-to-face interactions to convey an impression that aligns with expected social roles. Empirical research has since provided robust support for this theory. First, individuals strategically regulate their self-presentation to maintain or enhance their reputations, adjusting their behavior based on how they expect to be judged by others (Baumeister, 1982; Leary & Kowalski, 1990). For example, when individuals anticipated discussing their personal achievements in front of others, they presented themselves more favorably—emphasizing positive qualities and downplaying weaknesses. In contrast, participants who did not expect any social evaluation were less inclined to engage in such self-enhancing behaviors (Baumeister, 1982). Second, people automatically adopt the perspective of their collaborators, even when doing so is irrelevant to the task. In the classic joint Simon effect, two co-actors each respond to different stimuli, yet their reaction times are significantly influenced by their partner's task. Responses are faster when partners' stimulus-response mappings are congruent and slower when incongruent, demonstrating spontaneous, task-irrelevant perspective-taking (Sebanz & Knoblich, 2009). Third, even the mere presence of others, without any direct interaction, can increase prosocial behaviors, leading to greater generosity and reduced dishonesty (Engelmann, Herrmann, & Tomasello, 2012; Engelmann & Rapp, 2018). Fourth, self-presentation likely relies on theory of mind (ToM)—the ability to infer others' mental states, including beliefs, desires, and intentions (Gopnik & Wellman, 1992; Wellman et al., 2014). Recently, ToM has been formalized as a Bayesian inference process, in which an observer infers another agent's latent mental states—such as beliefs and desires—using Bayes' rule. Under this framework, the most likely combination of mental states is inferred as the one that best explains the agent's observed actions, assuming the agent acts rationally to maximize utility given those mental states (Baker et al., 2009; Jara-Ettinger et al., 2016). More recently, ToM has been extended beyond interpreting others' actions to explain a range of social behaviors, including cooperation (Wu et al., 2021; Tang et al., 2022) and communication (Ho et al., 2021; Stacy et al., 2024). It has also been applied to inward-facing processes, such as how individuals use ToM to regulate their own actions to make their mental states more easily readable to others (Ho et al., 2021; Jiang et al., 2022; Royka et al., 2023). This inward ToM enables proactive self-regulation to create a favorable impression. Supporting this, research shows that even young children can infer how others evaluate their abilities and strategically adjust their behavior to shape these perceptions (Asaba & Gweon, 2022).

Placing individual commitment in a social context

Given that social contexts effectively trigger a cascade of social-cognitive processes (e.g., Engelmann & Rapp, 2018; Shteynberg, 2018), they can serve as a valuable tool for examining whether individual commitment, like joint commitment, can adapt to different social contexts.

One hypothesis suggests that self-commitment, unlike joint commitment, functions as a purely individual cognitive mechanism and is not further strengthened by social context. If anything, social context may actually decrease self-commitment, as it promotes adaptability and flexibility through enhanced arousal and increased cognitive resources—a general effect of social context on all tasks. This is because, as demonstrated by the sunk cost fallacy (Staw, 1976; Arkes & Ayton, 1999; Tversky & Kahneman, 1973), self-commitment can manifest as a rigid and persistent pursuit of a suboptimal goal, which negatively impacts task performance. Therefore, improved task performance in a social context may indicate a reduction in such rigid commitment.

Alternatively, like joint commitment, self-commitment may be socially relevant, causing it to function differently across various social contexts. Given the importance of commitment in early social interactions (Tomasello, 2024), it is possible that this distinctly human capacity first emerged in collaborative contexts, where individuals had to interpret one another's intentions and mutually rely on each other to remain committed to those intentions. Over time, this commitment for social coordination could have been internalized and adapted for individual planning (Cheng et al., 2022). One major advantage of such commitment is the predictability and stability it confers, allowing individuals to coordinate with their "future selves" and plan actions far into the future (Bratman, 1987, 2018). For instance, it is possible to commit to traveling abroad next year—outlining specific cities to visit—yet have no firm plans for the upcoming month. This hypothesis aligns with recent findings indicating that the capacity for sequential, individual commitment develops around age six, which is notably later than the earliest observations of joint commitment (Zhai et al., 2024).

If self-commitment arises from social coordination, then even seemingly “individual” commitments possess an inherently social dimension: they render future actions more predictable, a quality highly valued in social settings. Accordingly, committing to one’s own goal can be a virtue worth signaling to others. In this sense, the function of displaying one’s intentional commitment through actions mirrors the function of visible sclera in human eyes, which facilitates gaze tracking and thereby makes intentions more transparent (Kobayashi & Kohshima, 1997; Tomasello et al., 2007). As a result, self-commitment not only serves personal goals but also fulfills a prosocial function, aligning individual behavior with social expectations and promoting mutual understanding.

From this perspective, humans may spontaneously use self-commitment to coordinate their actions, even when no such coordination is required. For example, when multiple individuals pursue their own goals in parallel—despite being explicitly instructed to focus solely on their own tasks—they may still spontaneously coordinate by reading others’ intentions, revealing their own commitments, and binding themselves to those commitments. This spontaneous coordination can manifest in various ways, including the goals people choose, the timing of their commitment (e.g., whether the presence of others accelerates or delays its revelation), and the strength of their adherence to that commitment. Interestingly, in contrast to the enhanced task performance due to arousal, heightened self-commitment—with its rigidity and persistence in pursuing even a suboptimal goal—could lead to a decrease in task performance.

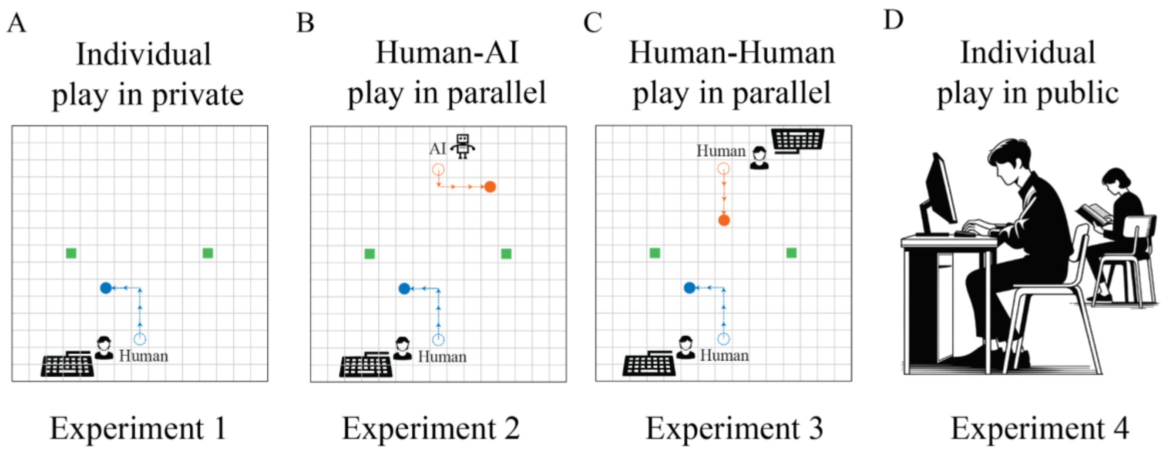
Current study

The currency study investigates how social context specifically shapes individual decision-making—namely, *which* intentions are chosen, *when* those commitments are revealed, and *how firmly* participants commit to them. Rather than assigning participants a predetermined goal and measuring their performance in goal pursuit, this study adopted the "Buridan's ass" paradigm (Cheng et al., 2023), examining how humans form an intention of pursuing one of two equally desirable destinations in a 2D grid map, and commit to their decisions over time. Participants were asked to finish the task as quickly as possible while taking the *fewest* steps. To highlight this, participants were informed that they could earn an additional bonus for completing the task efficiently by following the instructions. Notably, intention and commitment were never mentioned in the instructions.

Experiment 1 of Cheng et al. (2023) demonstrated a "goal perseverance" effect: A majority of participants, after an unexpected drift shifted them closer to an unintended destination and farther from their intended one, still pursued their original, now suboptimal, destination, resulting in poorer navigation performance. We replicated the "goal perseverance" effect in our Experiment 1 as a baseline to measure self-commitment strength when participants acted alone in a private room (Figure 1A). This paradigm was then adapted and integrated into an experimental game setting (Van Dijk & De Dreu, 2021), where two players could potentially adjust to each other's actions in real time while playing on the same game board. Experiments 2 and 3 introduced a social context, where participants played the same individual game alongside another agent, either controlled by artificial intelligence (AI) or a human (Figure 1B and 1C). Experiment 4 further minimized the social context by having participants as the sole players while merely adding the physical presence of another person in the room (Figure 1D). By comparing the "goal perseverance" effects across various social contexts to the baseline from Experiment 1, we can determine whether self-commitment is strengthened or weakened in these settings.

Figure 1

Manipulations of social contexts in an individual decision-making task



Note. Panel A. Experiment 1 set up a baseline where individuals play a decision-making game alone. Panels B and C. A second player will be present in the game, playing the same individual task in parallel with participants. Participants were explicitly told they could pursue any goal regardless of the other agent's plans. In Experiment 2, the second agent was controlled by an AI agent. Panel C. In Experiment 3, the second player was controlled by another human player. Panel D. In Experiment 4, the second agent did not participate in the game, but was present in the same room, sitting diagonally behind the participant, reading a book.

Moreover, the multiplayer game adaptation of the "Buridan's ass" paradigm offers a novel lens for exploring self-commitment within social contexts, moving beyond the "goal-perseverance" effect examined by Cheng et al. (2023). This approach enables us to further examine: (1) Whether participants spontaneously use ToM to infer the intentions of other players; (2) How this ToM inference shapes their own choice of destination to commit to; (3) Whether this ToM-driven coordination influences the temporal dynamics of their intentional commitment, as reflected in how quickly they reveal their intentions through actions.

Transparency and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. Experiments 1, 3, and 4 were pre-registered on

OSF (a link is provided for each experimental section). Data is available at https://osf.io/c8wpu/?view_only=78a34c89ff63481b97bbb0d655b9d752

Experiment 1: Replicate "goal perseverance" in a private room

In this experiment, participants completed the task alone in a private room, where their real-time actions were not observed by others.

Method

Participants

The sample size was set to $N = 50$ to match both the original experiment (Cheng et al., 2023) and recommendations from social science research (Simmons, Nelson, & Simonsohn, 2013). The sample sizes for subsequent experiments were also set to $N = 50$ to allow for direct comparisons. This sample size provided 80% power to detect an effect size of $w = 0.39$ or greater in a chi-square test with a 5% false-positive rate (by G*Power 3.1). This sensitivity power analysis information will not be repeated in subsequent experiments, as the sample sizes are the same.

A total of 50 participants ($M_{age} = 22.55$, range = 18-29) were recruited from the pool at XXX University. Each participant received monetary compensation (~ $1.5) for their participation, and an additional bonus (~ $0.5) for completing the task following instructions. Both this experiment and the subsequent experimental sessions lasted approximately 15 minutes.

This experiment, as well as the subsequent ones, was pre-reviewed and approved by the Institutional Review Board of the Department of Psychology at XXX University. All participants in this study provided informed consent prior to participating in the experiments. No participants were excluded from the analysis. This experiment was preregistered at https://osf.io/zw8by/?view_only=f785755c640c4d09b4f5cb7001cbba51

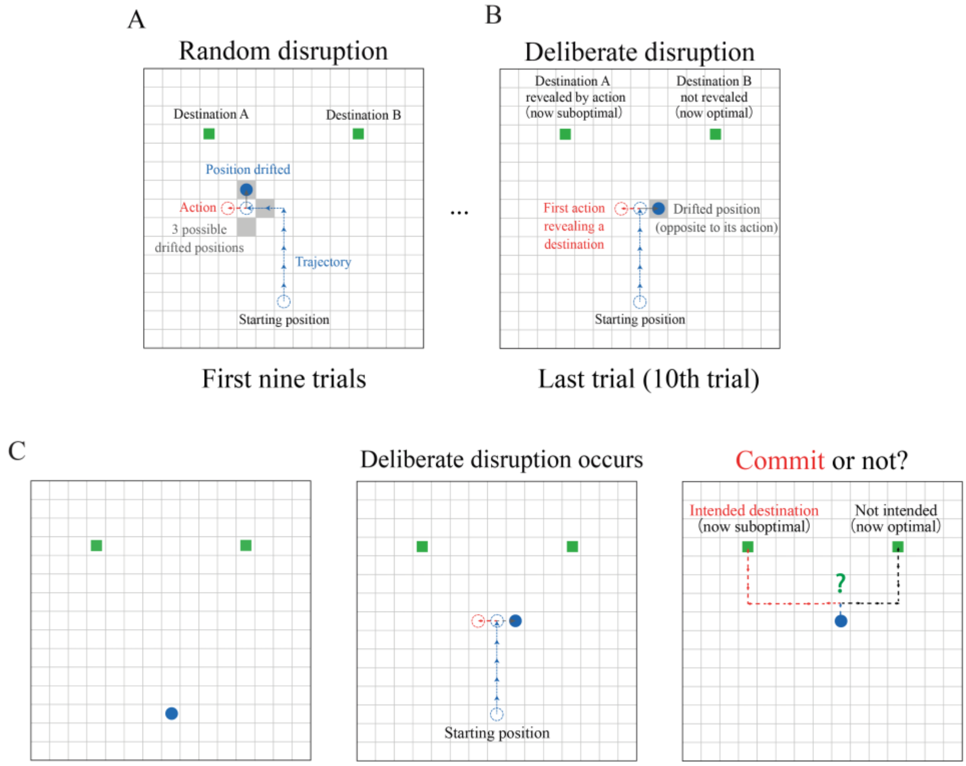
Design and Procedure

There were 10 trials per experiment. Each trial consisted of a 15×15 grid map presenting an agent and two destinations. The map was presented at a visual angle of approximately 21° × 21° from a viewing distance of 60 cm on a 24-inch monitor. Both destinations were presented in the same color and were always positioned either vertically or horizontally on the map, such that they and the agent's starting position formed an isosceles triangle with equal Manhattan distances to each target. This design ensured that the two goals were visually and spatially equivalent, making them equally preferable from the participants' perspective.

Participants were explicitly informed that the environment was not deterministic: at each step, there was a 10% probability that the agent's action could be interrupted by a random drift pushing it to an adjacent cell (Figure 2A). In the initial nine trials, as per the instructions given to the participants, disruptions occurred randomly, resulting in approximately one drift per trajectory. For each participant, both the timing and direction of these disruptions were independently randomized. However, the disruption in the final trial was not random, but deliberately generated. It was triggered precisely when the participant-controlled agent first indicated its destination by moving towards one destination while away from the other, causing the agent to drift one cell and end closer to the unchosen destination (Figure 2B). "Goal perseverance" was measured by the agent's decision to continue towards its originally intended destination, despite the unintended one now being closer (Figure 2C). To prevent arousing participants' suspicion, the deliberate disruption was introduced only once in the last trial as the "critical trial" (Mack & Rock, 1998) of the experiment. This disruption was disguised by the first nine trials of random disruptions, which served to establish the impression that all disruptions were "random," according to the instructions given to the participants. Following Cheng et al. (2023), in the analysis of "goal perseverance", we will focus on utilizing data from the "critical trial". However, we will also utilize data from the first nine trials as a baseline for other measurements of self-commitment, which will be covered in Experiments 2-4.

Figure 2

Measure commitment as "goal perseverance"



Note. Panel A. In the first nine trials, both the time step and the direction of the disruptions were randomly sampled. Panel B: Design of deliberate disruptions. In the last trial, both the time step and the direction of the disruptions were deliberately designed to push the agent away from the destination the moment it was revealed. Panel C: This is a sampled critical trial with deliberate disruption. Commitment is measured by whether the agent would still pursue the originally intended but now sub-optimal destination after the deliberate disruption. All dashed lines are for illustration; they are not visible in the real experiment. Figures adapted from Cheng et al. (2023).

Results

In the final trial with deliberate disruption, 56% of participants persisted with their originally intended, yet ultimately suboptimal, goal—a rate far exceeding the ideal of 0%

expected from a perfectly rational, utility-maximizing agent (binomial exact test, $p < .001$). Recognizing that human decision-making always involves some degree of noise, we established a human rationality baseline by measuring the percentage of non-optimal path selections in the first nine trials with random disruptions. In those trials, participants reached a suboptimal destination in only 1.8% of the cases (95% CI: [0.6%, 3.0%]). This stark contrast—56% with deliberate disruptions versus 1.8% with random disruptions—is statistically significant (binomial exact test, $p < .001$). These findings suggest that participants navigated rationally—except when a deliberate disruption turned their intentional commitment against the optimal path. In these trials, the high rate of suboptimal choices reflected a systematic bias toward maintaining initial intentions rather than random error. This replicates the "goal perseverance" effect in Cheng et al. (2023).

Experiment 2: human-RL play in parallel

In this experiment, we introduced a social context by having participants play alongside another “person” in parallel (Figure 1B). Pair of participants played the game simultaneously but in separate rooms. They were aware of each other’s presence and were informed that they could observe the other player’s actions on their monitors. However, the “other player” they were observing was actually an artificial intelligence (AI) agent whose actions were generated by a reinforcement learning (RL) model based on a Markov Decision Process (MDP). This MDP-RL was a commitment-free agent (Cheng et al., 2023), meaning that it does not pre-commit to a single goal or plan. Instead, it relies on a soft-max policy to make decisions probabilistically, prioritizing actions with higher expected rewards. When faced with two equally desirable destinations, many state-action pairs share the same expected rewards (see Figure 4A for a visual of expected values and policy). In such cases, the policy assigns equal probability to these actions and resolves ties through random sampling. Notably, this classic single-agent algorithm functions independently, with no interaction with humans.

To emphasize the individual nature of the task, participants were explicitly informed that their actions in the game did not interact or affect one another in any way and that it did not matter whether they chose the same destination.

Method

Participants

The sample size for this experiment, as well as subsequent ones, was set to match Experiment 1 to enable a direct comparison. Fifty participants (comprising 25 pairs; M_{age} = 20.98, range = 18-27) recruited from the pool at XXX University. Each participant received monetary compensation (~ $1.5) for their participation, and an additional bonus (~ $0.5) for completing the task following instructions.

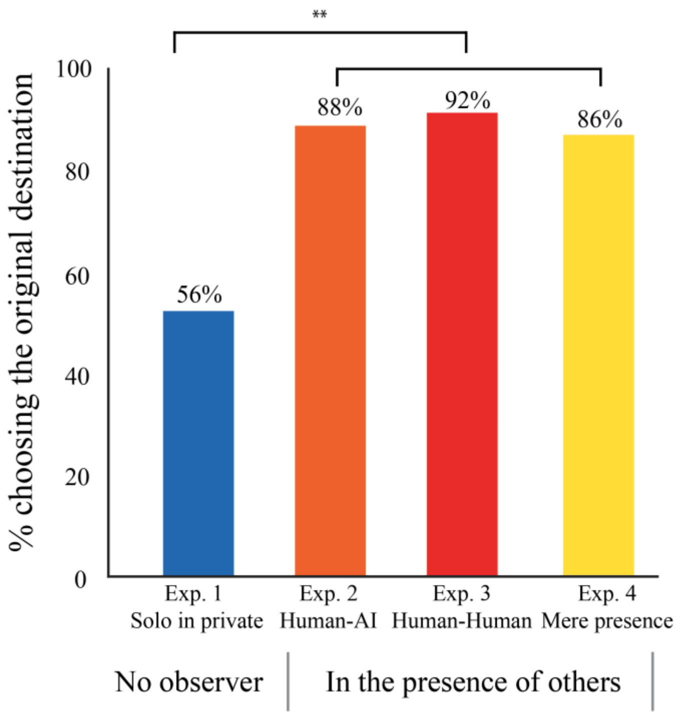
Design and Procedure

Pairs of participants received instructions simultaneously from an experimenter. The main task was identical to that in Experiment 1, requiring participants to reach one of the destinations as quickly as possible using the fewest steps. Participants were explicitly told that their tasks were independent: there would be no interaction or interference between their actions or rewards. They were also informed that they were free to enter the same grid and reach the same endpoint in the game. After receiving instructions, participants were placed in separate rooms to begin the main experiment. Each participant used a computer display in their respective rooms to view the game. The two computers were connected by a long cable, creating the impression that participants were seeing each other’s moves on a shared screen. In reality, they were observing the actions of an MDP-RL agent. The MDP-RL’s actions were synchronized with those of the human participants, who had full control over when to move. After each human keyboard input, the MDP-RL immediately executed its own action, so that both agents' actions appeared simultaneously on screen.

At the start of each round, both agents were positioned on the map such that (1) each agent was equidistant from both destinations, and (2) both agents had the same distance to each destination. The random disruptions are randomized for each agent.

Figure 3

Enhanced "goal perseverance'' across different social contexts



Note. Percentage of participants who reached the originally intended destination in the last deliberate disruption trial. ***p* < .01

Results

Enhanced "goal perseverance"

In the last trial with deliberate disruption, 88% of participants chose the originally intended but later suboptimal goal—significantly more than in Experiment 1 of solo play

(88% vs. 56%, $\chi^2(1) = 12.7$, $p < .001$, Cramer's $\varphi = 0.36$; Comparative analysis (e.g., chi-square test, *t*-test) in this and subsequent studies was done using Python's (3.10) researchpy package). Consistently, participants deviated more from the optimal number of steps here when playing with others than when playing alone in Experiment 1 (mean deviated steps: 1.88 vs. 1.24, $t(98) = 3.67$, $p = .0004$, Cohen's $d = 0.73$). These results suggest a significant increase in "goal perseverance" and a significant reduction in task performance during the deliberate disruption trial, due to the presence of a second agent.

In the first nine trials with random disruptions, we also found a drop of task performance, where the presence of others led participants to deviate more from the optimal number of steps taken to reach a goal, defined as the number of additional steps relative to the best possible path that accounts for the effects of random disruptions (1.67 vs 0.51, $t(98) = 8.47$, $p < .001$, Cohen's $d = 1.69$). These additional steps are likely due to coordination triggered by the presence of a second agent, as we will discuss below.

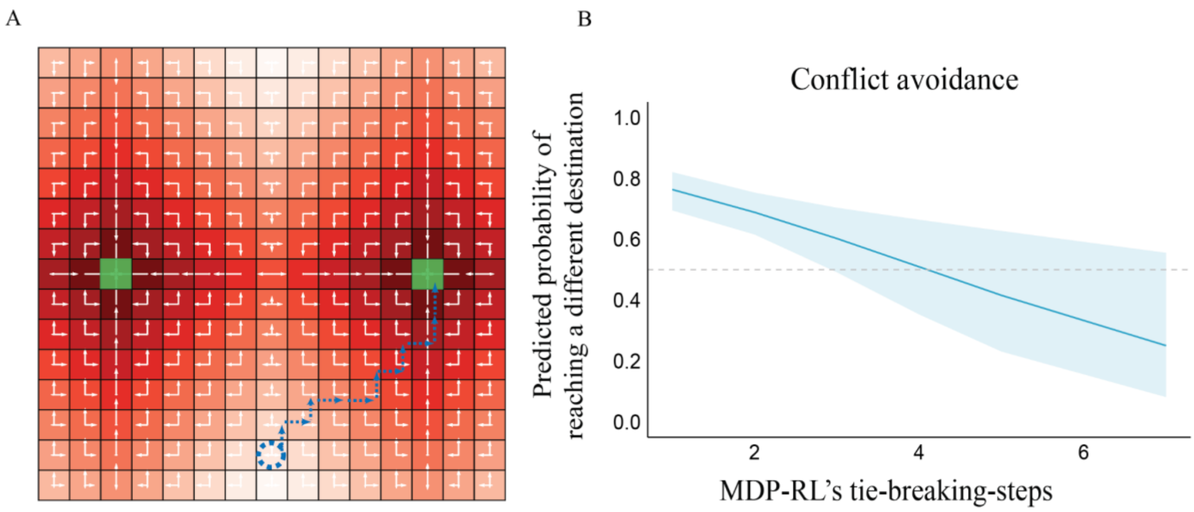
We now move beyond the sole focus on the last deliberate disruption trial as in Cheng et al. (2023), and analyze navigation data from all 10 trials to explore a variety of potential spontaneous intention-based coordination due to the multilayer nature of the task. Since the MDP-RL was a purely individual model without any consideration of the human actions, any observed interaction would be attributed unilaterally to the human participants.

Spontaneous intention coordination

We first examined whether the human participants' choice of destination was influenced by the destination selected by the MDP-RL. If participants strictly adhered to the instructions and made their decisions independently of the MDP-RL, there should be an equal probability of choosing the same destination as the AI or a different one. Nevertheless, the results showed that participants chose a different destination from the MDP-RL in 70.2% (95% CI [0.64, 0.76]) of the trials. A fitted mixed-effect logistic regression model (R, lmer package), predicting the likelihood of reaching different destinations with a random by-subject intercept, revealed that this human tendency to choose different destinations occurred more frequently than the chance level of 50% ($\beta_{intercept}$ = 1.05, OR = 2.87, 95% CI [1.99, 4.15], $p < .001$). These results revealed spontaneous, task-irrelevant coordination driven by humans accommodating the MDP-RL. This behavior involves at least two cognitive processes: (1) spontaneously predicting the MDP-RL's destination, a process clearly demanding ToM, and (2) interpreting heading toward the same destination as the other agent as a "conflict," leading to a bias toward choosing a different one, possibly to avoid "stepping on others' toes."

Figure 4

The impact of MDP-RL's actions on human choice of destination



Note. Panel A: An illustration of the MDP-RL's softmax policy in our grid world task environment. The color of each grid represents the expected future reward starting from the current state until reaching the goal, with darker colors representing higher expected rewards. The white arrows in each grid represent the optimal policy as the distribution of actions in each state. In the states where the agent is placed in between the two destinations (green squares), there is an equal probability to move in three directions (left, right and straight). The blue dashed lines represent a trajectory with randomly sampled actions from the softmax policy. Panel B. A fitted mixed-effect logistic regression predicting "conflict avoidance" from MDP-RL's tie-breaking-steps (plotting by R, ggeffects package). The error shadows reflect 95% confidence intervals.

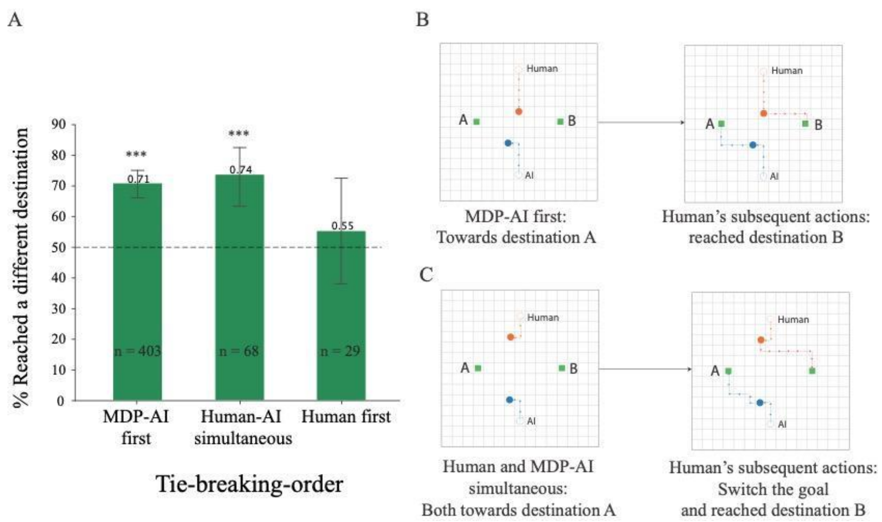
We next investigated how "conflict avoidance" interacts with the timing of when two agents—a human and an MDP-RL—reveal their destinations. Specifically, we examined whether the human's tendency to avoid selecting the same destination as the MDP-RL depends on which agent reveals its destination first. One hypothesis suggests a generic bias: humans switch to a different destination whenever they can predict that both agents are targeting the same one, regardless of who reveals their intention first. If true, the rate of "conflict avoidance" would remain constant, unaffected by the speed of the MDP-RL's revelation. Alternatively, "conflict avoidance" might stem from a social norm of respecting the first agent to "claim" a destination by signaling intent. In this case, humans would choose a different destination only if the MDP-RL reveals its intention first; if the MDP-RL's revelation is delayed, the human may have already committed and would stick to their choice, thereby reducing "conflict avoidance".

To test these hypotheses, we defined the "tie-breaking step" as the critical moment when the MDP-RL first moves toward one destination while moving away from the other, allowing an observer to predict its final destination. For example, Figure 4A shows a sample path of an MDP-RL that breaks the tie at step 2. Figure 4B depicts the strength of "conflict avoidance" as a function of the MDP-RL's tie-breaking step. We conducted a mixed-effects logistic regression model to predict "conflict avoidance" from the MDP-RL's tie-breaking step, with random intercepts included for individual players. There was a significant main effect of the MDP-RL's tie-breaking-step (β = -0.33; OR = 0.72, 95%CI [0.57, 0.91]; p = .005), suggesting that the later the MDP-RL revealed a destination, the less likely humans were to choose a different destination.

To further confirm “conflict avoidance” is constrained by intention-order, we bin the data into three groups based on the tie-breaking-order of the two agents: MDP-RL first, human and MDP-RL simultaneous, and human first (Figure 5A). Indeed, when the MDP-RL was the first to break the tie, which constituted the majority of the trials (80.6% of trials or 403/500), humans were more likely to choose a different destination compared to the chance level (71% or 285/403, binomial $p < .001$; Figure 5B illustrates a typical example). Nevertheless, in the few trials where humans revealed a destination first (in 5.8% of trials, or 29/500), the percentage of reaching a different destination (55.2% or 16/29) did not significantly differ from the chance level (binomial $p = .71$, $BF_{01} = 2.57 \times 10^8$). These results support the hypothesis that human "conflict avoidance" is not merely a generic bias but is shaped by the social norm of deferring to the first agent who signals their intention to claim a destination.

Figure 5

The effect of tie-breaking-order on human's "conflict avoidance"



Note. Panel A. Percentage of reaching a different destination based on who (human or MDP-RL) first breaks the tie. Panels B and C. Two representative trajectories of human "conflict avoidance". ***$p < .001$. Error bars indicate a 95% confidence interval.

Interestingly, we also found that when the MDP-RL and humans revealed their destinations simultaneously (in 13.6% of trials, or 68/500), humans still tended to arrive at a different destination (73.5% or 50/68, binomial $p < .001$; See Figure 5C as an example). This avoidance behavior, which cannot occur at the moment humans break the tie but only afterwards, indicates that humans retracted their commitments by changing their intention to a different destination, after recognizing a conflict with the MDP-RL's actions (Figure 5C). Consistently, we found that humans demonstrated high commitment to their initial choice in both the MDP-RL first and human first groups (96.77% and 93.1%, respectively; mean 96.53%), but this commitment dropped significantly to 77.4% in the human-RL simultaneous group when they revealed same destination at the same time (96.53% vs 77.4%, Fisher's exact test, odds ratio = 0.12, $p = .0003$, Cramer's $\varphi = 0.22$). The high-level commitment in human-first groups again suggested that if humans commit to a destination first, they are unwilling to "give in" when another agent later navigates toward it.

Overall, these results suggest that humans spontaneously apply a ToM to infer the MDP-RL's destination and use it to adjust their own actions. Moreover, humans did not yield to the other agent unconditionally, but only when their intention was shown later than the MDP-RL's, suggesting that humans may develop a sense of ownership simply because the other agent demonstrated an intention to reach it—a form of intention-based ownership. This interpretation will be further discussed in the General Discussion. After uncovering the subtle and delicate coordination between the two agents, it's worth reminding readers that these effects reflect unilateral accommodation by humans to the MDP-RL, which completely ignored human actions. Moreover, all these human accommodations were spontaneous and task-irrelevant, given the individual nature of their task.

Human intention in the eyes of a Bayesian observer

All the extra task-irrelevant consideration of coordination may complicate human decision-making and influence when humans reveal their intentions through their actions. To investigate, we applied a Bayesian ToM (BToM) model (Baker et al., 2009) to infer human intentions from their actions over time, using trajectories from both Experiment 1 and Experiment 2.

$$P(\text{ Intention} \mid \text{Action}_{1:T}, \text{Environment}) \propto \prod_{t=1}^{T} P(\text{Action}_t \mid \text{Intention, Environment}) P(\text{Intention} \mid \text{Environment}) \quad (1)$$

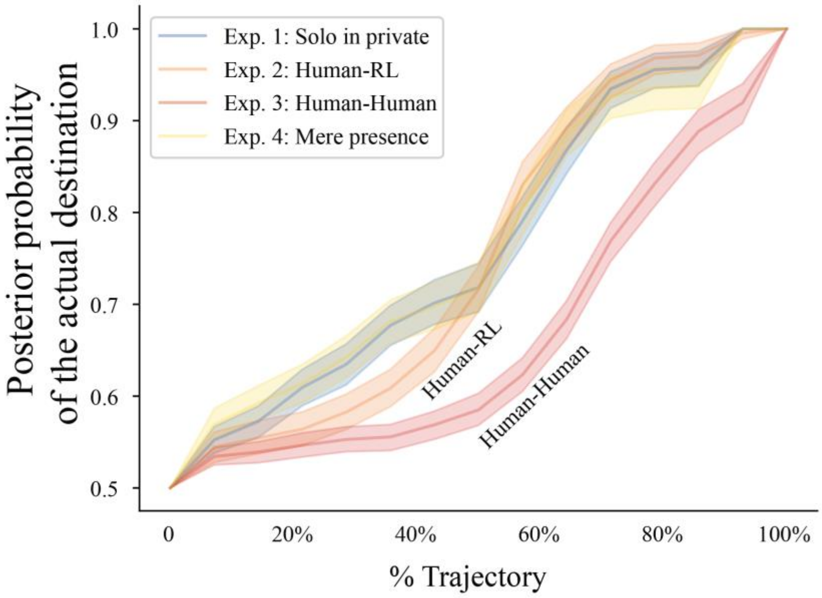
This BToM model assumes that human action is a rational process, and it infers the most likely intention that best explains the action trajectory so far (see Equation 1). The output of this inference process is the posterior probability of each destination, representing the intention that the agent is likely pursuing. We implemented this BToM model (following the detailed math in Baker et al., 2009) using Python (version 3.10). A key component of the model is the likelihood functions $P(\text{Action} \mid \text{Intention, Environment})$, which were derived from an optimal MDP policy that treats one destination as its goal. This policy was solved using value iteration based on the Bellman optimality equation (Bellman, 1957). To account for the inherent variability in human actions, we used a softmax policy with a temperature parameter β that controls action noise, which was fitted to $\beta = 2.5$ based on human navigation data in prior work (Cheng et al., 2023).

The impact of the MDP-RL on human decision-making can be revealed by applying BToM to infer the posteriors of human intentions over time, and comparing them to the intention posteriors of Experiment 1 conducted in a no-social-context setting. The BToM results showed that initially, humans revealed their intentions more slowly during navigation (Figure 6). The cluster-based permutation tests (Maris & Oostenveld, 2007) identified a significant gap from 20% to 46.7% of the trajectories (1000 permutations, $p = .011$, $T_{\max} =$

6.82, Cohen's $d = 0.004$). One factor contributing to the prolonged human decision process might be that humans were waiting for the MDP-RL's decision, to avoid potential conflicts.

Figure 6

Human intentions in the eyes of a BToM observer



Note. The posterior of the BToM inference for the agent's actual destination as a function of the agent's steps over time. The error shadows reflect 95% confidence intervals.

However, one must be cautious in generalizing the above findings of human tendencies, as they might be specific to interactions with the MDP-RL, which acts very differently from humans. These differences are highlighted by (1) the MDP-RL always broke a tie quickly using random sampling and (2) the self-centered nature of the MDP-RL, which ignores human actions as irrelevant to its own reward. Practically, due to this certain type of MDP-RL's actions, there were very few trials where a human player broke the tie first or simultaneously with the MDP-RL. We aimed to address these limitations in a subsequent experiment with two human players playing in parallel.

Experiment 3: human-human play in parallel

This experiment mirrored the design of Experiment 2, with the key difference being that the second agent in the game was actually controlled by a real human player.

Method

Participants

The sample size and recruiting procedure for this experiment were consistent with Experiment 2. We recruited 50 participants (comprising 25 pairs; M_{age} = 20.44, range = 18-25) from the pool of XXX University. Each participant received monetary compensation (~ $1.5) for their participation, and an additional bonus (~ $0.5) for completing the task following instructions.

Design and Procedure

The procedure for this experiment was identical to that of Experiment 2, except that after the instructions, participants played the game in a single room using one computer equipped with two monitors and two keyboards. The participants sat opposite each other on either side of a large table, each facing their respective monitor. The two monitors were wired together, displaying identical content. In this experiment, both human players had full control over when to move, and their actions were not synchronized. Each participant moved at their own pace in real time. [note: although the two players were instructed to proceed at their own pace and there was no action synchronization, they nonetheless navigated at roughly the same speed, with a mean decision time difference of only 44 ms per step]. The design and analysis plan for Experiment 3 was preregistered at https://osf.io/gzvf3/?view_only=9c51337dec84440dad5d199679bc4111

Results

Enhanced "goal perseverance"

In the last trial with deliberate disruption, participants were more likely to choose the originally-intended but later-suboptimal destination compared to Experiment 1 (92% vs. 56%, $\chi^2(1) = 16.84$, $p < .0001$, Cramer's $\varphi = 0.41$). Consistently, participants deviated more from the optimal number of steps compared to Experiment 1 (mean deviated steps: 2.48 vs. 1.24, $t(98) = 5.52$, $p < .001$, Cohen's $d = 1.10$). This result replicates the findings of Experiments 2, indicating that the presence of a second agent can indeed lead to higher "goal perseverance" and decreased task performance.

In the first nine trials with random disruptions, the presence of another human player also led participants to deviate more from the optimal number of steps compared to Experiment 1 (1.00 vs 0.51, $t(98) = 6.36$, $p < .001$, Cohen's $d = 1.27$). These additional steps are likely attributable to potential coordination triggered by the presence of a second agent, as we will discuss below.

Spontaneous intention coordination

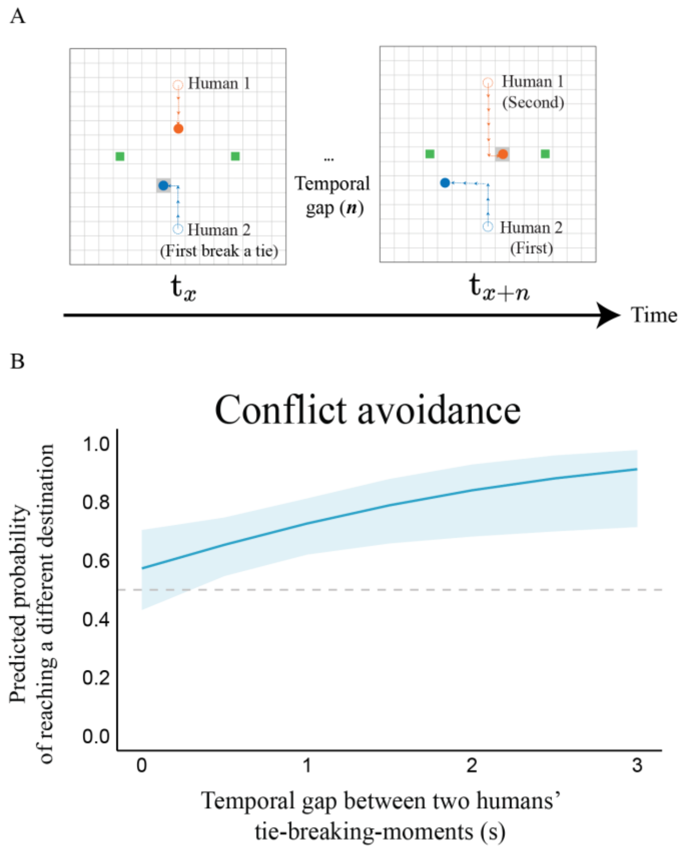
Across all trials, the "conflict avoidance" results showed that, in 66% (95% CI [0.57, 0.75]) of the trials, pairs of participants ultimately reached different destinations. A mixed-effect logistic regression model, predicting the likelihood of reaching different destinations with a random by-subject intercept, revealed that this human tendency to reach different destinations occurred more frequently than the chance level of 50% ($\beta_{\text{intercept}} = 0.75$, OR = 2.12, 95% CI [1.4, 3.2], $p < .001$).

We then measured the “temporal gap” between the two player’s tie-breaking-moments (Figure 7A), to explore how "conflict avoidance" was influenced by when the two humans revealed their intentions. We conducted a mixed-effect logistic regression to predict "conflict avoidance" from the temporal-gap, with each individual included as a random effect. Figure 8B shows the fitted curve. This analysis revealed a significant positive coefficient for the temporal-gap ($\beta = 0.68$, OR = 1.98, 95% CI [1.10, 3.57], $p = .023$; Figure

7B), suggesting that a larger temporal-gap leads to stronger "conflict avoidance". One possible interpretation is that the longer the temporal gap, the more likely it is that the second players, who reveal their intentions later, will take into account the intentions of the first player and are more likely to choose a different destination to avoid conflict. Together, these results extended the findings of "conflict avoidance" from unilateral human-RL interactions to bilateral human-human interactions, suggesting that humans spontaneously coordinate their actions based on inferred intentions regarding the other's destination.

Figure 7

Spontaneous intention coordination



Note. Panel A. Temporal gap (n) between two players' tie-breaking-moments. Panel B. Predicting "conflict avoidance" from the temporal gap between two humans' tie-breaking-moments. The error shadows reflect 95% confidence intervals.

Human intention in the eyes of a Bayesian observer

In addition to the preregistered tests of “goal perseverance” and “conflict avoidance,” we now conduct the preregistered exploratory analysis to explore how human intentions manifest in the eyes of a rational observer implemented by the BToM model.

Across all trials, the BToM analysis showed that, compared to Experiment 1 with no-social-context, participants revealed their intentions more slowly during their navigation (Figure 6). Cluster-based permutation tests (1000 permutations) found significant differences between 13.3% to 93.3% of the trajectories ($p = .001$, $T_{max} = 105.96$, Cohen's $d = 0.48$). These periods of delayed intention revelation suggest that humans may exert additional effort to coordinate their intentions with each other. It is notable that the delay in intention revelation here was much more extensive than that in Experiment 2 (cluster-based permutation tests found significant differences between 33.3% to 93.3% of the trajectories, $p < .001$, $T_{max} = 157.33$, Cohen's $d = 0.54$). This is likely because the two-way coordination between two real human participants is much more subtle and engaging than the one-way unilateral human accommodation of the MDP-RL.

Experiment 4: Mere presence of another person

In Experiments 2 and 3, we observed several effects of social context on individual commitment: strengthened "goal perseverance", decreased navigation performance (even in trials with random disruptions), and delayed intention revelation. A critical question arises: to what extent are these effects attributable to the mere presence of another person versus the presence of another agent as a potential partner in the same game environment? This distinction aligns with the theoretical contrast between an objective third-person perspective (a public nature of environment) and a specific second-person perspective (interacting with a potential partner). Both perspectives represent forms of social context that can influence individual decision-making, but they may operate through different mechanisms. The mere-presence effect might stem from evaluating and even censoring one's own action from an objective third-person perspective capturing generic social norms, while the partner effect could involve taking a second-person perspective for coordinating with a partner in a specific task. The current experiment aims at disentangling these influences, so as to better understand how social context shapes self-commitment.

In this experiment, participants completed the navigation task while a passive observer was merely present in the room – not a participant controlling an agent in the game environment (Figure 1D). We intentionally introduced only a weak social context, wherein the observer was instructed to sit quietly and read a book, without interacting with the participants or directly observing them. To avoid arousing the participants' suspicions, this individual was introduced as an experimental assistant who would remain in the room to assist with any computer-related issues. Both the participants and this second person were unaware of the experiment's true purpose.

Method

Participants

The sample size was the same as in Experiment 1. A total of 50 adult participants (M_{age} = 22.2, range = 18-25) were recruited from the participant pool at XXX University for the experiment. Each participant received monetary compensation (~ $1.5) for their participation, and an additional bonus (~ $0.5) for completing the task following instructions.

Design and Procedure

Except for the introduction of a second person, the design and procedure remain the same as Experiment 1. This experiment was preregistered at https://osf.io/tjwc3/?view_only=64c9220e3605442389d9ba7e3f58f81c

Results

In the last trial with deliberate disruption, 86% of participants reached the originally-intended but later-suboptimal goal, which was significantly higher than in Experiment 1 (86% vs. 56%, $\chi^2(1) = 10.93$, $p < .001$, Cramer's $\varphi = 0.33$). Consistently, participants deviated more from the optimal number of steps compared to Experiment 1 (mean deviation steps: 1.88 vs. 1.24, $t(98) = 3.67$, $p < .001$, Cohen's $d = 0.73$).

In contrast, in the first nine trials with random disruptions, task performance—as measured by deviation from optimal shortest path—did not statistically differ from that observed in Experiment 1 (0.53 vs 0.51, $t(98) = 0.46$, $p = .65$, Cohen's $d = 0.09$; $BF_{01} = 0.90$). This null result contrasts sharply with Experiments 2-3, where social contexts significantly decreased navigation performance. These findings suggest that while the mere presence of an observer was sufficient to reinforce "goal perseverance" and impair navigation performance during the deliberate disruption trial, it did not affect general performance under random disruptions. This indicates that the decreased overall navigation performance in Experiments 2-3 was driven by spontaneous coordination with the other agent in the game.

The BToM results (Figure 6) showed that there was also no difference in intention-revealing speed between this experiment with a mere observer and Experiment 1 with no observer. Cluster-based permutation tests did not reveal any significant differences in the posterior probabilities of inferred destination between the two experiments (no cluster was found with 1000 permutations using a threshold of $p = .05$, $T_{\max} = 2.20$), suggesting that the mere observer had no impact on participants' intention-revealing speed.

Overall, the results show that both the third-party observer and second-party partner perspectives influence the effects of social context on self-commitment in Experiments 2–3, albeit in distinct ways. The mere presence of an observer is sufficient to heighten self-commitment, resulting in a stronger "goal perseverance" effect. However, the general decline in navigation performance and the delay in intention revelation are attributed to spontaneous coordination with a potential partner.

General Discussion

Our study demonstrates that spontaneous self-commitment is significantly shaped by social contexts, as evidenced by variations in its strength, content, and timing across diverse settings (see Table 1 for a summary of effects.). These findings reveal that acting in the presence of others can significantly influence how humans commit to their intentions in their individual tasks, underscoring the intrinsic social dimension of self-commitment. Moreover, the distinct ways in which these aspects respond to different social environments provide a robust framework for investigating the underlying social-cognitive mechanisms that govern both the formation of intentions and the processes by which commitment to those intentions is enforced.

Table 1

Summary of different effects of social context on self-commitment

	Enhanced "goal perseverance" with deliberate disruption?	Drop of navigation performance with deliberate disruption?	Drop of navigation performance with random disruptions?	Delayed intention revelation?	"Conflict avoidance"?	Temporal gap between intentions affects "conflict avoidance"?
Experiment 2 (human-RL)	Yes	Yes	Yes	Yes	Yes	Yes
Experiment 3 (human-human)	Yes	Yes	Yes	Yes	Yes	Yes
Experiment 4 (mere presence)	Yes	Yes	No	No	Not apply	Not apply

This study investigates the influence of three different social contexts on spontaneous self-commitment, each compared to a baseline of private, solitary play. Experiment 3 involved the strongest social setting, where two humans played the game in parallel. Experiment 2 featured a weaker social context, with participants unknowingly playing alongside an AI agent. Experiment 4 had the weakest social setting, where participants played alone in a public environment with another person seated nearby. The results showed that the strength of the social context influenced the extent of self-commitment effects: Experiment 3 exhibited the most significant effects, while Experiment 4 showed the most limited results.

It is noteworthy that the specific effects of social context on self-commitment should be distinguished from the boosted task performance due to a general effect of enhanced arousal (e.g., Zajonc, 1965) or increased cognitive resources (e.g., Shteynberg & Galinsky, 2011) accompanying social contexts. Commitment to a goal or coordinating with another agent was never part of the task instructions and was not how task performance was measured. In fact, across various experiments, we consistently found that strengthened self-commitment and spontaneous intention-based coordination led to decreased navigation performance, not an increase. Prior work showing performance benefits from social presence typically involves tasks with externally assigned goals, where social presence boosts motivation and cognitive effort toward goal execution (Zajonc, 1965). In contrast, our task required participants to select their own goals. In such a process, social presence may introduce additional considerations—such as coordination through ToM reasoning—that shape goal selection and commitment, sometimes at the expense of task efficiency. This distinction helps explain why performance decreased in our setting despite the established benefits of mere presence in other contexts.

Importantly, our results not only show that social context distracts participants from optimal behavior, but also reveal the underlying social-cognitive mechanisms such as spontaneous ToM, sensitivity to intention ordering, and the temporal delay in revealing intentions. The discussion below will focus on the special mechanism through which social contexts impact self-commitment. We will first address the minimal effects observed in Experiment 4, followed by an analysis of the findings from Experiments 2 and 3, both of which involved parallel play.

Enhanced goal preservation in a public space

In Experiment 4, participants were the sole players of the game, meaning there was no explicitly mentioned second-person component that was typically required in studies on prosocial behaviors (Engelmann et al., 2012; Ma et al., 2020). Consequently, any observed effects stem solely from an objective third-person perspective, triggered by another person's presence in the room. Previous research suggests that socialization shapes behavior through a normative third-person perspective, where individuals assess their actions from a societal observer's viewpoint. Supporting this, studies show that young children naturally adopt this normative perspective when evaluating others' decisions in social contexts, rating those who consider others' goals in mind as "nicer," even when such considerations are irrelevant to their own goals (Zhao et al., 2021). More importantly, this normative perspective can be utilized to evaluate both others' actions and control one's own. This aligns with the broader literature on moral action and obligation (Tomasello, 2020), suggesting that children and adults actively enforce social norms to regulate behavior, even if it may be against their own personal interests. By judging whether their behavior aligns with societal norms (e.g. fairness), they filter out unapproved actions, executing only those passing this self-regulatory process (Schmidt, & Tomasello, 2012; Blake & McAuliffe, 2011). Within this framework, commitment to personal goals emerges as a virtue encoded in the normative third-person perspective. Research on public commitment supports this, showing individuals are more likely to follow through on intentions after public declarations (e.g., Hollenbeck et al., 1989; Bryan et al., 2010).

In line with this perspective, Experiment 4 revealed a significantly stronger "goal perseverance" effect. Specifically, a growing percentage of participants continued to pursue their original destination, even though it became a suboptimal choice following a deliberate disruption in the final trial. This result indicates that a potential observer's presence shifts the perceived environment from private to public, heightening awareness of being observed. In this public context, participants spontaneously enhance self-regulation, boosting self-commitment—as if abandoning their goal post-declaration were unacceptable.

Meanwhile, it is also noteworthy that the timing of intentional commitment was not significantly affected by merely adding another person to the testing room. General navigation performance with random disruptions was also not impaired by adding an observer. Together, these findings indicate the limitations of a social context that only involves a third-party perspective.

"Conflict avoidance" through spontaneous intention-based coordination

In Experiments 2 and 3, participants shared a game board with another agent (AI or human) but were instructed to focus on their independent tasks. Unlike Experiment 4's observer, who remained outside the game, the shared game environment and real-time visibility of the other agent's actions in Experiments 2 and 3 permitted potential coordination, despite explicit instructions to work independently. This created a social context rooted in a second-person perspective ("you and I") between potential partners (Engelmann & Rapp, 2018; Asaba & Gweon, 2022), beyond the abstract third-party perspective ("being observed by someone") of Experiment 4. The magnitude of "goal perseverance" is increased in these two experiments, leading to poorer task performance in the final trial, as measured by the deviation from the shortest path. However, since this effect is also present in Experiment 4 and does not vary across the three social contexts, it appears insensitive to whether the social context involves a second-person perspective or a third-person perspective. Therefore, we will primarily focus on the "conflict avoidance" effect, which only exists in the presence of a potential partner.

Despite the independent nature of their tasks, the two agents' navigation is clearly interdependent, as they ended up reaching different destinations more often than would be expected by chance. Why do participants exhibit such an unnecessary "conflict avoidance" ? We suggest that participants may instinctively treat each destination as an object claimable by only one agent. Picture a busy parking lot: would you swoop into a spot another driver is approaching, even if they're still far off? Most would pause, sensing an unspoken claim tied to the other driver's effort. This intuition suggests our simple individual navigation task triggers a deep intuition of ownership – a question profoundly embedded in law, philosophy, and psychology. In his *Second Treatise of Government* (Locke, 1690/1980), John Locke famously asked: when does something become yours? Is it when you first gather it, consume it, or finally possess it? Locke argued that ownership emerges when you mix your labor with a natural resource, turning it into private property.

The challenge is that there was typically a temporal lag between the start of investment of labor, and the completion of possession, which could fuel disputes in the real world. As illustrated by the 1805 case *Pierson v. Post*, Post pursued a fox, only for Pierson to kill and claim it. The New York Supreme Court ruled that possession, not pursuit, determines ownership, favoring Pierson. Yet, our participants align more with Post's perspective, avoiding destinations others are pursuing as if invested effort stakes a claim. Notably, unlike the legal case, our task offered equal access to non-exclusive goals in a non-competitive context, and participants could "share" the same destination without interference. Still, participants demonstrated a spontaneous avoidance tendency, as if they were actually in a social interaction task where ownership of a destination cannot be shared. This suggests that the cognitive tendencies illustrated in that real-life ownership dispute may be embedded in our social intuitions, such that people spontaneously invoke them even in situations where explicit conflict is absent. Moreover, our results suggest that humans recognize invested labor through ToM—invested labor begins when the intention to pursue is inferred from actions. Participants avoided pursuing the same destination once they recognized that the other agent had signaled an intention to pursue it.

While prior work on ownership has largely focused on possession-based claims (Friedman et al., 2008; Rossano et al., 2011), our findings suggest that ownership can also emerge from an agent's expressed intention. Importantly, this sense of intention-based ownership not only applies to cases of avoiding conflicts, but it also predicts that humans should not yield if they are the ones who first expressed intention. Indeed, we found that in Experiment 2, when humans are the ones who first revealed their intention, they tend to commit to their goal even if the MDP-RL later chooses the same goal.

Certainly, people do not always avoid others' goals—imitation and goal alignment are also common in social contexts, particularly when social affiliation is salient. Developmental research shows that even young children are highly sensitive to the goal-directed actions and often imitate others' goals automatically (Nielsen, 2006; Lyons, Young & Keil; 2007). Notably, many imitation tasks are conducted in pedagogical contexts, typically involving a learner–demonstrator dyad in which the child interacts with an adult rather than a peer who may have conflicting goals. Moreover, the measures of goal imitation often focus on reproducing the manner of an action toward a demonstrated goal, rather than on selecting between competing goals. It would be interesting for future research to explore how spontaneous goal avoidance versus goal alignment varies with different social contexts and how this tendency changes over development. Such research could shed light on the emergence of normative perspectives on commitment and how children navigate competing intentions in more complex, peer-based social settings.

Relatedly, while our studies focus on spontaneous coordination and therefore design a neutral task setting without explicit social instructions, it is possible that participants were nonetheless influenced by their intuition of the social context (e.g., cooperative or competitive). For example, they may have construed the situation as having competitive potential and proactively avoided such a setting by steering away from the same goal. Future work could benefit from directly manipulating task framing to assess how such characteristics interact with spontaneous coordination.

Different social contexts impact the intention formation process differently

While the strength of commitment was enhanced across all social contexts after an intention was formed, our BToM analysis showed that different social contexts had varying impacts on the process of intention formation itself. In Experiments 2 and 3, the presence of a second player in a shared visual display delayed participants' demonstration of intention. This delay likely arose from a spontaneous coordination process, where participants implicitly negotiated who should act first or which goals to pursue.

Importantly, the commitment delay effect was strongest in Experiment 3 with human-human dyads, where two human players may implicitly infer each other's intentions and coordinate actions. This caused a bidirectional negotiation process that took time, making participants more cautious in revealing their commitments. In Experiment 2 with human-RL dyads, the commitment delay effect was present but not as strong, as it was only the human who unilaterally accommodated the MDP-RL. The MDP-RL acted only to maximize its own utilities, ignoring human actions. Still, participants tried to consider the MDP-RL's actions, leading to a delay in revealing their commitments, likely because they waited for the MDP-RL to reveal its intentions first so that they could choose a different destination. Finally, while adding a second agent generally increased commitment, merely having an observer (Experiment 4) did not affect the revealing speed of intention, indicating that commitment delay effect is specific to a negotiation process with a potential partner, and does not extend while just being observed.

From avoidance to presentation of commitment: two opposing stages of intentional actions

Our findings also revealed an intriguing pattern in how humans form commitments. In philosophy and neuroscience, there have been rich discussions regarding how humans can break a tie when facing the equilibrium of two equally desirable choices (Ullmann-Margalit, 1977; Furstenberg et al., 2015). One hypothesis suggests that humans, like MDP-RL, may simply break ties through random sampling. Yet, our findings indicate that human intention formation is more sophisticated than mere random selection, as evidenced by humans' prolonged periods of maintaining the tie. Contrary to MDP-RL, which acts without commitment throughout its trajectories, human intentional actions display two distinct deliberation phases: initially avoiding commitment to any option, followed by committing to a chosen intention. While human decision-making seems unnecessarily sophisticated compared to the MDP-RL, it is reasonable when considering the importance and consequences of making a commitment in real life. These findings suggest that one may not simply assume that humans, being goal-directed agents, will quickly manifest their intentions. Due to the inherent “inertia” and “resistance to re-planning” nature commitment, humans appear to be cautious about entering into a commitment. When individuals have not yet decided which intention to commit to, they tend to deliberately keep their options open by proactively concealing or delaying the revelation of their intentions – a process opposite to the showing or signaling process (e.g., Ho et al., 2021). This concealment period was further extended in Experiment 3, which involved human–human interaction—possibly because potential goal conflicts raised the threshold for commitment. This dynamic between intention concealment and eventual commitment is a compelling phenomenon that warrants further investigation in future research.

Distinguishing two types of third-person perspective in social contexts

Our findings from Experiment 4 suggest that self-commitment can be shaped by social context even when no direct interaction or monitoring occurs. However, the notion of a "third-person perspective" warrants further clarification, as it can refer to distinct psychological constructs. In the present study, we adopt an interpretation grounded in the idea of an objective, normative perspective—sometimes described as a "bird's-eye view" (Tomasello, 2014) or a "view from nowhere" (Nagel, 1986). This perspective is not tied to any particular observer but instead reflects internalized social norms and the generalized viewpoint of the social public. It has been proposed that this objective perspective plays a key role in self-regulation and may serve as one origin of metacognitive awareness (e.g., Carruthers, 2009).

In contrast, another common usage of the third-person perspective refers to the concrete viewpoint of a specific observer who is actively attending to and evaluating the participant's behavior. While our design targets the former—the objective third-person perspective—it remains an open question whether "goal perseverance" in such contexts also depends on the belief that one is being specifically observed. Future research could explicitly compare these two forms of social context by manipulating the potential observer's attentional state or perceptual access (e.g., see design in Sartori et al., 2009 or Engelmann, Herrmann, & Tomasello, 2015). Such work would help clarify whether commitment effects are driven by abstract social norms or by reasoning about a particular individual's evaluative stance, and would contribute to a more precise understanding of the mechanisms linking social presence, ToM, and self-regulation.

Commitment in pure individual context

While our findings emphasize the role of social context in shaping self-commitment, it is also important to note that participants in Experiment 1—conducted in a purely non-social setting—still exhibited a relatively high level of commitment. This suggests that spontaneous commitment can occur even in the absence of immediate social cues. Therefore, it’s worth considering non-social factors that may contribute to self-commitment. One proposed function of commitment is to reduce computational and cognitive cost (Bratman, 1987): commitment stabilizes one's course of action and minimizes the need for constant reevaluation, thus supporting more efficient decision-making in complex environments. Recent work suggests that goal commitment may serve individual cognitive processes in at least two ways. First, it reduces computational load, as participants tend to rely on retrospective valuation—persisting with goals based on prior progress—rather than constantly re-evaluating future outcomes, offering a more efficient strategy in dynamic environments (Aenugu & O'Doherty, 2025). Second, it can guide selective attention. Recent work showed that commitment is reinforced through attentional narrowing: selective attention is biased toward goal-relevant features and away from alternatives, reducing the need to continuously reprocess competing options (Holton et al., 2024).

Although attention has recently been recognized as a critical mechanism in individual commitment, it remains unclear how social contexts may impact the underlying attentional processes. Notably, much of the attention literature has focused on non-agentive settings, where attention is externally guided by task instructions rather than spontaneously generated (Rosenholtz, 2024). In contrast, research in social cognition and ToM has emphasized mental states such as beliefs, desires, and intentions, often overlooking attention as a contributing factor in agentive decisions (Tomasello, 2022). Bridging these gaps may reveal how attention and social reasoning jointly shape commitment—a promising direction for future research.

Conclusion and Future Directions

This study demonstrates that self-commitment, traditionally viewed within individual decision-making frameworks, has profound social roots. Our experiments consistently show that the presence of others—even without direct interaction or explicit coordination demands—influences the strength, content, and timing of commitment. The findings suggest that humans spontaneously incorporate social considerations into ostensibly individual tasks, highlighting an underlying cognitive mechanism shared between self-commitment and social coordination.

Building on the insights gained from this research, several promising avenues for future investigation emerge. First, subsequent studies should further disentangle the distinct contributions of explicit reputation management and implicit normative self-regulation processes. Employing experimental paradigms that manipulate anonymity and visibility more explicitly could clarify the boundary conditions under which reputation concerns specifically drive commitment. Additionally, extending the current findings to developmental populations would provide valuable insights into how the capacity for self-commitment emerges and evolves through childhood. Investigating age-related shifts could further elucidate the cognitive prerequisites and developmental trajectory of social influences on commitment.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT 4o and Grok 3 in order to proofread. After using these tools, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

Aenugu, S., & O'Doherty, J. P. (2025). Building momentum: A computational account of persistence toward long-term goals. *PLOS Computational Biology*, 21(5), e1013054.

Aiello, J. R., & Douthitt, E. A. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice, 5*(3), 163–180. https://doi.org/10.1037/1089-2699.5.3.163

Ajzen, I., & Kruglanski, A. W. (2019). Reasoned action in the service of goal pursuit. *Psychological Review, 126*(5), 774–786. https://doi.org/10.1037/rev0000155

Ajzen, I., Czasch, C., & Flood, M. G. (2009). From intentions to behavior: Implementation intention, commitment, and conscientiousness. *Journal of Applied Social Psychology*, 39(6), 1356–1372. https://doi.org/10.1111/j.1559-1816.2009.00485.x

Arkes, H. R., & Ayton, P. (1999). The sunk cost and Concorde effects: Are humans less rational than lower animals? *Psychological Bulletin, 125*(5), 591–600. https://doi.org/10.1037/0033-2909.125.5.591

Asaba, M., & Gweon, H. (2022). Young children infer and manage what others think about them. *Proceedings of the National Academy of Sciences*, 119(32), e2105642119. https://doi.org/10.1073/pnas.2105642119

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. Cognition, 113(3), 329–349. https://doi.org/10.1016/j.cognition.2009.07.005

Barclay, P. (2013). Strategies for cooperation in biological markets, especially for humans. *Evolution and Human Behavior, 34*(3), 164–175. https://doi.org/10.1016/j.evolhumbehav.2013.02.002

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*(1), 3–26. https://doi.org/10.1037/0033-2909.91.1.3

Bellman, R. (1957). A Markovian decision process. *Journal of mathematics and mechanics*, 679-684.

Blake, P. R., & McAuliffe, K. (2011). “I had so much it didn’t seem fair”: Eight-year-olds reject two forms of inequity. *Cognition*, *120*(2), 215-224.

Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press.

Bratman, M. (2018). *Planning, time, and self-governance: Essays in practical rationality*. Oxford University Press.

Bryan, G., Karlan, D., & Nelson, S. (2010). Commitment devices. *Annu. Rev. Econ.*, *2*(1), 671-698.

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32(2), 164-182.

Cheng, S., Zhao, M., Tang, N., Zhao, Y., Zhou, J., Shen, M., & Gao, T. (2023). Intention beyond desire: Spontaneous intentional commitment regulates conflicting desires. *Cognition*, 238, 105513. https://doi.org/10.1016/j.cognition.2023.105513

Cheng, S., Zhao, M., Zhu, J., Zhou, J., Shen, M., & Gao, T. (2022). Intentional commitment through an internalized theory of mind: Acting in the eyes of an imagined observer. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44, No. 44).

Chu, J., & Schulz, L. (2022). " Because I want to": Valuing goals for their own sake. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44, No. 44).

Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K., & Tomasello, M. (2014). Coordination strategies of chimpanzees and human children in a Stag Hunt game. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1796), 20141973.

Elkind, D. (1967). Egocentrism in adolescence. *Child Development, 38*(4), 1025–1034. https://doi.org/10.2307/1127100

Elster, J. (1987). *The multiple self.* Cambridge University Press.

Engelmann, J. M., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology, 20,* 92–95. https://doi.org/10.1016/j.copsyc.2017.08.024

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PloS one*, *7*(10), e48433.

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2016). The effects of being watched on resource acquisition in chimpanzees and human children. *Animal Cognition*, 19, 147-151.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual review of psychology*, 50(1), 21-45. https://doi.org/10.1146/annurev.psych.50.1.21

Friedman, O., & Neary, K. R. (2008). Determining who owns what: Do children infer ownership from first possession? *Cognition, 107*(3), 829–849. https://doi.org/10.1016/j.cognition.2007.12.002

Furstenberg, A., Breska, A., Sompolinsky, H., & Deouell, L. Y. (2015). Evidence of change of intention in picking situations. *Journal of Cognitive Neuroscience*, 27(11), 2133-2146.

Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, 54(7), 493–503. https://doi.org/10.1037/0003-066X.54.7.493

Gollwitzer, P. M., & Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology*, 73(1), 186–199. https://doi.org/10.1037/0022-3514.73.1.186

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. In M. P. Zanna (Ed.), *Advances in experimental* *social psychology,* Vol. 38, pp. 69–119). Elsevier Academic Press. https://doi.org/10.1016/S0065-2601(06)38002-1

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really *is* a theory. *Mind & Language, 7*(1-2), 145–171. https://doi.org/10.1111/j.1468-0017.1992.tb00202.x

Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children’s understanding of joint commitments. *Developmental Psychology, 45*(5), 1430–1443. https://doi.org/10.1037/a0016122

Hamann, K., Warneken, F., & Tomasello, M. (2012). Children’s developing commitments to joint goals. *Child Development*, 83(1), 137–145. https://doi.org/10.1111/j.1467-8624.2011.01695.x

Higgins, E. T., Rossignac-Milon, M., & Echterhoff, G. (2021). Shared reality: From sharing-is-believing to merging minds. *Current Directions in Psychological Science, 30*(2), 103–110. https://doi.org/10.1177/0963721421992027

Ho, M. K., Cushman, F., Littman, M. L., & Austerweil, J. L. (2021). Communication in action: Planning and interpreting communicative demonstrations. *Journal of Experimental Psychology: General*, 150(11), 2246.

Hollenbeck, J. R., Williams, C. R., & Klein, H. J. (1989). An empirical examination of the antecedents of commitment to difficult goals. *Journal of applied psychology*, *74*(1), 18.

Holton, E., Grohn, J., Ward, H., Manohar, S. G., O’reilly, J. X., & Kolling, N. (2024). Goal commitment is supported by vmPFC through selective attention. *Nature Human Behaviour*, 8(7), 1351-1365.

Jiang, K., Stacy, S., Dahmani, A. L., Jiang, B., Rossano, F., Zhu, Y., & Gao, T. (2022). What is the point? a theory of mind model of relevance. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44, No. 44).

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, *387*(6635), 767-768.

Kushnir, T. (2022). Imagination and social cognition in childhood. *Wiley Interdisciplinary Reviews: Cognitive Science*, 13(4), e1603.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin, 107*(1), 34–47. https://doi.org/10.1037/0033-2909.107.1.34

Locke, J. (1980). *Second treatise of government* (C. B. Macpherson, Ed.). Hackett Publishing.

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, 104(50), 19751-19756.

Ma, F., Zeng, D., Xu, F., Compton, B. J., & Heyman, G. D. (2020). Delay of gratification as reputation management. *Psychological Science*, *31*(9), 1174-1182.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. The MIT Press

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, 164(1), 177-190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Mead, G.H. (1934). *Mind, Self, and Society from the Standpoint of a Social Behaviorist.* University of Chicago Press: Chicago.

Michael, J. (2022). *The philosophy and psychology of commitment*. Taylor & Francis.

Michael, J., & Pacherie, E. (2015). On commitments and other uncertainty reduction tools in joint action. *Journal of Social Ontology*, 1(1), 89-120.

Molinaro, G., & Collins, A. G. (2023). A goal-centric outlook on learning. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2023.08.011

Nagel, T. (1989). *The view from nowhere*. Oxford University Press.

Nenkov, G. Y., & Gollwitzer, P. M. (2012). Pre- versus postdecisional deliberation and goal commitment: The positive effects of defensiveness. *Journal of Experimental Social Psychology, 48*(1), 106–121. https://doi.org/10.1016/j.jesp.2011.08.002

Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology, 42*(3), 555–565.

Rajecki, D. W., Ickes, W., Corcoran, C., & Lenerz, K. (1977). Social facilitation of human performance: Mere presence effects. *The Journal of Social Psychology, 102*(2), 297–310. https://doi.org/10.1080/00224545.1977.9713277

Rochat, P. (2009). *Others in mind: Social origins of self-consciousness.* Cambridge University Press. https://doi.org/10.1017/CBO9780511812484

Rosenholtz, R. (2024). Visual attention in crisis. *Behavioral and Brain Sciences*, 1-32.

Rossano, F., Rakoczy, H., & Tomasello, M. (2011). Young children's understanding of violations of property rights. *Cognition*, 121(2), 219-227. https://doi.org/10.1016/j.cognition.2011.06.007

Royka, A. L., Török, G., & Jara-Ettinger, J. (2023). Guiding Inference: Signaling intentions using efficient action. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 45, No. 45).

Sartori, L., Becchio, C., Bara, B. G., & Castiello, U. (2009). Does the intention to communicate affect action kinematics?. *Consciousness and cognition*, 18(3), 766-772.

Schmidt, M. F., & Tomasello, M. (2012). Young children enforce social norms. *Current Directions in Psychological Science*, *21*(4), 232-236.

Searle, J. (2003). *Rationality in action*. The MIT Press.

Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science, 1*(2), 353–367. https://doi.org/10.1111/j.1756-8765.2009.01024.x

Shteynberg, G. (2015). Shared attention. *Perspectives on Psychological Science, 10*(5), 579–590. https://doi.org/10.1177/1745691615589104

Shteynberg, G. (2018). A collective perspective: Shared attention and the mind. *Current Opinion in Psychology, 23,* 93–97. https://doi.org/10.1016/j.copsyc.2017.12.007

Shteynberg, G., & Galinsky, A. D. (2011). Implicit coordination: Sharing goals with similar others intensifies goal pursuit. *Journal of Experimental Social Psychology, 47*(6), 1291–1294. https://doi.org/10.1016/j.jesp.2011.04.012

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2013). Life after P-hacking. In *Meeting of the Society for Personality and Social Psychology*. https://doi.org/10.2139/ssrn.2205186

Stacy, S., Gong, S., Parab, A., Zhao, M., Jiang, K., & Gao, T. (2024). A Bayesian theory of mind approach to modeling cooperation and communication. *Wiley Interdisciplinary Reviews: Computational Statistics*, *16*(1), e1631.

Staw, B. M. (1976). Knee-deep in the Big Muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior & Human Performance, 16*(1), 27–44.

Tang, N., Gong, S., Zhao, M., Gu, C., Zhou, J., Shen, M., & Gao, T. (2022). Exploring an imagined "we" in human collective hunting: Joint commitment within shared intentionality. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 44, No. 44).

Tomasello, M. (2010). *Origins of human communication*. MIT press.

Tomasello, M. (2014). *A natural history of human thinking*. Harvard University Press.

Tomasello, M. (2020). The moral psychology of obligation. *Behavioral and Brain Sciences*, *43*, e56.

Tomasello, M. (2022). *The evolution of agency: Behavioral organization from lizards to humans.* MIT Press.

Tomasello, M. (2024). *Agency and cognitive development*. Oxford University Press.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), 675–735. https://doi.org/10.1017/S0140525X05000129

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: the cooperative eye hypothesis. *Journal of human evolution*, *52*(3), 314-320.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2), 207-232.

Ullmann-Margalit, E., & Morgenbesser, S. (1977). Picking and choosing. *Social research*, 757-785.

Van Dijk, E., & De Dreu, C. K. (2021). Experimental games and social decision making. *Annual Review of Psychology*, 72(1), 415-438.

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative Activities in Young Children and Chimpanzees. *Child Development, 77*(3), 640–663. https://doi.org/10.1111/j.1467-8624.2006.00895.x

Wellman, H. M. (2014). *Making minds: How theory of mind develops.* Oxford University Press.

Wu, S. A., Wang, R. E., Evans, J. A., Tenenbaum, J. B., Parkes, D. C., & Kleiman-Weiner, M. (2021). Too many cooks: Bayesian inference for coordinating multi-agent collaboration. *Topics in Cognitive Science*, *13*(2), 414-432.

Zajonc, R. B. (1965). Social facilitation. *Science, 149*(Whole No. 3681), 269–274. https://doi.org/10.1126/science.149.3681.269

Zhai, S., Cheng, S., Moskowitz, N., Shen, M., & Gao, T. (2024). The development of commitment: Attention for intention. *Child Development*. https://doi.org/10.1111/cdev.13955

Zhao, X., Zhao, X., Gweon, H., & Kushnir, T. (2021). Leaving a choice for others: Children's evaluations of considerate, socially-mindful actions. *Child Development*, 92(4), 1238–1253. https://doi.org/10.1111/cdev.13480