RATIONAL USE OF TIME FOR SOCIAL COORDINATION IN HUMANS AND AI

Anonymous authors
Paper under double-blind review

ABSTRACT

All social interactions are situated in *time*. Previous studies have shown that humans are able to use implicit temporal cues, such as the duration of a pause, to infer others' mental states. But how do humans use temporal cues for dynamic, fine-grained social coordination? And can AI systems do the same, without "experiencing" continuous real-time as humans do? We study these questions using a cooperative card game task called *The Mind*, which requires players to achieve a shared goal across repeated interactions but forbids explicit communication: the only available signal is the passage of time. We evaluate 150 humans, and two classes of AI models which reflect different ways of operationalizing time in artificial systems: computer-use agents (CUAs), and standard text-prompted LLMs. We find that human behavior is best captured by a Bayesian strategy that infers a partner's belief from their timing—and similarly for frontier models. However, unlike humans, none of the tested models seem to efficiently adapt their behaviors to their partners' over the course of the game. From a cognitive perspective, our findings suggest that humans perform implicit probabilistic reasoning about continuous-time signals in dynamic social settings, and that such reasoning mechanisms may emerge without real-time experience. Practically, our study highlights new challenges for AI systems to coordinate efficiently with other agents by learning to interpret and adapt to their implicit signals in dynamic settings.

1 INTRODUCTION

Humans draw rich social inferences through all kinds of signals, from linguistic expressions to gestures to the placement of objects. But humans also infer complex information about other agents' mental states in the *absence* of explicit communicative signal: that is, based on **the passage of time**. For example, the duration of silence before an utterance in a conversation can provide cues into a speaker's trustworthiness or cognitive state (e.g., Christenfeld, 1995; Kohtz & Niebuhr, 2017; Ziano & Wang, 2021), and solving a complex task in a short amount of time might suggest that the agent had prior knowledge about the task (e.g., Berke et al., 2026). These findings provide evidence that humans use temporal information to draw one-shot social inferences, but we still lack a systematic account of how humans use time as a social signal to dynamically coordinate with other participants.

It also remains unclear whether AI systems can similarly use temporal information for social coordination. As AI agents are increasingly used in real-time interactions (e.g., Nakano et al., 2021; Li et al., 2026), it seems desirable for AI agents to leverage time as an information source about human users or other agents. However, while large language models (LLMs) have been evaluated in diverse social tasks (Tran et al., 2025; Sidji et al., 2024; Zhu et al., 2025; Yim et al., 2024), most studies have investigated communication based on explicit signals, such as words or non-linguistic symbols. And studies of temporal reasoning typically represent time using explicit timestamps or durations (Fatemi et al., 2025; Gupta et al., 2023; Wang & Zhao, 2024), leaving underexplored how AI systems perform in settings where time unfolds continuously and must be inferred through interaction.

Moreover, investigating time-based social coordination in AI models is not just motivated on practical grounds. It is not even clear whether current AI models are the right *kind of thing* to be able to use time like humans do. LLMs can only iteratively receive discrete prompts, instead of existing in the continuous flow of real-time like biological organisms. Even Computer-Use Agents (CUAs), which can dynamically interact with graphical user interfaces, still perceive the environment through

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

discrete screenshots rather than continuous visual input. It is therefore worth investigating whether such AI systems can perform complex collaborative tasks that require time-based inferences.

Building upon prior work using games to investigate cognition in humans (Newell & Simon, 1972; Dubey et al., 2018; Allen et al., 2024) and AI models (Silver et al., 2016; Vinyals et al., 2019; Bakhtin et al., 2022; Liang et al., 2025), we evaluate humans and models on a cooperative task based on a card game called *The Mind*. In our simplified version of the game, there are two players. On level n, each player is randomly dealt n cards from a deck of cards numbered 1 through 100. The players share a cooperative goal: play all dealt cards in ascending order without any communication (no talking, gesturing, winking, etc.). The only information players have is the passage of time. This form of coordination lacks any kind of explicit signaling or grounding of symbols into meanings—and yet, people successfully play the game, with minimal training. They even find it fun.[1]

Using *The Mind*, we investigate two main questions: How do humans use temporal cues for dynamic social coordination? And can AI systems also coordinate using such cues, despite not "experiencing" continuous real time? We break these questions down into the following subquestions:

- **RQ1 (Performance):** How successfully and efficiently do humans and models perform?
- **RQ2 (Strategy):** What cognitive strategy best characterizes human and model behavior?
- **RQ3 (Adaptation):** Do humans and models adapt their behaviors to increase success and efficiency with their partner across repeated interactions?
- **RQ4 (Behavior Consistency):** Do models show consistent behavior when prompted to reason about time at different granularities?

We implement five types of cognitive strategies as fixed partners, and evaluate how three types of players interact with them: (1) humans; (2) CUAs, which operate through a real-time graphical user interface; and (3) standard LLMs, which can only reason about time through prompts. We find that both humans' and frontier LLMs' behaviors are best characterized by a Bayesian strategy that reasons about time to infer the partner's state (Section 4.2). However, although frontier LLMs can match or even outperform humans in task success (Section 4.1), humans show greater adaptation to their partners' behavior over the course of the game (Section 4.3). LLMs' performance is also sensitive to how temporal information is represented in the prompt (Section 4.4).

Our results suggest that humans perform implicit probabilistic reasoning about continuous-time signals in dynamic social settings. Frontier LLMs may have internalized the same rational inferential strategy that humans use to coordinate under temporal uncertainty. However, this alignment does not appear to be driven by the same experience-driven, online adaptation to a specific partner that characterizes human coordination. LLMs may already outperform humans in aggregate coordination outcomes, but efficiently and consistently adapting to a partner during an interaction remains an open challenge—especially when using implicit temporal signals without linguistic grounding.

2 Related Work

Humans can rapidly establish shared conventions with a stranger, even when they have no prior coordination history (e.g., Clark & Wilkes-Gibbs, 1986; Garrod & Doherty, 1994). Importantly, convention formation does not rely exclusively on linguistic communication. Humans can coordinate through discrete linguistic cues (Steels et al., 2005), discrete non-linguistic cues (Misyak et al., 2016), continuous non-linguistic cues (Chen et al., 2025), and body gestures (Pezzulo et al., 2019). A growing body of work has also investigated convention formation in artificial agents from different perspectives, including coordination games (Sidji et al., 2024; Shih et al., 2021) and pragmatic reasoning (Yu et al., 2026). Various benchmarks have also evaluated current agent systems' collaboration ability in complex tasks (Zhu et al., 2025; Shih et al., 2021; Yim et al., 2024; Yee & Koh, 2026). Other studies have specifically examined LLM collaboration in games under implicit or communication-constrained settings (Liang et al., 2025; Buscemi et al., 2026).

[1] To quote Paul Dean, the co-founder of board game site and show *Shut Up and Sit Down* (Dean, 2018): "Somehow, there are few things in life more satisfying than laying down cards in the right order, something that gets the blood pumping through your veins in a way nothing else quite can, something that gives you the feeling of being a true hero, a feeling that you know few others will ever really experience in their lives. *The Mind* is one of the purest, simplest and best games I have ever played."

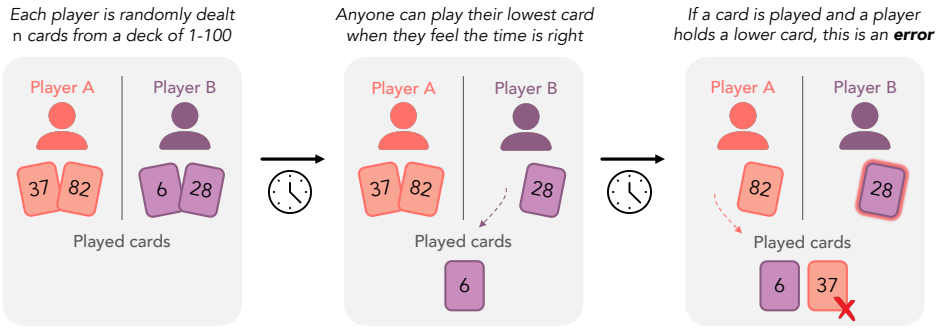


Figure 1: Overview of task requiring time-based inference (based on *The Mind*). Each player holds their own private cards (dealt uniformly at random from 1-100), and can see the previously played pile of cards. The players jointly share the goal of playing all dealt cards in ascending order, from lowest to highest, without any communication. Players simply play when they feel the time is right.

Another relevant line of work has investigated temporal inference in humans and LLMs. Humans infer other people's beliefs or preferences from response or decision time (Gates et al., 2021; Bavard et al., 2024; Konovalov & Krajbich, 2023). Recently, Berke et al. (2026) showed that people can infer how long another agent spent distracted purely from the duration of a pause, and that this inference is well captured by a Bayesian model of computation. Temporal reasoning has also received increasing attention in LLMs, with a growing number of benchmarks designed to assess models' ability to reason about temporal information (e.g., Fatemi et al., 2025; Islakoglu & Kalo, 2025b; Gupta et al., 2023; Wang & Zhao, 2024; Chu et al., 2024; Fatemi et al., 2025; Islakoglu & Kalo, 2025a; Wei et al., 2026). More recently, Sehgal et al. (2026) show that LLMs exhibit fragile adaptation to wall-clock time pressure in negotiating settings, while Wang et al. (2025) find that LLMs can associate token counts with the passage of continuous time. However, these evaluations primarily evaluate whether LLMs can reason about explicitly specified and static temporal information, leaving how LLMs represent and track the passage of time during dynamic interaction underexplored.

3 Methods

3.1 Overview of Experiments

Our task is based on a cooperative card game called *The Mind* (see Figure 1). There are two players in the game, and a deck of cards numbered 1 through 100. The game proceeds across ten levels. On level 1, each player is randomly dealt 1 card from the deck; on level 2, each player is randomly dealt 2 cards from the new deck; and so on. On each level, the players share a cooperative goal: play all dealt cards in ascending order (lowest to highest). Critically, there are no turns, and players cannot communicate with each other. Players simply play a card when they feel that it is the next lowest. If a player plays a card that is higher than the lowest card held by another player, this counts as an error, and all the lower card(s) are automatically played.

We evaluate three types of players on the task: humans, CUAs, and LLMs. In each experiment, we pair the agent of interest against artificial partners that implement different kinds of cognitive strategies. By manipulating the partner strategies, we can perform controlled analyses on what kind of partners leads to the best cooperation with each agent. We describe the rule-based partner strategies in more detail in Section 3.2, and then describe the methods of each experiment in Section 3.3.

3.2 Rule-based Partners

We pair each agent (humans, CUAs, and LLMs) with five kinds of partners implementing different cognitive strategies: Linear +/-Calibration, Bayesian +/-Calibration, and Random. These strategies differ along two separate dimensions: first, the way that time is used (Linear, Bayesian, or Random), and second, whether play speed is adjusted based on their partner's hesitation time (+/-Calibration). The Linear strategy is motivated by the evidence that adults represent numbers linearly (Siegler & Opfer, 2003), making a linear mapping between the card value and waiting time a plausible simple

default. The Bayesian strategy is motivated by prior work that show human temporal reasoning or perception are well described by Bayesian models (e.g., Berke et al., 2026; Jazayeri & Shadlen, 2010). The Random strategy is taken as a baseline. Additional details are provided in Appendix A.

At the start of level n, player A is dealt a hand of n cards $\langle a_1, a_2, \cdots, a_n \rangle$, where $a_1 < a_2 < \cdots < a_n$, and player B similarly is dealt $\langle b_1, b_2, \cdots, b_n \rangle$. Cards are played one at a time. Let c_i denote the i-th card played, and use t to denote elapsed wall-clock time since the most recent card was played. We write $c^{(t)}$ to denote the most recently played card at time t, and $\mathbf{c}^{(t)} = \{c_1, c_2, \ldots, c^{(t)}\}$ to denote the history of cards played until time t. We also write $a_1^{(t)}$ to denote player A's lowest remaining card, and $b_1^{(t)}$ to denote player B's lowest remaining card at time t. The t "clock" is reset to 0 at the start of each round and each time a card is played.

3.2.1 Strategy types

Linear player. The linear player waits a fixed amount of time proportional to the gap between the most recently played card $c^{(t)}$ and their lowest card. In other words, they use a fixed scaling factor—the "tempo" τ—to determine when to play their lowest card. For a given tempo τ, player A waits to play their lowest card $a_1^{(t)}$ at elapsed time $\tau \cdot \left(a_1^{(t)} - c^{(t)}\right)$ (and similarly for player B).

If both players share the same tempo τ, and follow this strategy perfectly and deterministically, then this strategy is guaranteed to succeed. However, this strategy is sensitive to miscalibration of τ, potentially leading to mistakes. Separately, this strategy ignores the *information* in the other player's silence, potentially leading to inefficiency (i.e., wait times that are longer than necessary).

Bayesian player. Instead of using a fixed mapping between card values and time units, the Bayesian player rationally uses their partner's silence as *evidence*. Without loss of generality, we focus on player A's perspective. Let $B_1^{(t)}$ be a random variable denoting player B's lowest card at time t, and let T_B be a random variable denoting the time at which player B plays their lowest card. This means that at time t, player A has observed that $T_B > t$. Player A can use Bayesian inference to compute the probability that player B's current lowest card $B_1^{(t)}$ is equal to some value k:

$$P\left(B_1^{(t)} = k \mid T_B > t, \mathbf{c}^{(t)}\right) \propto P\left(T_B > t \mid B_1^{(t)} = k, \mathbf{c}^{(t)}\right) \cdot P\left(B_1^{(t)} = k \mid \mathbf{c}^{(t)}\right) \tag{1}$$

To compute the prior probability that player B's lowest card $B_1^{(t)}$ is equal to some card value k, player A assumes a uniform draw from all cards that are not in their hand and have not been played. To compute the likelihood of player B waiting at least until time t to play card with value k, we model each player's play time as a random variable drawn from a hazard process that depends on their card value. That is, $\left(T_B \mid B = k, c^{(t)}\right) \sim \text{Weibull}\left(\gamma, \frac{1}{\tau \cdot (k - c^{(t)})}\right)$, where τ is the tempo parameter and γ is a parameter controlling the change of the hazard rate. If $\gamma > 1$, the hazard rate increases over time, which captures impatience in the game. A full derivation is provided in Appendix A.3.

Player A waits to play until their belief that they hold the lower card exceeds a confidence threshold $\theta > 0$. In other words, they play at the smallest t such that $P(B_1^{(t)} > a_1^{(t)} \mid T_B > t) \geq \theta$.

Random player. As a baseline, we consider a random player who decides whether to play or wait every second with probability 0.5.

3.2.2 Tempo calibration

The Linear and Bayesian players both make use of a "tempo" parameter $\tau > 0$. Intuitively, the tempo of the game serves as a scaling factor between card value units and time units. A smaller τ means that everyone plays more quickly; a larger τ means that everyone plays more slowly. A **-Calibration** player begins with a fixed tempo τ_0 and does not update it throughout the course of the game. A **+Calibration** player begins with some τ_0 and updates it after each played card. After each card c_i is played at time t_i since the previous play, the players observe a new data point and can immediately update their estimate of τ. The single-observation MLE implied by this data point is $\hat{\tau}_i = \frac{t_i}{c_i - c_{i-1}}$, where c_{i-1} is the previously played card. For players with tempo calibration, we

define the players as taking an exponential moving average, weighted toward recent observations: $\hat{\tau}_i = (1-\eta)\cdot\hat{\tau}_{i-1}+\eta\cdot\frac{t_i}{c_i-c_{i-1}}$, where $\eta \in (0,1)$ controls the rate of adaptation to recent observations.

3.3 Experimental Methods

3.3.1 Experiment 1: Humans vs. Rule-based partners

We randomly generated 10 game seeds, each defining 10 levels with 2 pairs of hands per level. Each seed was evaluated under 5 different partner conditions: Bayesian +/-Calibration, Linear +/-Calibration, and Random, resulting in 50 seed-by-partner combinations. For all Bayesian and Linear players, we set τ to 0.5, γ to 2, and θ to 0.85. For players with calibration enabled, we set η to 0.3. These parameter values were selected based on the results and feedback from the pilot study.

For each seed-by-partner combination, we recruited 3 human participants through Prolific, resulting in 150 total participants. Participants played the game with the rule-based partners through a web-based game interface. Before playing, participants provided informed consent and received detailed instructions on the game rules and the layout of the game interface. They were then required to answer three comprehension questions based on the instructions before proceeding to the game. To encourage efficient cooperation, we provided a bonus up to $2 based on the game duration and success rate apart from the base payment. Experiment details can be found in Appendix C.

3.3.2 Experiment 2: Computer-use agents (CUAs) vs. Rule-based partners

CUAs are built on large vision-language models. They take multimodal content as input and can use computer tools to interact with and operate GUIs in real-time, making them a natural choice for investigating whether AI agents can use temporal information during social coordination. CUAs interact with the game under the same experimental conditions and through the same interface used by human participants, including instructions and comprehension questions. Therefore, their interaction time is measured using wall-clock time, as it is for human participants.

We evaluate three frontier CUAs: `Gemini-3.7-flash` (Gemini Team et al., 2023), `GPT-5.6-luna`, and `GPT-6-astra`.[2] Due to computational costs, we run each CUA only once for each game seed-partner combination, resulting in 50 (10 seeds$\times$5 partners) data points per CUA.

3.3.3 Experiment 3: LLMs vs. Rule-based partners

Finally, we consider standard LLMs which can only "use" time by reasoning about it in text-based prompts. We evaluate 10 instruct-tuned LLMs from 5 families, spanning across sizes and open-/closed-source settings: `GPT-6-astra` and `GPT-5.6-luna`; `Gemini-3.7-flash`; `GPT-oss-120b` and `GPT-oss-20b` (OpenAI, 2025); `Qwen3-32B` and `Qwen3-8B` (Qwen Team, 2025); and `LLaMa-3.1-70b`, `LLaMa-3.1-8b`, and `LLaMa-3.1-1b` (Touvron et al., 2023). The experiment uses the same conditions as the CUA experiment, with 50 data points per model. We also pair LLMs with the same LLM model to see how well they cooperate with each other. For LLMs that show performance comparable to humans, we additionally pair them with `Gemini-3.7-flash` to examine how they coordinate with a higher- or similarly performing partner. Conversely, we pair `Gemini-3.7-flash` with `gpt-oss-120b` to examine how it coordinates with a lower-performing partner. Hyperparameter details can be found in Table 1.

As LLMs cannot directly perceive the passage of time, we prompt the LLM at simulated time steps with the current game state and ask the model whether it decides to wait (`WAIT`) or play the lowest card (`PLAY`). To assess whether models behave consistently with different temporal granularities, we consider the following conditions: (i) **Time-based**: The elapsed waiting time is explicitly provided. We use four variants: the time is updated in increments of 1 second, 0.5 seconds, or 0.25 seconds, or the model directly generates an absolute waiting-time estimate based on the current game state. (ii) **Wait-based (count)**: The number of previous `WAIT` decisions is provided as a numerical value. (iii) **Wait-based (action)**: The history of previous decisions is provided as a sequence of tokens (e.g., `WAIT, WAIT, WAIT`). The Wait-based conditions test whether LLMs can perform the task

[2] We refer to the CUA models as *[model]-CUA* to distinguish them from their underlying language models.

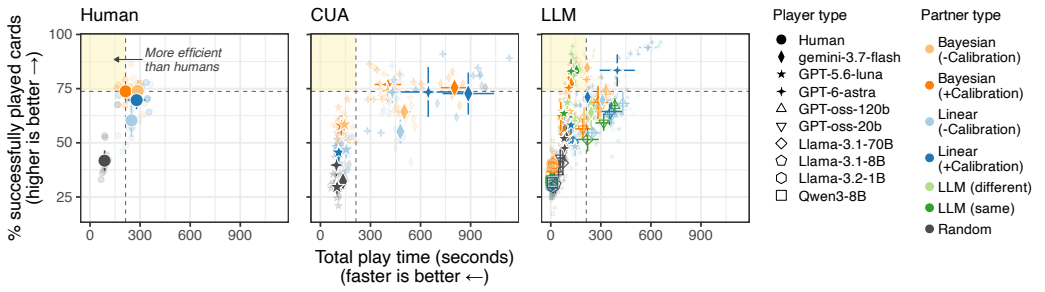


Figure 2: Success rate versus total play time, for humans, CUAs, and LLMs with error bars showing 95% CI. Total play time is directly measured in wall-clock time for humans and CUAs, and indirectly computed based on prompt responses from LLMs. Opaque large markers represent means across game seeds; more transparent smaller markers represent individual game seeds.

by relying on numerical information, even when it is divorced from real time units.[3] We evaluate all LLMs under these conditions except `GPT-6-astra`, for which we use only the 1-second Time-based condition due to costs. Details of prompts can be found in Section D.2 and Table 2.

4 Results

4.1 RQ1: Overall Performance of Humans, CUAs, and LLMs

We first analyze overall performance, using two complementary measures: **game duration**, defined as the time elapsed between the start and end of the game, and **success rate**, defined as the number of successfully played cards divided by the total number of cards (110). An ideal player would successfully play as many cards as possible while completing the game in as little time as possible.

The results are presented in Figure 2.[4] We observe three noteworthy findings. First, across all player types—humans, CUAs, and LLMs—the best performance (i.e., trade-off between success rate and game duration) is generally achieved with Bayesian partners. This may suggest that people's (and models') default expectations about partner behavior are more consistent with a Bayesian strategy than alternatives like the Linear strategy. Second, frontier models can achieve performance comparable to, and in some cases exceeding, that of humans. The rightmost subplot shows that certain LLM players achieve both higher success and faster play times than the most efficient human games.[5] Finally, although CUAs can achieve human-level success rates, they take much longer to complete the game than humans. This suggests that, under the same experimental conditions, CUAs do not yet use temporal information as efficiently as humans for social coordination.

4.2 RQ2: Cognitive Strategies Used by Humans, CUAs, and LLMs

Next, we ask which strategies best capture human, CUA, and LLM behavior. We evaluated the fit of six types of probabilistic models to each game record: Random, Constant Wait,[6] Linear +/- Calibration, and Bayesian +/-Calibration. We evaluated a wide range of parameters for each model class, and used each model M to compute the likelihood of the critical player's (i.e., human, CUA, or LLM) sequence of actions. When the critical player played first, we calculated the probability

[3] In order to "play" with LLMs in the Wait-based conditions, the rule-based partners estimate the elapsed waiting time from the number of `WAIT`s, assuming that each `WAIT` token corresponds to one second.

[4] For LLMs, we report results from the time-based (1s) prompt. We compare across prompts in Section 4.4.

[5] To jointly evaluate success rate and game duration, we compute an "efficiency" score by z-scoring time and success and computing $-0.5 \times time_z + 0.5 \times success_z$. We visualize top-20 most efficient player-partner combinations in Figure 11. Closed-source LLMs are among the most efficient player types, showing higher efficiency than human players in these combinations.

[6] The Constant Wait model assumes that Player A's waiting time is drawn from the same fitted distribution regardless of the cards played or the preceding game history. This strategy was not included as a rule-based partner condition, but we included it as a potential strategy explaining player behaviors.

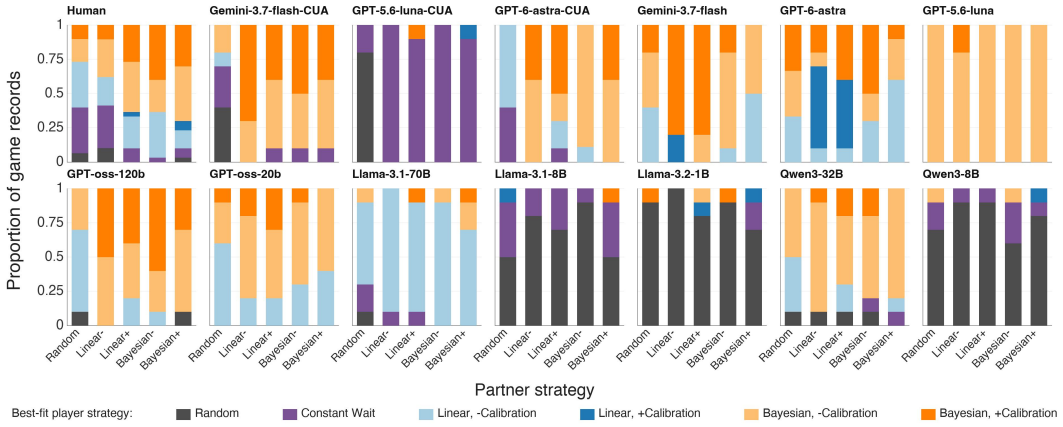


Figure 3: Strategies that best capture human, CUA, and LLM behaviors across partner conditions.

under M that the player would play within the observed time interval. When the rule-based partner played first, we calculated the probability under M that the critical player would still be waiting at that time. Forced autoplay and stand-alone cards were excluded because they did not reflect an active timing decision. The complete likelihood definitions are given in Appendix B.

Using the maximized likelihood for each model and game record, we computed $\text{AIC} = 2K - 2\log\hat{L}$, where K is the number of fitted parameters and $\hat{L}$ is the maximized likelihood.[7] We selected the model with the lowest AIC for each game record, and then report the proportion of game records won by each model class. AIC values were never compared across different game records.

The results are shown in Figure 3. First, we find that humans adopt strategies more flexibly when playing with different partners. The Constant strategy best fit humans paired with simpler partners like Random or Linear players without calibration, while Bayesian strategies best fit humans paired with other more complex partners. In contrast, both LLMs and CUAs tend to be best fit by one strategy across player types. Second, the best fitting strategy for LLM and CUA behavior varies substantially across model families and sizes. Bayesian strategies best fit the behavior for larger models, while the Random strategy best fit the behavior for smaller ones. `Llama 70B` and `GPT-5.6-luna (CUA)`, however, are best fit with Linear and Constant strategies, respectively.

4.3 RQ3: Behavior Adaptation in Humans, CUAs, and LLMs

Next, we ask whether players adapt to their partners to improve efficiency across the levels of a game. Mirroring the analysis in Section 4.1, we use two behavioral measures to capture changes of efficiency: normalized hesitation time and success rate.[8] Because cards that are automatically played following an error by the other partner and standalone cards do not meaningfully reflect the player's own timing decision, we exclude these cards from the calculation of hesitation time.

To examine how these measures change over the course of the game, we divide the 110 cards into 10 deciles and model the effect of decile on success rate and normalized hesitation time. For success rate, we fit generalized linear mixed-effects models (GLMMs).[9] A decrease in hesitation time suggests that players become more confident in their partner's strategy, while an increase in success rate suggests that the coordination becomes more successful over time. Neither measure alone, however,

[7] We also considered BIC, which further penalizes model complexity and would therefore disfavor the Bayesian strategies. BIC selected the same model class as AIC for 80.3% of game records. See Appendix E.1.

[8] Raw hesitation time is not comparable across plays because it depends on the gap between the current card and the previously played card. For example, for the same current card (e.g., 55), a player would be expected to wait for different amounts of time when the previously played card is 12 versus 42. We therefore normalize hesitation time by the value gap between the current and previously played cards.

[9] The model formulas were `outcome ~ decile + (1|hand_id) + (1|group)`, where `group = participant` for humans and `prompt_type` for LLMs and CUAs; `outcome` was `cbind(success, total-success)` or `normalized_hesitation_time`, respectively.

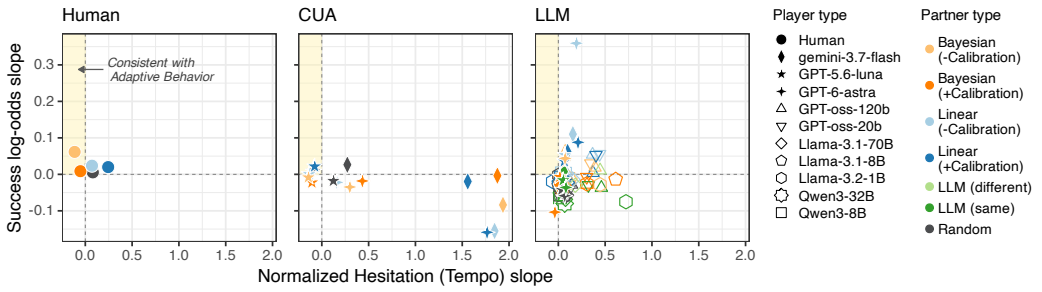


Figure 4: Regression slopes of success rate and normalized hesitation time.

provides a complete picture of behavior adaptation: players may become more cautious over time and achieve higher accuracy, or they may simply become less patient and act more quickly.

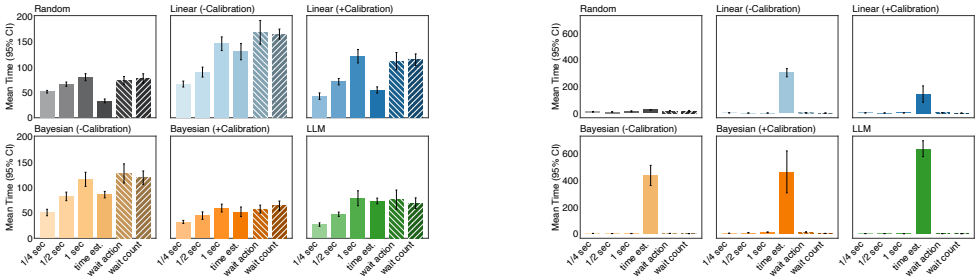
We plot the slopes of success rate and normalized hesitation time in Figure 4. We find that human players show better learning results across all partner conditions with two conditions showing statistically significant or marginally significant improvements in performance: LINEAR (-CALIBRATION) condition ($\hat{\beta} = 0.024$, $p = 0.07$) and BAYESIAN (-CALIBRATION) condition ($\hat{\beta} = 0.062$, $p < 0.01$). Humans also tend to respond more quickly as the game progresses when playing with Bayesian partners, with negative estimates in both the BAYESIAN (-CALIBRATION) ($\hat{\beta} = -0.11$, $p = 0.08$) and BAYESIAN (+CALIBRATION) ($\hat{\beta} = -0.05$, $p = 0.35$) conditions.

In contrast, across all CUA and LLM players, only `GPT-5.6-luna (CUA)` show both faster speed and higher success rate over the course of the game.[10] However, neither effect is statistically significant ($\hat{\beta}_{\text{tempo}} = -0.08$, $p = 0.64$; $\hat{\beta}_{\text{success}} = 0.02$, $p = 0.30$). Most other player-partner combinations instead exhibit a trade-off between success rate and speed: models either become more successful but slower over the course of the game, or become faster at the cost of a lower success rate. Human participants show patterns consistent with convention formation in two of their five partner conditions, whereas no individual model shows such a pattern in more than one of its five partner conditions. Additionally, human players are the **only** group that consistently shows increasing success rates over the course of the game across 5 partner types. This pattern suggests that humans adapt to their partners' behavior more consistently across conditions than individual LLMs or CUAs. Details of the statistical results are in Section E.4.

4.4 RQ4: BEHAVIOR CONSISTENCY ACROSS PROMPTS (LLMS ONLY)

Finally, we evaluated whether the text-prompted LLMs show consistent behavior across different ways of encoding temporal information in the prompt. We first analyze the four Time-based prompts (see Section 3.3.3 for details). An agent with a consistent interpretation of time should exhibit similar hesitation times when the temporal granularity or temporal information changes (and all other aspects of the prompt remain identical). We use the coefficient of variation (CV) to measure the variation of each player–partner combination across the four Time-based prompt variants. Figure 5 shows the hesitation times of the two models with the smallest and largest average CVs (see Appendix E.3 for additional models and CV results). We find that simply changing the temporal granularity used to represent waiting time in the prompt can lead to substantially different behaviors, even when all other aspects of the game remain unchanged. In particular, it seems that models may make decisions based on how many times they are prompted to wait, rather than *how long* they've been waiting. This effect is observed consistently across LLMs, but its magnitude varies systematically with model size: larger models tend to be more robust to prompt variations. Turning to the Wait-based prompts, we also observe in some cases (such as with `GPT-5.6-luna`; Figure 5a) that the Wait-based prompts lead to similar playing time as certain Time-based prompts. Since the Wait-based prompts do not make any reference to time units, this finding suggests that LLMs might be able to perform the task based on numerical information instead of reasoning about real-time.

[10] We do not interpret the `GPT-6 (CUA)`-Random combination as evidence of convention formation, since Random partners have no temporal structure for a player to adapt to.



(a) GPT-5.6-Luna

(b) Llama-3.2-1B-Instruct

Figure 5: Mean play time by partner type and prompt type (x-axis) for the LLMs with the smallest (a) and largest (b) average coefficient of variation across prompt types. Error bars show 95% CI.

5 Discussion

Social interactions in the real world unfold in *time*. Whether and how different agents use time as a signal is therefore central to understanding social coordination for both humans and AI. In this work, we investigated how humans and AI systems reason about the passage of time, extending prior work on one-shot inferences to a collaborative task involving repeated interactions. We found that humans perform the best with partners that perform Bayesian reasoning about time, and humans themselves also seem to adopt Bayesian reasoning strategies instead of simply mapping between time units and card units. Moreover, humans show evidence of adaptation over the course of the game when playing with Bayesian partners. Taken together, these results suggest that humans make rational use of temporal signals in dynamic interactions, and expect social partners to do the same.

Turning to the AI models, we found that some frontier models can achieve human-level performance and, like humans, show behaviors that are best captured by Bayesian reasoning strategies. However, unlike humans, the models do not seem to adapt to their partner's behavior as the game progresses. Some actually show worse success rates over the course of a game. The text-based LLMs are also sensitive to time granularity presented in the prompts, and show similar timing patterns whether or not time units are mentioned in the prompt. These findings suggest that while AI models might adopt human-like strategies for temporal coordination, the process by which they arrive at this behavior may differ fundamentally from how humans learn to coordinate with a partner. We also find that only prompt-based LLMs achieve human-comparable efficiency in the task, even though CUAs act in real-time like humans do. This pattern cannot be explained simply by implementation or reasoning latency, since CUAs do show hesitation times comparable to humans when playing against certain kinds of partners. This observation may instead reflect a difficulty current CUAs face in balancing speed and accuracy as an interaction unfolds, trading longer waiting times for higher accuracy.

We note a few limitations regarding our study. First, we focused on players interacting with rule-based partners, rather than directly pairing "live" players with each other. While this design gave us greater control for addressing our research questions, it does not tell us what kinds of strategies would organically arise through player-player interactions. Second, the strategies we use to model the players' behavior are necessarily simplified. Players may mix strategies, change strategies across time, and so on. Additionally, the Bayesian strategy we used in our experiment represents a rational approach to inferring the state of the game, but a deeper social strategy might involve recursive reasoning about the other agent's mental state. We leave a detailed treatment of these more complex strategies for future work. Third, we note a key difference in the way the task is realized for humans and text-prompted LLMs: humans experience continuous wall-clock time, whereas for LLMs we approximate continuous-time interaction by explicitly specifying waiting durations in prompts. While this distinction is motivated by our broader question—how a non-time-embodied system can perform the task—we note that this complicates direct comparisons between human and LLM playing times. A more direct comparison would test humans in a text-based version of the task. Finally, we acknowledge that our experiments use a relatively constrained task setting based on a card game. Future work can examine how temporal signals interact with other forms of communication or environmental cues in more naturalistic settings. We see our study as suggesting new ways to investigate the flexibility and efficiency of social cognition in humans and artificial systems.

AI USE STATEMENT

In this work, we used generative AI tools for *design or provide feedback on research methodology or experiments, interpret results*. We have not used generative AI tools for *implement methods, clean and reformat dataset, support qualitative and thematic data analysis, propose or refine hypotheses*, and all other tasks are not applicable to this work. Additionally, we used generative AI tools for *create or modify scientific figures or images, sourcing/searching for information, edit a research paper to improve readability, identify relevant literature*. We have reviewed all AI-assisted work. The implementation python code is written by authors and then reviewed by AI tools. The JavaScript implementation of the game interface was translated from the Python code with the assistance of AI tools and was tested by the authors before the experiments. We also used AI tools to search for information and identify relevant literature; all such information was independently verified against the original sources. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This study involved human participants and was reviewed and approved by an institutional review board (identifying details withheld for double-blind review). All participants provided informed consent before participating. Participants were compensated for their time in accordance with the approved study protocol ($12 per hour). The data were collected and analyzed under the approved protocol, and personally identifying information was not included in the released or analyzed dataset.

Our work focuses on how humans and AI coordinate with a rule-based partner by using implicit signals, i.e., passage of time, through repeated interactions. As AI systems increasingly serve as personal assistants and collaborative agents, the ability to infer information from behavioral signals may help build more efficient and personalized interactions. However, it is also worth noting that such ability may lead to unintended social inference and raise privacy concerns, particularly during interactions in which users may not be fully aware of such subtle cues.

REPRODUCIBILITY STATEMENT

The hyperparameters used by the models are listed in Table 1. Code will be available at acceptance.

REFERENCES

Kelsey Allen, Franziska Brändle, Matthew Botvinick, Judith E. Fan, Samuel J. Gershman, Alison Gopnik, Thomas L. Griffiths, Joshua K. Hartshorne, Tobias U. Hauser, Mark K. Ho, Joshua R. de Leeuw, Wei Ji Ma, Kou Murayama, Jonathan D. Nelson, Bas van Opheusden, Thomas Pouncy, Janet Rafner, Iyad Rahwan, Robb B. Rutledge, Jacob Sherson, Özgür Şimşek, Hugo Spiers, Christopher Summerfield, Mirko Thalmann, Natalia Vélez, Andrew J. Watrous, Joshua B. Tenenbaum, and Eric Schulz. Using games to understand the mind. *Nature Human Behaviour*, 8(6):1035–1043, June 2024. ISSN 2397-3374. doi: 10.1038/s41562-024-01878-9. URL https://doi.org/10.1038/s41562-024-01878-9.

Anton Bakhtin, Noam Brown, Emily Dinan, Gabriele Farina, Colin Flaherty, Daniel Fried, Andrew Goff, Jonathan Gray, Hengyuan Hu, Athul Paul Jacob, Mojtaba Komeili, Karthik Konath, Minae Kwon, Adam Lerer, Mike Lewis, Alexander H. Miller, Sasha Mitts, Adithya Renduchintala, Stephen Roller, Dirk Rowe, Weiyan Shi, Joe Spisak, Alexander Wei, David Wu, Hugh Zhang, and Markus Zijlstra. Human-level play in the game of Diplomacy by combining language models with strategic reasoning. *Science*, 0(0):eade9097, 2022. doi: 10.1126/science.ade9097. URL https://doi.org/10.1126/science.ade9097.

Sophie Bavard, Erik Stuchlỳ, Arkady Konovalov, and Sebastian Gluth. Humans can infer social preferences from decision speed alone. *PLoS Biology*, 22(6):e3002686, 2024.

Marlene Berke, Ben Sterling, Abigail L Tenenbaum, and Julian Jara-Ettinger. People Reconstruct Others' Cognitive Processes by Representing Their Minds as Computational Systems, March 2026. URL osf.io/preprints/psyarxiv/gmwav_v1.

Alessio Buscemi, Daniele Proverbio, Alessandro Di Stefano, The-Anh Han, German Castignani, and Pietro Liò. When numbers start talking: Implicit numerical coordination among LLM-based agents. *arXiv preprint arXiv:2601.03846*, 2026.

Alicia M Chen, Matthias Hofer, Moshe Poliak, Roger Levy, and Noga Zaslavsky. Discrete and systematic communication in a continuous signal-meaning space. *Journal of Language Evolution*, 10(1):lzaf003, 2025.

Nicholas Christenfeld. Does it hurt to say um? *Journal of Nonverbal Behavior*, 19(3):171–186, September 1995. ISSN 1573-3653. doi: 10.1007/BF02175503. URL https://doi.org/10.1007/BF02175503.

Zheng Chu, Jingchang Chen, Qianglong Chen, Weijiang Yu, Haotian Wang, Ming Liu, and Bing Qin. Timebench: A comprehensive evaluation of temporal reasoning abilities in large language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1204–1228, 2024.

Herbert H Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. *Cognition*, 22(1): 1–39, 1986.

Paul Dean. Review: The Mind, 2018. URL https://www.shutupandsitdown.com/review-the-mind/.

Rachit Dubey, Pulkit Agrawal, Deepak Pathak, Thomas L. Griffiths, and Alexei A. Efros. Investigating Human Priors for Playing Video Games. In *Proceedings of the 35th International Conference on Machine Learning*, 2018. URL https://arxiv.org/abs/1802.10217.

Bahare Fatemi, Seyed Mehran Kazemi, Anton Tsitsulin, Karishma Malkan, Jinyeong Yim, John Palowitch, Sungyong Seo, Jonathan Halcrow, and Bryan Perozzi. Test of time: A benchmark for evaluating LLMs on temporal reasoning. In *International Conference on Learning Representations*, volume 2025, pp. 94426–94447, 2025.

Simon Garrod and Gwyneth Doherty. Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, 53(3):181–215, 1994.

Vael Gates, Frederick Callaway, Mark K Ho, and Thomas L Griffiths. A rational model of people's inferences about others' preferences based on response times. *Cognition*, 217:104885, 2021.

Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, Katie Millican, et al. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023.

Vivek Gupta, Pranshu Kandoi, Mahek Vora, Shuo Zhang, Yujie He, Ridho Reinanda, and Vivek Srikumar. TempTabQA: Temporal question answering for semi-structured tables. In *Proceedings of the 2023 conference on empirical methods in natural language processing*, pp. 2431–2453, 2023.

Duygu Sezen Islakoglu and Jan-Christoph Kalo. ChronoSense: Exploring temporal understanding in large language models with time intervals of events. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar (eds.), *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 590–602, Vienna, Austria, July 2025a. Association for Computational Linguistics. ISBN 979-8-89176-252-7. doi: 10.18653/v1/2025.acl-short.46. URL https://aclanthology.org/2025.acl-short.46/.

Duygu Sezen Islakoglu and Jan-Christoph Kalo. Chronosense: Exploring temporal understanding in large language models with time intervals of events. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pp. 590–602, 2025b.

Mehrdad Jazayeri and Michael N Shadlen. Temporal context calibrates interval timing. *Nature neuroscience*, 13(8):1020–1026, 2010.

Lea S. Kohtz and Oliver Niebuhr. How Long is Too Long? How Pause Features After Requests Affect the Perceived Willingness of Affirmative Answers. In *Interspeech 2017*, pp. 3792–3796, 2017. doi: 10.21437/Interspeech.2017-50.

Arkady Konovalov and Ian Krajbich. Decision times reveal private information in strategic settings: Evidence from bargaining experiments. *The Economic Journal*, 133(656):3007–3033, 2023.

Xiaopeng Li, Pengyue Jia, Derong Xu, Yi Wen, Yingyi Zhang, Wenlin Zhang, Wanyu Wang, Yichao Wang, Zhaocheng Du, Xiangyang Li, et al. A survey of personalization: From rag to agent. *ACM Transactions on Information Systems*, 44(4):1–39, 2026.

Fangzhou Liang, Tianshi Zheng, Chunkit Chan, Yauwai Yim, and Yangqiu Song. LLM-hanabi: Evaluating multi-agent gameplays with theory-of-mind and rationale inference in imperfect information collaboration game. *arXiv preprint arXiv:2510.04980*, 2025.

Jennifer Misyak, Takao Noguchi, and Nick Chater. Instantaneous conventions: The emergence of flexible communicative signals. *Psychological science*, 27(12):1550–1561, 2016.

Reiichiro Nakano, Jacob Hilton, Suchir Balaji, Jeff Wu, Long Ouyang, Christina Kim, Christopher Hesse, Shantanu Jain, Vineet Kosaraju, William Saunders, Xu Jiang, Karl Cobbe, Tyna Eloundou, Gretchen Krueger, Kevin Button, Matthew Knight, Benjamin Chess, and John Schulman. WebGPT: Browser-assisted question-answering with human feedback. *ArXiv*, abs/2112.09332, 2021. URL https://api.semanticscholar.org/CorpusID:245329531.

Allen Newell and Herbert A. Simon. *Human Problem Solving*. Prentice-Hall, Oxford, 1972.

OpenAI. gpt-oss-120b & gpt-oss-20b model card, 2025. URL https://arxiv.org/abs/2508.10925.

Giovanni Pezzulo, Francesco Donnarumma, Haris Dindo, Alessandro D'Ausilio, Ivana Konvalinka, and Cristiano Castelfranchi. The body talks: Sensorimotor communication and its brain and kinematic signatures. *Physics of life reviews*, 28:1–21, 2019.

Qwen Team. Qwen3 technical report, 2025. URL https://arxiv.org/abs/2505.09388.

Neil KR Sehgal, Sharath Chandra Guntuku, and Lyle Ungar. Real-time deadlines reveal temporal awareness failures in LLM strategic dialogues. *arXiv preprint arXiv:2601.13206*, 2026.

Andy Shih, Arjun Sawhney, Jovana Kondic, Stefano Ermon, and Dorsa Sadigh. On the critical role of conventions in adaptive human-AI collaboration. *arXiv preprint arXiv:2104.02871*, 2021.

Matthew Sidji, Wally Smith, and Melissa J Rogerson. Human-AI collaboration in cooperative games: A study of playing codenames with an LLM assistant. *Proceedings of the ACM on Human-Computer Interaction*, 8(CHI PLAY):1–25, 2024.

Robert S Siegler and John E Opfer. The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological science*, 14(3):237–250, 2003.

David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, January 2016. ISSN 1476-4687. doi: 10.1038/nature16961. URL https://doi.org/10.1038/nature16961.

Luc Steels, Tony Belpaeme, et al. Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and brain sciences*, 28(4):469–488, 2005.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aur'elien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. Llama: Open and efficient foundation language models. *ArXiv*, abs/2302.13971, 2023. URL https://api.semanticscholar.org/CorpusID:257219404.

Khanh-Tung Tran, Dung Dao, Minh-Duong Nguyen, Quoc-Viet Pham, Barry O'Sullivan, and Hoang D Nguyen. Multi-agent collaboration mechanisms: A survey of LLMs. *arXiv preprint arXiv:2501.06322*, 2025.

Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, Laurent Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander S. Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, November 2019. ISSN 1476-4687. doi: 10.1038/s41586-019-1724-z. URL https://doi.org/10.1038/s41586-019-1724-z.

Minghan Wang, Ye Bai, Thuy-Trang Vu, Ehsan Shareghi, and Gholamreza Haffari. Discrete minds in a continuous world: Do language models know time passes? In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng (eds.), *Findings of the Association for Computational Linguistics: EMNLP 2025*, pp. 18703–18729, Suzhou, China, November 2025. Association for Computational Linguistics. ISBN 979-8-89176-335-7. doi: 10.18653/v1/2025.findings-emnlp.1016. URL https://aclanthology.org/2025.findings-emnlp.1016/.

Yuqing Wang and Yun Zhao. Tram: Benchmarking temporal reasoning for large language models. In *Findings of the Association for Computational Linguistics: ACL 2024*, pp. 6389–6415, 2024.

Shaohang Wei, Wei Li, Feifan Song, Wen Luo, Tianyi Zhuang, Haochen Tan, Zhijiang Guo, and Houfeng Wang. Time: A multi-level benchmark for temporal reasoning of LLMs in real-world scenarios. *Advances in Neural Information Processing Systems*, 38, 2026.

Brandon Yee and Pairie Koh. Benchmarking emergent coordination in large-scale llm populations: An evaluation framework on the moltbook archive. *arXiv preprint arXiv:2603.03555*, 2026.

Yauwai Yim, Chunkit Chan, Tianyu Shi, Zheye Deng, Wei Fan, Tianshi Zheng, and Yangqiu Song. Evaluating and enhancing LLMs agent based on theory of mind in guandan: A multi-player cooperative game under imperfect information. In *2024 IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pp. 461–465. IEEE, 2024.

Kefan Yu, Qingcheng Zeng, Weihao Xuan, Wanxin Li, Jingyi Wu, and Rob Voigt. The pragmatic mind of machines: Tracing the emergence of pragmatic competence in large language models. In Vera Demberg, Kentaro Inui, and Lluís Marquez (eds.), *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 192–213, Rabat, Morocco, March 2026. Association for Computational Linguistics. ISBN 979-8-89176-380-7. doi: 10.18653/v1/2026.eacl-long.9. URL https://aclanthology.org/2026.eacl-long.9/.

Kunlun Zhu, Hongyi Du, Zhaochen Hong, Xiaocheng Yang, Shuyi Guo, Zhe Wang, Zhenhailong Wang, Cheng Qian, Xiangru Tang, Heng Ji, and Jiaxuan You. MultiAgentBench : Evaluating the collaboration and competition of LLM agents. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar (eds.), *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 8580–8622, Vienna, Austria, July 2025. Association for Computational Linguistics. ISBN 979-8-89176-251-0. doi: 10.18653/v1/2025.acl-long.421. URL https://aclanthology.org/2025.acl-long.421/.

Ignazio Ziano and Deming Wang. Slow lies: Response delays promote perceptions of insincerity. *Journal of personality and social psychology*, 120(6):1457–1479, June 2021. ISSN 1939-1315 0022-3514. doi: 10.1037/pspa0000250.

A Details of rule-based players

A.1 State Space

Throughout this appendix we use the notation introduced in Section 3.2. There are two players, A and B. At the start of level n, each player is dealt a hand of n cards uniformly at random from a deck $\{1, 2, \ldots, 100\}$. We denote the players' hands as the following:

$$\text{Player A holds } \mathbf{A} = \langle a_1, a_2, \cdots, a_n \rangle, \text{ where } a_1 < a_2 < \cdots < a_n \tag{2}$$

$$\text{Player B holds } \mathbf{B} = \langle b_1, b_2, \cdots, b_n \rangle, \text{ where } b_1 < b_2 < \cdots < b_n \tag{3}$$

We will denote $a_1^{(t)}$ as player A's lowest remaining card, $b_1^{(t)}$ as player B's lowest remaining card, and $c^{(t)}$ the most recently played card at time t. When no card has yet been played in the current level, we set $c^{(t)} = 0$. There is no 0 card in the deck.

At any given time t, we can define the state as:

- $\mathbf{A}^{(t)} \subseteq \mathbf{A}$ denotes player A's remaining cards (known only to A)
- $\mathbf{B}^{(t)} \subseteq \mathbf{B}$ denotes player B's remaining cards (known only to B)
- $\mathbf{c}^{(t)} = \{c_1, c_2, \ldots, c^{(t)}\}$ is the history of played cards (known to both players)
- $\overline{\mathbf{c}}^{(t)} = \{1, \ldots, 100\} \setminus \mathbf{c}^{(t)}$ is the set of cards not yet played (known to both players, since all played cards $\mathbf{c}^{(t)}$ are known)

A.2 Linear player

As described in Section 3.2.1, the Linear player simply waits a fixed amount of time proportional to the gap between their lowest card and the previously played card before playing. We denote this scaling factor as with a "tempo" parameter $\tau > 0$.

Therefore, player A would play their lowest card at time

$$t^* = \tau \cdot \left(a_1^{(t)} - c^{(t)}\right), \tag{4}$$

and similarly for player B. When a single player holds the only card(s) remaining in the level (i.e., the standalone cards), they can execute the rest of the plays perfectly, so they play the remaining card(s) immediately rather than waiting the full time.

If both players share the same τ and follow this strategy perfectly and deterministically, this strategy will never make mistakes. However, as stated in Section 3.2.1, this strategy ignores the *information* in the other player's silence, and is sensitive to miscalibration of τ. For Linear players with tempo calibration (see Section 3.2.2), we replace τ by the updated estimate after each play.

A.3 Bayesian player

Consider the game state at time t. Let m_A denote the number of cards that player A still holds ($|\mathbf{A}^{(t)}|$), and let m_B denote the number of cards that player B still holds ($|\mathbf{B}^{(t)}|$). Both players know m_A and m_B at all times.

The prior. After some cards $\mathbf{c}^{(t)}$ have been played, player A knows their own remaining hand $\mathbf{A}^{(t)}$, and knows that player B's remaining cards $\mathbf{B}^{(t)}$ are drawn from $\overline{\mathbf{c}}^{(t)} \setminus \mathbf{A}^{(t)}$.

The prior probability that player B's current lowest card $B_1^{(t)}$ is some card value k is the probability that k is the minimum of m_B cards drawn uniformly without replacement from a set of $N = |\overline{\mathbf{c}}^{(t)} \setminus \mathbf{A}^{(t)}|$ cards. For this to be true, two conditions must be satisfied: (1) k needs to be in B's hand, and (2) all remaining $m_B - 1$ of B's cards are greater than k. This gives us the following formula for the history-conditioned **prior** over player B's lowest remaining card $B_1^{(t)}$:

$$P(B_1^{(t)} = k \mid \mathbf{c}^{(t)}) = \frac{\binom{|\overline{\mathbf{c}}^{(t)} \setminus \mathbf{A}^{(t)}| - 1 - \mathrm{rank}(k)}{m_B - 1}}{\binom{|\overline{\mathbf{c}}^{(t)} \setminus \mathbf{A}^{(t)}|}{m_B}} \tag{5}$$

The denominator is the total number of outcomes (i.e., the number of ways to choose B's hand of m_B cards from the available cards). The numerator is a bit more involved. Given that k is in player B's hand, the remaining $m_B - 1$ cards must be chosen from the cards in $\overline{\mathbf{c}}^{(t)} \setminus \mathbf{A}^{(t)}$ that are $> k$. We can denote this as $\mathrm{rank}(k)$. So, the number of elements greater than k is given by $|\overline{\mathbf{c}}^{(t)} \setminus \mathbf{A}^{(t)}| - 1 - \mathrm{rank}(k)$. Notice that the revealed card acts as a lower bound on all remaining cards in that player's hand, which tightens the prior at each step.

Likelihood (survival function). Player A observes that player B has been silent for time t since the last card was played. The likelihood of observing silence up to time t, given $B_1^{(t)} = k$, is the survival function of a hazard process. To capture the intuition that the hazard rate may change over time, we take $(T_B \mid B_1^{(t)} = k, \mathbf{c}^{(t)})$ to be distributed as Weibull with shape γ and scale $\tau \cdot (k - c^{(t)})$, where τ is the tempo parameter, and γ is a parameter controlling the change of the hazard rate. If $\gamma = 1$, this is equivalent to an Expontential distribution. If $\gamma > 1$, the hazard rate increases over time, which captures impatience and pressure to act. This gives us the following survival function:

$$P(T_B > t \mid B_1^{(t)} = k, \mathbf{c}^{(t)}) = e^{-\left(t/(\tau \cdot (k - c^{(t)}))\right)^{\gamma}} \tag{6}$$

In our experiments, we set $\gamma = 2$ for the Bayesian rule-based partner, based on feedback from a pilot study.

For Bayesian players with tempo calibration (see Section 3.2.2) we replace τ by the updated estimate after each play, which is itself a function of the previous game history.

Bayesian updating. At each time point t, Player A can update their belief over $B_1^{(t)}$, the lowest card in player B's remaining hand. By Bayes' rule,

$$P(B_1^{(t)} = k \mid T_B > t, \mathbf{c}^{(t)}) = \frac{P(T_B > t \mid B_1^{(t)} = k, \mathbf{c}^{(t)}) \cdot P(B_1^{(t)} = k \mid \mathbf{c}^{(t)})}{P(T_B > t \mid \mathbf{c}^{(t)})} \tag{7}$$

where the prior and likelihood are given by Equation (5) and Equation (6), respectively.

Decision rule. Player A plays their lowest card ($a_1^{(t)}$) at the first time t^* such that they are sufficiently confident (according to some threshold θ) that $a_1^{(t)} < B_1^{(t)}$. That is:

$$t^* = \min\{t \geq 0 : P(B_1^{(t)} > a_1^{(t)} \mid T_B > t, \mathbf{c}^{(t)}) \geq \theta\} \tag{8}$$

To compute $P(B_1^{(t)} > a_1^{(t)} \mid T_B > t, \mathbf{c}^{(t)})$, we sum the relevant values of the posterior:

$$P(B_1^{(t)} > a_1^{(t)} \mid T_B > t, \mathbf{c}^{(t)}) = \sum_{k \in \overline{\mathbf{c}}^{(t)} \setminus \mathbf{A}^{(t)} \,:\, k > a_1^{(t)}} P(B_1^{(t)} = k \mid T_B > t, \mathbf{c}^{(t)}) \tag{9}$$

A.4 Random player

At every second, random player A samples from a Bernoulli distribution with $p = 0.5$ to decide whether to play $a_1^{(t)}$ or wait.

B Computing Data Likelihood

To compute AIC, we need to compute the likelihood of the plays observed for a single game record.

Let the data be given by a sequence of D datapoints, each of the form (x_i, c_i), where x_i is the play (card and time) and c_i is the context of previous plays. The log-likelihood of the data for a given model with parameters θ is given by:

$$\sum_{i=1}^{D} \log P(x_i \mid c_i; \theta) \tag{10}$$

B.1 LINEAR MODEL

The Linear strategy predicts Player A's waiting time from the gap between Player A's lowest card and the previously played card. It multiplies this gap by the current tempo parameter. We used a Gamma distribution centered on this predicted waiting time, which allows the actual plays to occur earlier or later, but still maintains nonnegative waiting times. Using the notation defined in the main text, we modeled

$$T_{A,i} \sim \text{Gamma}\left(\kappa,\ \frac{\max\left\{\tau_i\left(a_1^{(t)} - c(t)\right), \delta_i/2\right\}}{\kappa}\right), \tag{11}$$

where the Gamma distribution is parameterized by shape and scale.

Its mean is

$$\max\left\{\tau_i\left(a_1^{(t)} - c(t)\right), \delta_i/2\right\},$$

and its variance is the square of the mean divided by κ. The parameter κ controls timing variability: larger values produce waiting times more tightly concentrated around the Linear prediction.

Without calibration, τ_i remained fixed throughout the game. With calibration, the tempo estimate was updated after each observed, non-forced play using the rule defined in the main text:

$$\hat{\tau}_i = (1-\eta)\hat{\tau}_{i-1} + \eta\frac{t_i}{c_i - c_{i-1}}. \tag{12}$$

Here, η determines how strongly the most recent observed play changes the current tempo estimate.

B.2 BAYESIAN MODEL

The Bayesian strategy uses Player B's silence to update Player A's confidence that Player A holds the lower card:

$$P\left(B_1^{(t)} > a_1^{(t)} \mid T_B > t, C(t)\right).$$

As described earlier, this posterior is based on the Weibull survival model:

$$P\left(T_B > t \mid B_1^{(t)} = k, C(t)\right) = \exp\left[-\left(\frac{t}{\tau(k - c(t))}\right)^{\gamma}\right].$$

Instead of using a fixed decision threshold, we converted Player A's confidence into an instantaneous play rate:

$$\lambda_A(t) = \exp\left[\beta_0 + \beta_1 P\left(B_1^{(t)} > a_1^{(t)} \mid T_B > t, C(t)\right)\right],$$

where $\beta_1 \geq 0$. The parameter β_0 determines the baseline play rate, and β_1 determines how strongly the play rate increases with confidence. We do this to give Player A a nonzero play rate even at low confidence.

The probability that Player A remains waiting until time t is then:

$$P(T_A > t) = \exp\left[-\int_0^t \lambda_A(u)\, du\right].$$

So, if Player A played during the observed interval $[t_1, t_2)$, we calculated:

$$P(t_1 \leq T_A < t_2) = P(T_A > t_1) - P(T_A > t_2).$$

If Player B played first at time t, we instead calculated $P(T_A > t)$, the probability that Player A would still be waiting at that time. In the +Calibration condition, τ was updated after each observed play using the calibration rule defined above.

B.3 RANDOM STRATEGY

The Random strategy assumes that Player A has a constant probability p of choosing to play during each one-second interval. Let N be the zero-indexed one-second interval in which Player A plays:

$$P(N = n) = p(1-p)^n, \qquad n = 0, 1, 2, \ldots. \tag{13}$$

Conditional on selecting interval N, the play time was uniform within that second:

$$T_A = N + U, \qquad U \sim \text{Uniform}(0, 1). \tag{14}$$

This allowed the model to assign probabilities to continuous and sub-second observations. For $t \geq 0$, let $n = \lfloor t \rfloor$ and $r = t - n$. The resulting CDF is

$$F(t) = 1 - (1-p)^n + p(1-p)^n r. \tag{15}$$

The parameter p was fitted separately to each game record.

B.4 CONSTANT STRATEGY

The Constant wait model assumed that Player A's waiting time did not depend on the card values or game history, so the expected waiting time is constant:

$$T_i \sim \text{Gamma}\left(\kappa, \frac{\mu}{\kappa}\right), \tag{16}$$

where the Gamma distribution is parameterized by shape and scale. It has mean μ and variance μ^2/κ. Both μ and κ were fitted separately to each game record.

B.5 PARAMETER FITTING AND MODEL COMPLEXITY

For each candidate model and game record, we selected the parameter values that maximized the likelihood. We evaluated the initial tempo parameter over

$$\tau \in \{0.1, 0.3, 0.5, 0.7, 1.0, 1.3, 1.5\}, \tag{17}$$

the calibration learning rate over

$$\eta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}, \tag{18}$$

and the Weibull shape parameter over

$$\gamma \in \{0.5, 0.7, 1.0, 1.3, 1.5, 2.0, 3.0\}. \tag{19}$$

The Random probability p was optimized continuously over $(0, 1)$. For the Gamma models, κ was optimized continuously over

$$\kappa \in [0.02, 10{,}000]. \tag{20}$$

For Constant Wait, μ was optimized continuously over a positive, record-specific range covering the observed waiting times. For Confidence Bayesian, β_0 and β_1 were optimized continuously over

$$\beta_0 \in [-12, 6], \qquad \beta_1 \in [0, 20]. \tag{21}$$

Values selected from the parameter grids were counted as fitted parameters. All candidate models were fitted to the same timing decisions within a game record, and values were compared only within that record.

C HUMAN EXPERIMENTS

All participants were native English speakers residing in the United States. To ensure data quality, we restricted participation to Prolific users with an approval rate above 95%. To guarantee the data collection quality, we used hesitation time on standalone cards (i.e., the cards played at the end of each level after the partner has finished playing) as an attention check, and filtered out participants whose hesitation time exceeded 1 minute (1 participant was excluded on this basis). Participants were compensated at an hourly rate of \$12. We provide participants a performance-based bonus of up to \$2. We computed a performance score as $-(0.5 \cdot \text{time}_z + 0.5 \cdot \text{error}_z)$, where time_z and error_z denote the z-scored completion time and error rate, respectively, computed across all participants. Participants in the top 25% of the performance-score distribution received a \$2 bonus, while those in the 25th-50th percentiles received a \$1 bonus.

C.1 Game Interface for Human and CUA Experiments

The game interface is shown in Figure 6 and the instruction pages are shown in Figures 7 to 9.

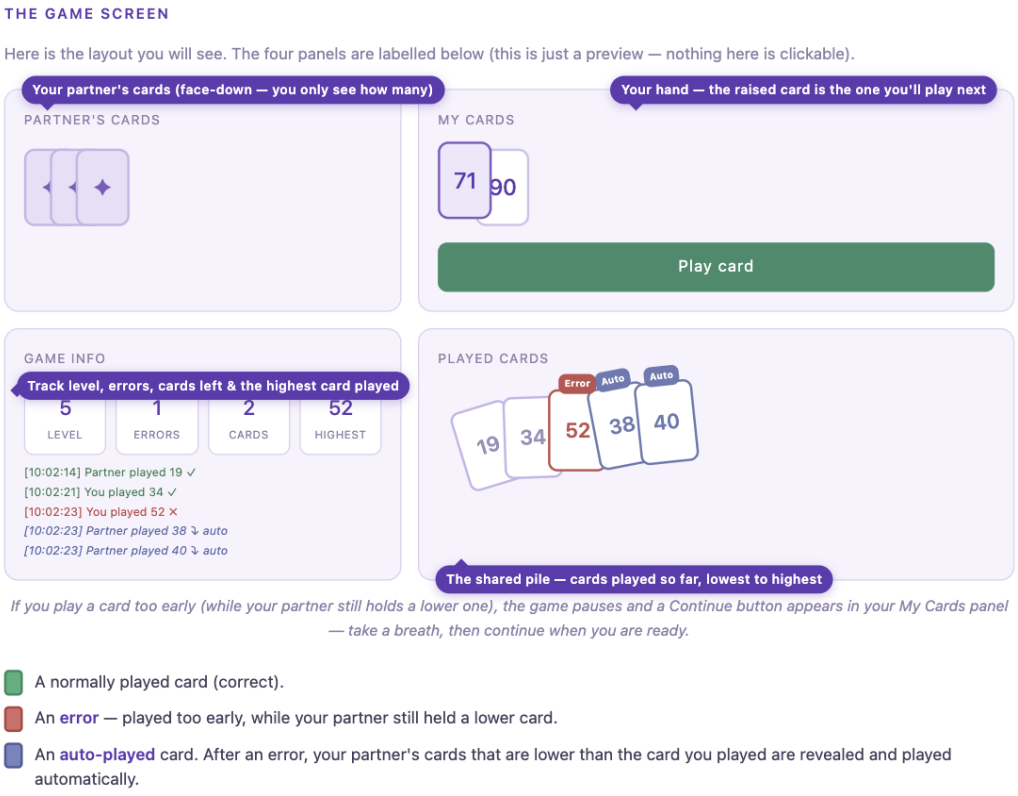


Figure 6: The game interface of *The Mind*.

D LLM Experiments

D.1 Hyperparameters

The hyperparameter settings are listed in Table 1. Unless otherwise specified, we use the default settings.

Models	Settings
`gemini-3.7-flash`	thinking_level: High; temperature: 1.0; TopP: 0.95; TopK: 64
`gpt-5.6-luna`	API model; default setting; thinking=medium
`gpt-6-astra`	API model; default setting
`gpt-oss-120b`	temperature: 0.7; max_tokens: 10,000
`gpt-oss-20b`	temperature: 0.7; max_tokens: 10,000
`Qwen3-32B`	temperature: 0.6; TopP: 0.95; TopK: 20; max_tokens: 10,000
`Qwen3-8B`	temperature: 0.6; TopP: 0.95; TopK: 20; max_tokens: 10,000
`Mistral-Medium-3.5-128B`	temperature: 0.7; max_tokens: 10,000
`Llama-3.1-70B-Instruct`	temperature: 0.6; max_tokens: 10,000
`Llama-3.1-8B-Instruct`	temperature: 0.6; max_tokens: 10,000
`Llama-3.2-1B-Instruct`	temperature: 0.6; max_tokens: 10,000

Table 1: Hyperparameter settings for the models.

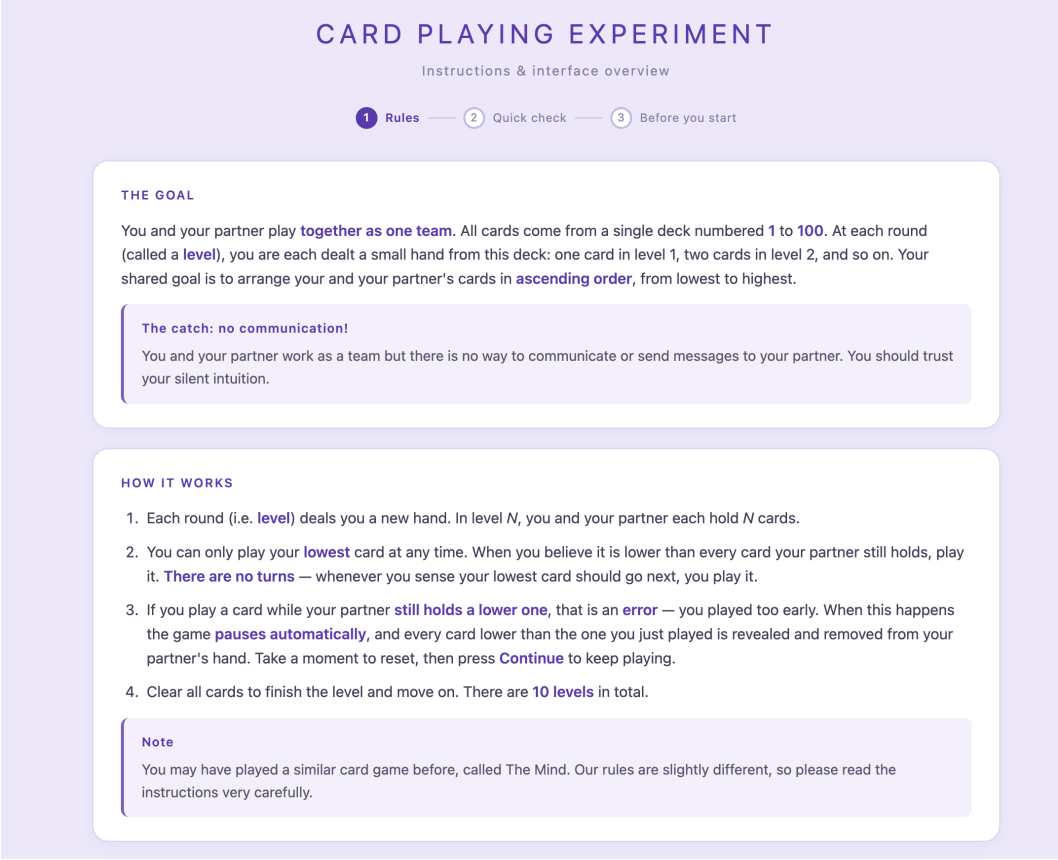


Figure 7: Instruction Page (1/3)

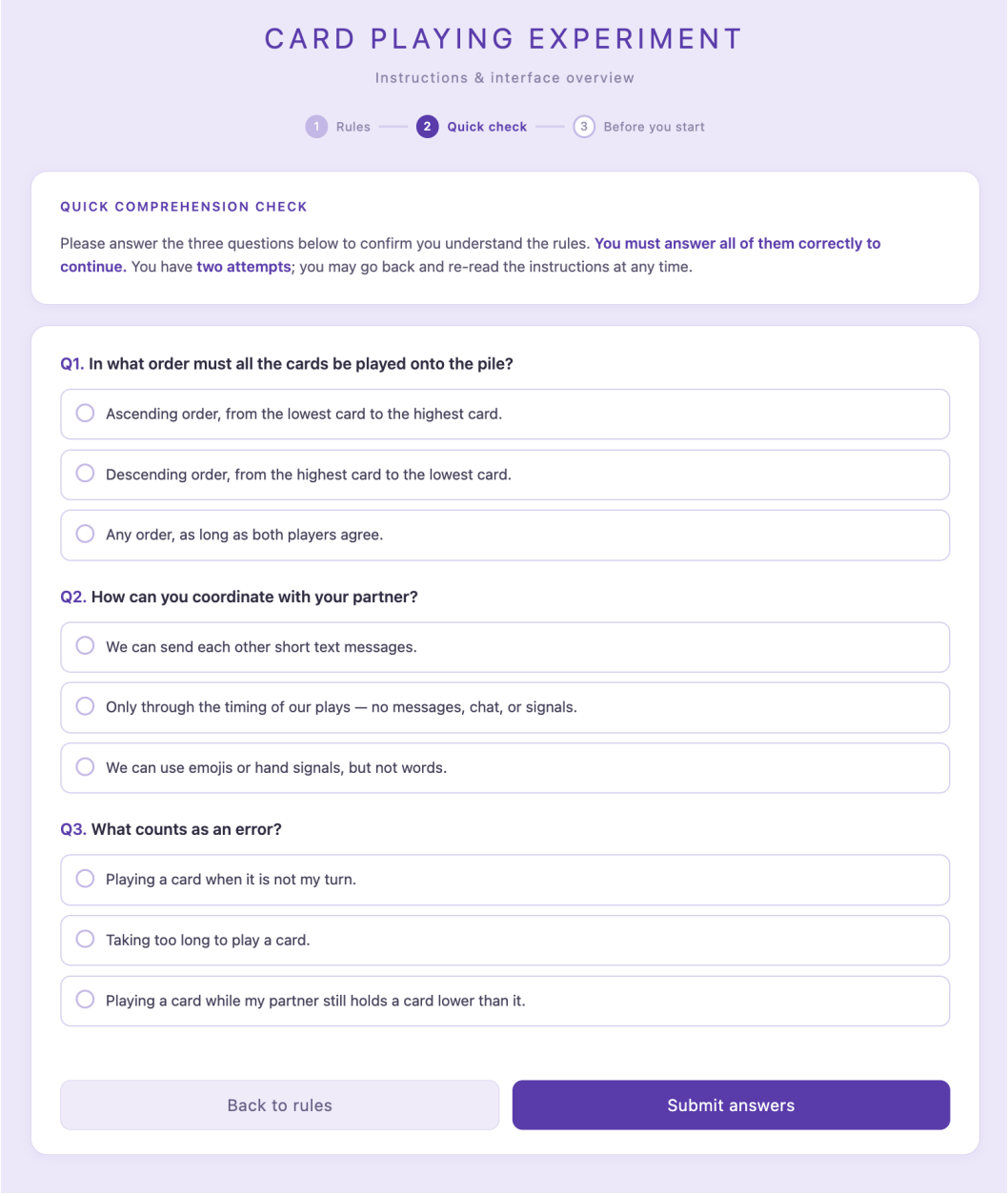


Figure 8: Instruction Page (2/3)

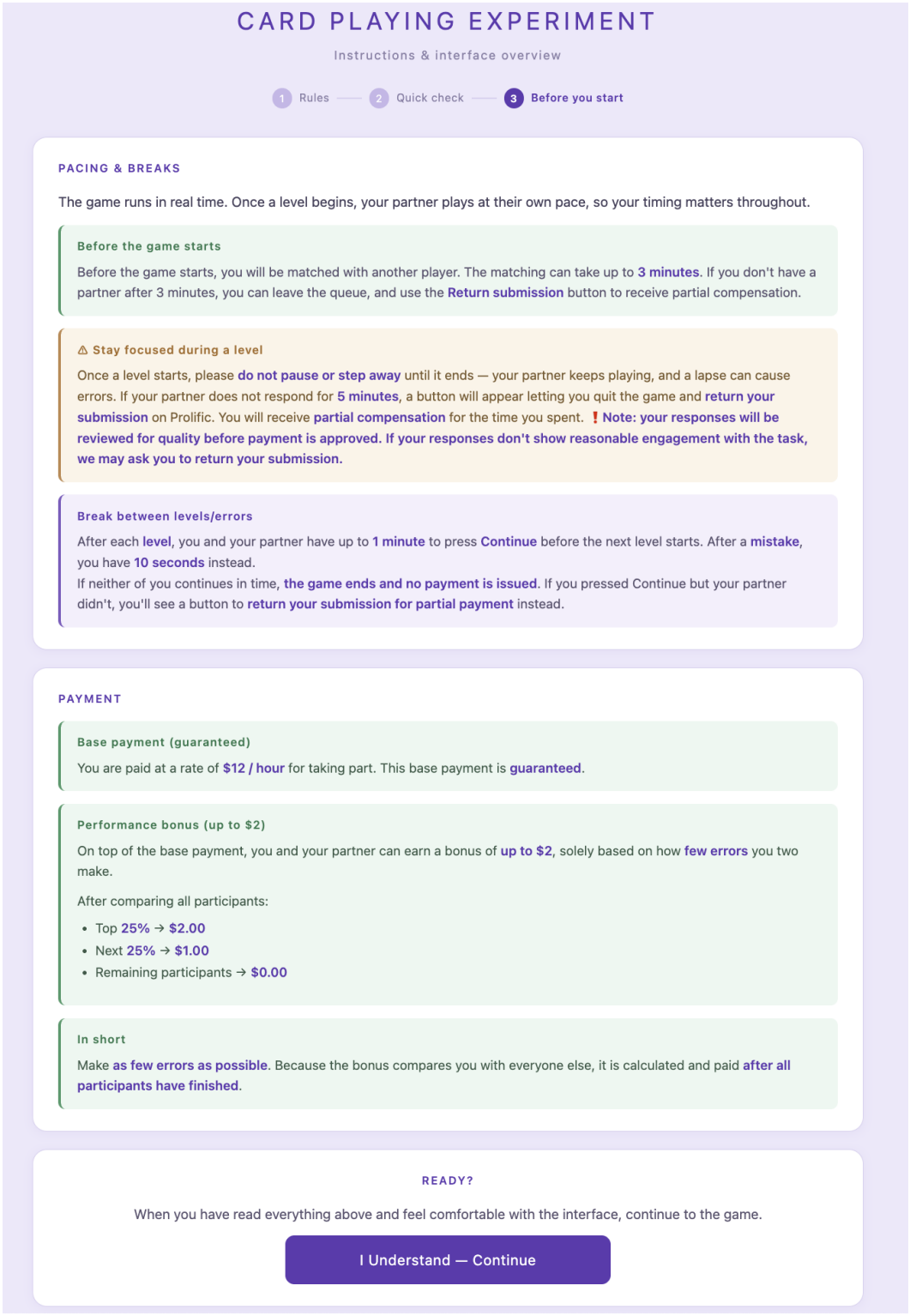


Figure 9: Instruction Page (3/3)

D.2 PROMPT

The game prompt is composed of two main parts: the game rule and the query. The query is composed of three parts: the most recent plays, the current game state, and the task instruction. Templates for these three parts across conditions are listed in Table 2.

Prompt of game rule

You and your partner are playing a cooperative card game. This game is similar to a game called The Mind. However, our rules are slightly different, so please read the instructions very carefully.

You and your partner play together as one team. At each round (called a level), you and your partner are each dealt a new hand from a fresh deck of cards numbered 1 to 100. Cards from previous levels do not carry over or affect the cards dealt in the new level. You receive one card each in level 1, two cards each in level 2, and so on.

Your shared goal is to arrange your and your partner's cards in ascending order, from lowest to highest. You cannot see your partner's cards. You only know how many cards they have left and what cards have been played.

How it works: 1. Each level deals you a new hand. In level N, you and your partner each hold N cards. 2. You can only play your lowest card at any time. When you believe it is lower than every card your partner still holds, play it. There are no turns — whenever you sense your lowest card should go next, you play it. 3. If you play a card while your partner still holds a lower one, that is an error — you played too early. When this happens the game pauses automatically, and every card lower than the one you just played is revealed and removed from your partner's hand. 4. Clear all cards to finish the level and move on. There are 10 levels in total.

The performance of the game is equally based on **how fast you finish** and **how few errors you make**. In other words, you should play QUICKLY and make as FEW errors as possible.

Prompt of the query

Below, you will see the 5 most recent plays in the game and the current game state. If the 5 most recent plays include the beginning of a new level, it will be clearly marked. Your task will be to decide whether to continue waiting or play your lowest card.

===5 MOST RECENT PLAYS===
——BEGIN LEVEL X, DEAL NEW HANDS AND RESET CLOCK—— *[This exact line appears only when a new level begins.]*
[recent play history; format varies by prompt condition]
===CURRENT GAME STATE===
[current hand, cards played, level info; format varies by prompt condition]
===YOUR TASK===
[decision prompt, phrased per condition]

Condition	5 Most Recent Plays (Example)	Current Game State	Your Task
Time-based (1 second)	Level: <level> Card played: <card_value> Player: <you_or_your_partner> Time since the start of level <level>: <seconds> second(s).	You are currently on level <level>. Current hand: [<card_values>] Your partner has <num_cards_left> card(s) left. The highest card played so far: <highest_card_value> You have been waiting for <seconds> second(s) since the last successfully played card <card_value>.	Based on the above information, do you want to wait one more second or play your lowest card, i.e., <card_value>? Respond with exactly one word: PLAY or WAIT.
Time-based (0.5 second)	Level: <level> Card played: <card_value> Player: <you_or_your_partner> Time since the start of level <level>: <seconds> second(s).	You are currently on level <level>. Current hand: [<card_values>] Your partner has <num_cards_left> card(s) left. The highest card played so far: <highest_card_value> You have been waiting for <seconds> second(s) since the last successfully played card <card_value>.	Based on the above information, do you want to wait another 0.5 second or play your lowest card, i.e., <card_value>? Respond with exactly one word: PLAY or WAIT.
Time-based (0.25 second)	Level: <level> Card played: <card_value> Player: <you_or_your_partner> Time since the start of level <level>: <seconds> second(s).	You are currently on level <level>. Current hand: [<card_values>] Your partner has <num_cards_left> card(s) left. The highest card played so far: <highest_card_value> You have been waiting for <seconds> second(s) since the last successfully played card <card_value>.	Based on the above information, do you want to wait another 0.25 second or play your lowest card, i.e., <card_value>? Respond with exactly one word: PLAY or WAIT.
Time Estimate	Level: <level> Card played: <card_value> Player: <you_or_your_partner> Time since the start of level <level>: <seconds> second(s).	You are currently on level <level>. Current hand: [<card_values>] Your partner has <num_cards_left> card(s) left. The highest card played so far: <highest_card_value> You have been waiting for <seconds> second(s) since the last successfully played card <card_value>.	Based on the information above, how many seconds would you like to wait before playing your lowest card, i.e., <card_value>? Respond with only a single numeric value. Do not include any explanation or units.
Wait-based (count)	Level: <level> Card played: <card_value> Player: <you_or_your_partner> Decisions since the last played card: <decision_history> *<decision_history> is a comma-separated sequence of each player's past decisions (e.g., WAIT, WAIT, WAIT, WAIT, WAIT, PLAY)*	You are currently on level <level>. Current hand: [<card_values>] Your partner has <num_cards_left> card(s) left. The highest card played so far: <highest_card_value> Your decision(s) since the last successfully played card <card_value> are: <decision_history>.	Based on the above information, do you want to wait one more time or play your lowest card, i.e., <card_value>? Respond with exactly one word: PLAY or WAIT
Wait-based (count)	Level: <level> Card played: <card_value> Player: <you_or_your_partner> Number of WAITs since the last played card <card_value>: <num_of_waits>.	You are currently on level <level>. Current hand: [<card_values>] Your partner has <num_cards_left> card(s) left. The highest card played so far: <highest_card_value> You have chosen to WAIT <num_of_waits> times since the last successfully played card <card_value>.	Based on the above information, do you want to wait one more time or play your lowest card, i.e., <card_value>? Respond with exactly one word: PLAY or WAIT

Table 2: Prompt template shared across conditions, with condition-specific content for each templated section.

E Complementary Results

E.1 BIC

As shown in Figure 10, where the AIC and BIC analysis differed, BIC preferentially selected simpler Random or Linear without calibration models.

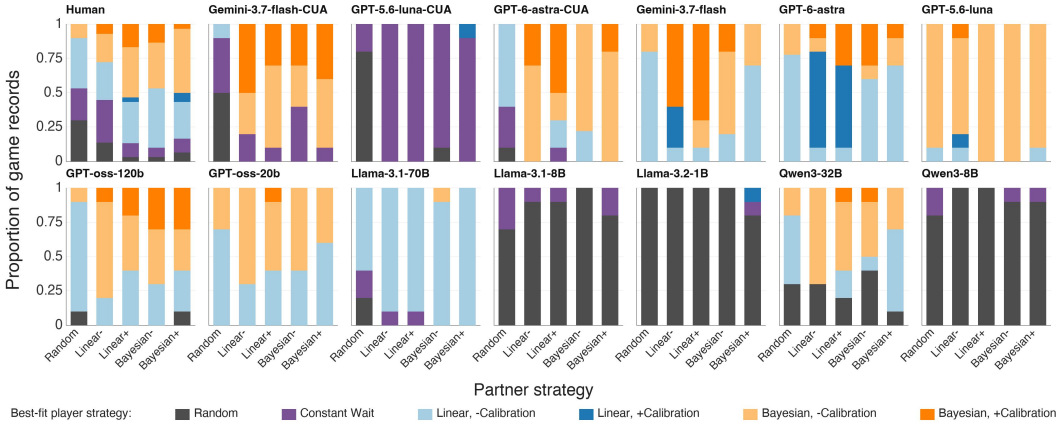


Figure 10: Proportion of Human, LLM, and CUA Game Records Best Fit by Strategy Models (BIC)

E.2 Efficiency comparison across players

The top-20 most efficient player-partner combination is reported in Figure 11.

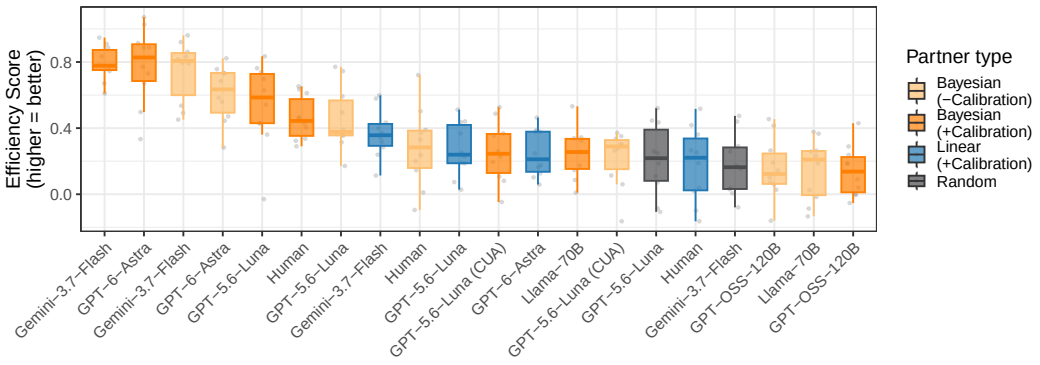


Figure 11: Top 20 most efficient player-partner combinations.

E.3 Time consistency and CV of time of models across prompt types

Figure 12 shows the Coefficient of variation of different models across prompt types.

Figure 14 shows the time consistency of other models that are not reported in the main text. We find similar patterns: simply changing the temporal granularity used to represent waiting time in the prompt can lead to substantially different behavioral patterns.

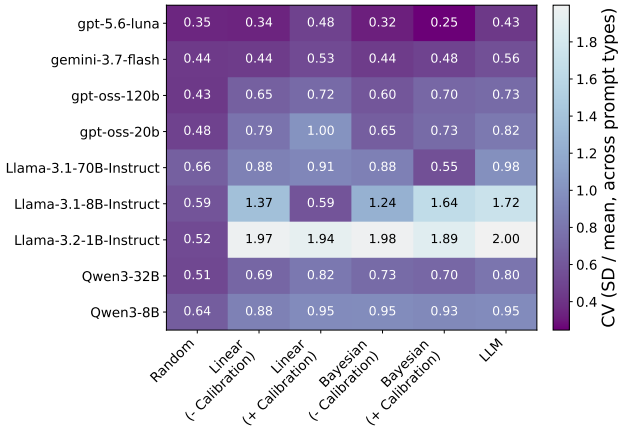
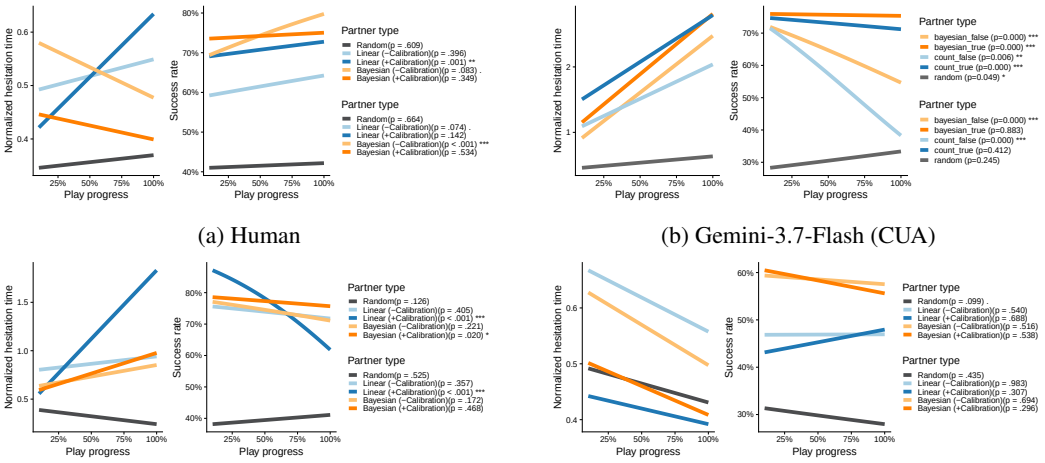


Figure 12: Coefficient of variation across prompt types.

E.4 Statistical Analysis of Player's Adaptive Behavior

Figures 15 and 16 show the fitted regression lines for normalized hesitation time and success rate for each LLM and prompt combination respectively. Figure 13 visualizes the fitted regression lines for normalized hesitation time and success rate over the course of a hand, for humans and three CUAs. We find that (i) most prompted-based LLMs show slower response time and lower success rate as time goes; (ii) even in more comparable human-CUA settings, only humans playing with Bayesian Player (without Calibration) show both (marginally) statistically significant faster speed and higher success rate over the course of the game.



(a) Human

(b) Gemini-3.7-Flash (CUA)

(c) GPT-6-astra (CUA)

(d) GPT-5.6-luna (CUA)

Figure 13: Fitted regression lines for normalized hesitation time (left) and success rate (right) over the course of a hand, for humans and three CUAs, by partner type.

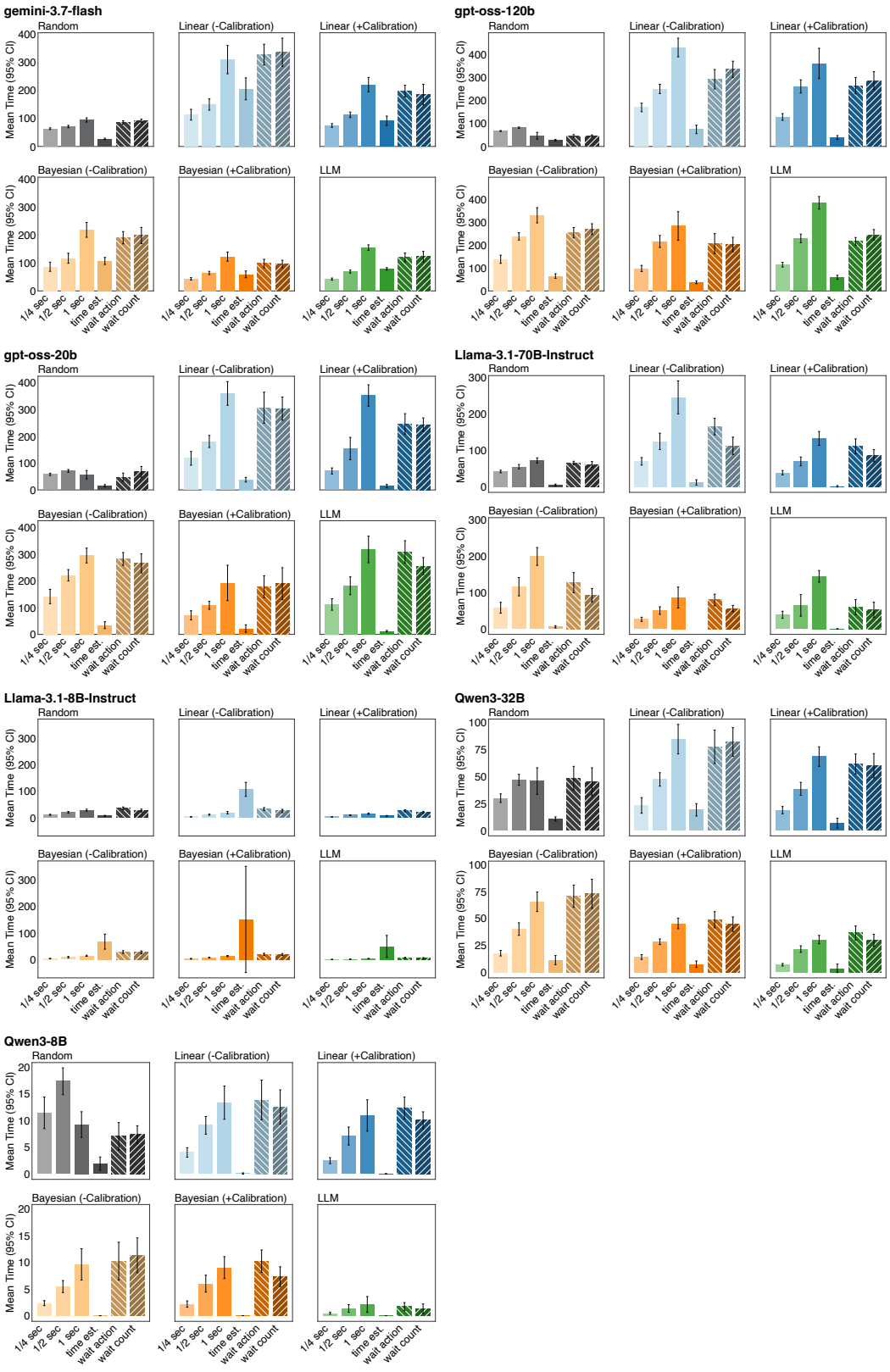


Figure 14: Mean play time by partner type and prompt type for the LLMs. Error bars show 95% CI.

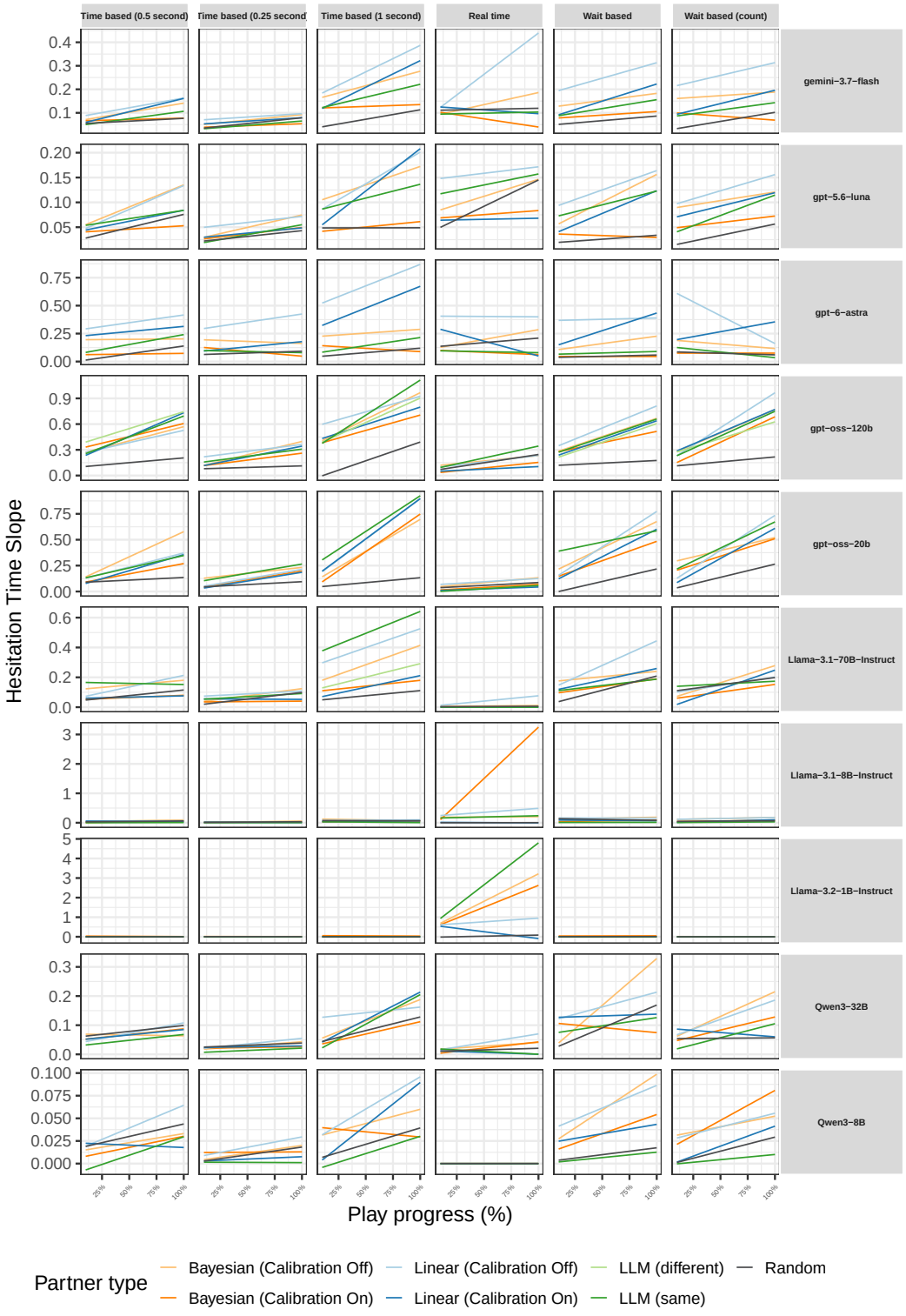


Figure 15: The hesitation slope from the mixed model for each model by prompt setting.

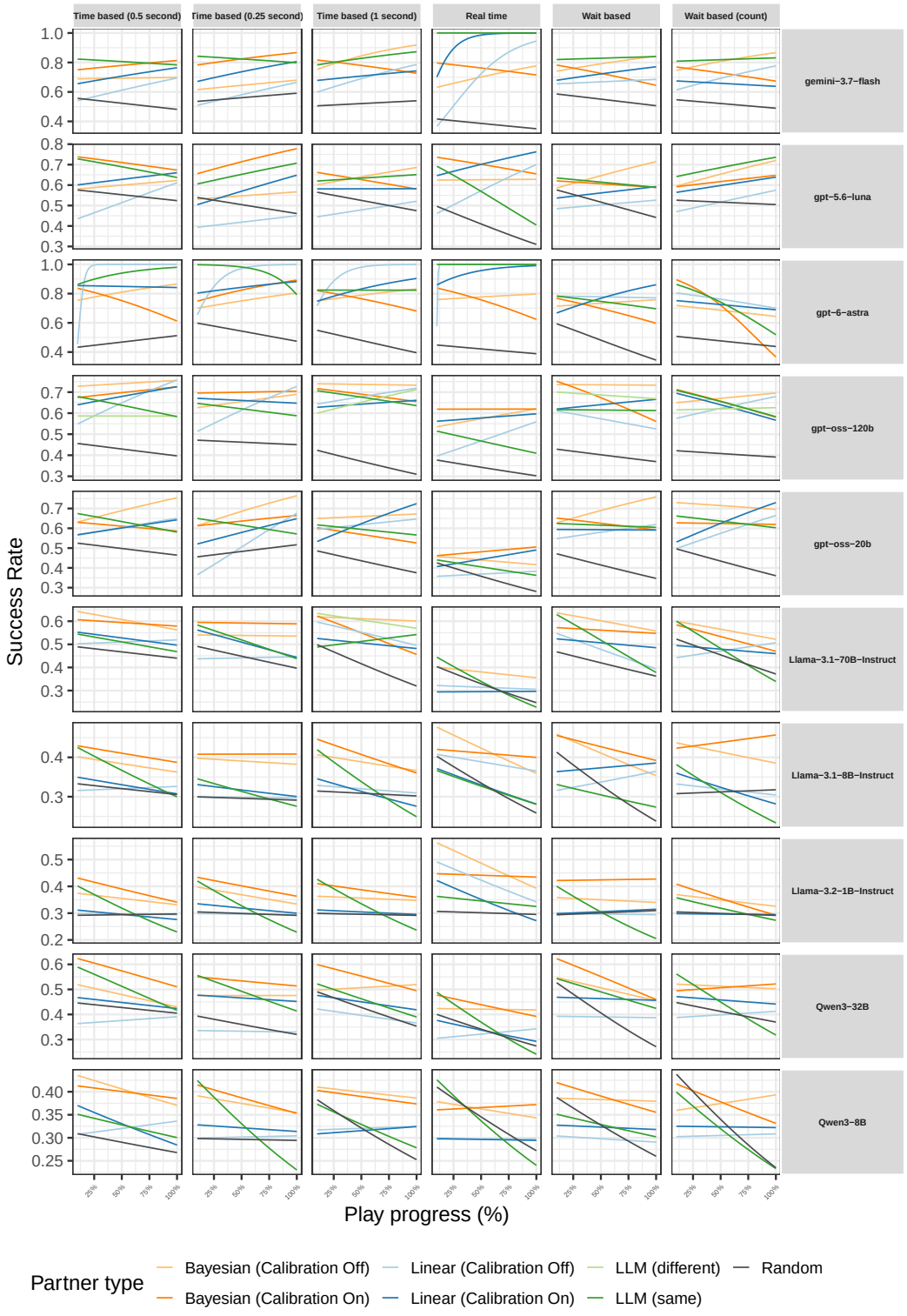


Figure 16: The success slope from the mixed model for each model by prompt setting.