Solving strategic social coordination via Bayesian learning

Amrita Lamba[1] (a_lamba@mit.edu), **Sean Dae Houlihan**[2] (dae.houlihan@dartmouth.edu),
Rebecca Saxe[1] (saxe@mit.edu)
[1]Brain and Cognitive Sciences, Massachusetts Institute of Technology
[2]Department of Computer Science, Dartmouth College

Abstract

Repeated social coordination is a crucial aspect of daily life in which individuals strategically distribute labor and resources, often to accomplish complex tasks and goals. However, social coordination is also very challenging because humans often have competing interests, especially when successful coordination persistently leaves one party better off, entrenching inequality. Here we use a novel task, the Asymmetric Social Exchange (ASE) Game, to study how individuals learn to coordinate with different kinds of social partners and how individual trait variability on key social dimensions related to negative evaluation (i.e., social anxiety), impacts compliance with disadvantageous conventions (N = 675). Using two kinds of Bayesian models, one that learns from experience and one that builds a causal model of others' hidden motivations, we show that differences in coordination strategies arise from both individual learning differences and from expressed social preferences. Further, we find that social anxiety increases compliance with disadvantageous payoffs.

Keywords: asymmetric social coordination; social learning; social anxiety; hierarchical Bayesian model; Bayesian causal model; reinforcement learning

Introduction

In daily life, people regularly coordinate social interactions that accomplish what would otherwise be difficult or impossible to manage alone. For example, families distribute chores amongst members and nations establish agreements to tackle major issues, such as climate change. In entering such social arrangements, especially repeatedly, labor and resources are often difficult to evenly split (Sákovics & Steiner, 2012). Asymmetric divisions of labor and resources can perpetuate inequality when one party benefits more from joint coordination, increasing tension between individuals and making sustained coordination more difficult to achieve. For instance, unevenly distributing household chores can result in interpersonal tension when inequality persists. In such cases, individual variability in one's desire to smooth social tension can impact how people manage disadvantageous roles. Thus, when we repeatedly coordinate with others to achieve a common goal or task, we may experience tension between advancing our own interests while also being helpful to others.

There is much to learn about the cognitive processes that influence how people respond to different kinds of inequality perpetuated in asymmetric coordination problems, an issue of increased relevance in an era where many of the most pressing problems, such as mitigating climate change, require sustained coordination. In the current study, we investigate how people dynamically form coordination strategies from direct experience interacting with different kinds of social partners. We also study how trait variability on key individual dimensions related to fear of negative evaluation are related to compliance with disadvantageous social conventions. As one index of individual variability we consider how social anxiety—a trait associated with negative social theories of others (Kocovski & Endler, 2000)—relates to compliance with unfair terms.

The cognitive challenge of repeated social coordination. A major challenge for adaptively managing social interactions is inferring the hidden intentions and motivations of others, crucial social variables that cannot be directly observed. This presents several unique challenges for social interaction: each individual we meet must be evaluated so we can tailor our behavior appropriately. A wealth of prior work has shown that humans leverage extensive social learning mechanisms to build generative causal models of other people's behavior (Houlihan et al., 2023; Kleiman-Weiner et al., 2016), and to incrementally learn the behavior of others through observation and interaction (Lamba et al., 2023, 2024; Yu & Thompson, 2024).

Moreover, repeated social coordination frequently results in the personal dilemma of advancing one's own interests or being a team player. Prior work from behavioral economics has extensively studied how people weigh out such interpersonal costs and benefits. In behavioral economics, rational choice in social dilemmas can be captured through utility functions which capture the linear combination of weights on inequality aversion and expected rewards, in which individuals can express preferences as disadvantageous inequality aversion (DIA) and advantageous inequality aversion (AIA) through weights on these utilities (Fehr & Schmidt, 1999; Fehr et al., 2005). Thus, the cognitive challenge of repeated social coordination is learning the hidden intentions and motivations of others and then tailoring a strategic response that reflects the learned values of others.

Social tension and fear of negative evaluation. Repeated coordination may also tap into a unique dimension of population phenotypic variation, namely how individuals respond to social tension. When it is clear that a social partner will persistently pursue the advantageous position, the resulting social tension can spark conflict. Individuals with increased desire to mitigate social tension may be more likely to com-

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

ply with disadvantageous social conventions to avoid conflict. Social anxiety, a form of pathological anxiety characterized by excessive social avoidance behavior and irrational fear of negative evaluation, can increase compliance with unfair social conventions. Although symptoms associated with social anxiety have been well-established in the clinical literature (Carleton et al., 2007), little is known about the cognitive and behavioral effects of social anxiety in the context of strategic coordination, where a reduced tendency to advocate for yourself could adversely affect one's well-being. In the current work, we examine the relationship between self-reported levels of social anxiety and compliance with disadvantageous social conventions.

Cognitive models of repeated coordination. Given the distinct challenges of social interaction, people likely recruit multiple learning mechanisms to solve strategic coordination (Kleiman-Weiner et al., 2018). Specifically, we may form distinct representations of other individuals through incremental learning and by constructing a mental causal model of their behavior to infer their motivations and intentions. Different modeling architectures can provide clarity on distinct aspects of this social reasoning process.

Bayesian Reinforcement Learning (B-RL). B-RL, a form of RL that updates belief distributions over a partner agent's expected actions based on observation (Dayan & Daw, 2008), can expediently learn optimal choice policies (i.e. what to do to maximize outcomes). Critically, B-RL is model-free and thus treats the partner agent's internal state as a black box, and is therefore only capable of forming a choice policy based on observed outcome history. However, in our model described in this work, B-RL can capture the distinct contributions of the learned action values and an individual's expressed social preferences by forming a choice policy aligned with one's goals. By orthogonalizing the distinct contributions of incremental learning and the expression of social preferences, B-RL can evaluate (1) whether these processes are sufficient to recapitulate human coordination dynamics and (2) how trait variability on the desire to smooth social tension (i.e., social anxiety) is related to these distinct components.

Bayesian causal model. Accurate prediction is no guarantee of ontological correctness—a model can make arbitrarily precise predictions without necessarily being structurally similar to the process being modeled. For Bayesian causal models, prediction is insufficient without understanding. These models target isomorphism with the abstract causal structure of an information processing system. Applied to human behavior, a model of this type formalizes a hypothesis, expressed in cognitive terms, about the computational structure of the mind. Behavior is treated as the output of approximately-rational reasoning over a mental model (an "intuitive theory"). By explicitly representing causal relationships between the world, mental states (like preferences and beliefs), and behavior, a Bayesian causal model aims to recapitulate human cognition.

In the context of strategic coordination, simulated players are endowed with preferences for monetary and social outcomes, along with beliefs about the world and their social partners. They choose actions that noisily maximize their expected utilities given their preferences and beliefs. These models are naturally recursive; simulated players choose actions based on predictions of their partners' choices by reasoning over an embedded mental model that represents their beliefs about their partners' preferences and beliefs, which can recursively include their partners' beliefs about the player (and so on). Over iterated trials with multiple partners, simulated players hierarchically update their beliefs, simultaneously learning about specific partners and the broader population. For example, inferring that one partner is highly motivated to avoid disadvantageous inequity will lead a player to think it is more probable that other partners are similarly motivated, and the player will require stronger evidence to believe otherwise. Like B-RL, the Bayesian causal modeling approach predicts *what* participants chose, trial by trial, by modeling participants' learning, preferences and utility-based decisions. However, whereas B-RL is model-free, a Bayesian causal model instantiates a richly structured model of players' minds to recapitulate *how* people learn, strategize, and act, and thereby explain *why* players behaved as they did.

Asymmetric Social Exchange Game. In the current study we use a novel task, the asymmetric social exchange (ASE) game, to study repeated coordination. The ASE game is a modified and iterated version of the two-player Leader-Follower game (Leitmann, 1978) in which players can only coordinate by accepting uneven payouts, a situation which forces inequality. In this task, we show that learning dynamics and utility-based social preferences play a key role in shaping strategic coordination in both advantageous and disadvantageous inequality scenarios. Further, participants coordinate at significantly lower rates when a partner forces the individual to accept a disadvantageous social convention. In this vein, individuals with increased social anxiety are especially likely to comply with disadvantageous social conventions, accepting lower payouts compared to their partner.

With our B-RL model, we show that learning and utility-based social preferences exert distinct influences on strategic coordination. By decoupling learning mechanisms from utility-based preferences, we compare the distinct influence of social anxiety on these factors, namely in responding to disadvantageous inequality. We show that people with social anxiety exhibit increased compliance with disadvantageous inequality due to lower expression of inequality aversion, not from a learning bias. The Bayesian causal model infers the hidden motivations of each partner by learning about each agent's social preferences. This causal model recapitulates our key finding, that social anxiety is associated with reduced expression of social preferences. This hints that the pathology of social anxiety is related to negative social theories and expectations, leading to increased compliance with disadvantageous outcomes.

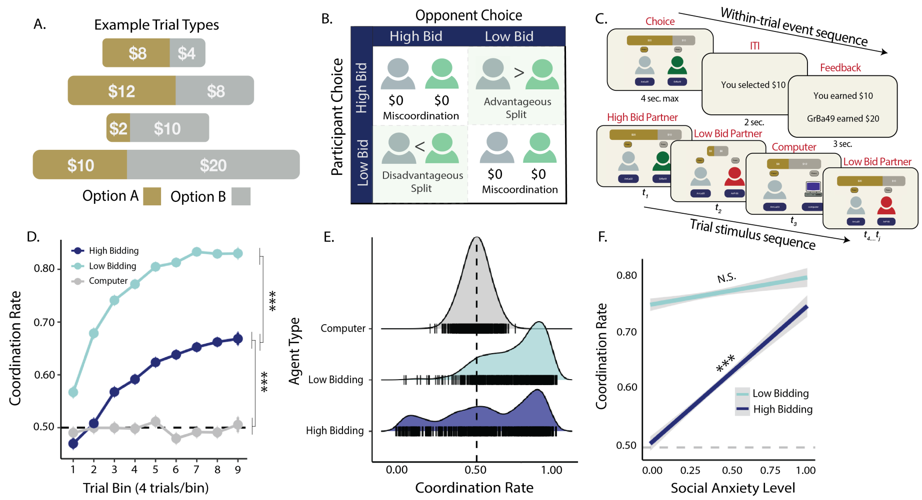

Figure 1: Asymmetric social exchange (ASE) game. **A.** Example pot sizes and offers. **B.** ASE game payoff matrix. Players only earn their bids by coordinating on the off-diagonal strategy, one partner bidding high and one partner bidding low. **C.** Task trial structure. Participants are given 4 seconds to place their bid simultaneously with their partner, before observing the trial outcome. Interactions with each partner algorithm are randomly interleaved. **D.** Coordination rates across task blocks. Participants learn to play the complementary policy with each agents but show significantly lower coordination with the high bidding agent. **E.** Distribution of individual coordination rates with each partner agent. Participants showed the most variability on trials with the high bidding agent. **F.** Effect of trait variability on coordination. Social anxiety predicts increased coordination with the high bidding agent, but not with the low bidding agent.

Methods

Participants. Data was collected from the Prolific online recruitment platform (N = 675). Study participants were required to reside within the United States and indicate proficiency in English. The experiment was approximately 45 minutes long, and all participants were paid $12 in compensation. Task performance was incentivized, the outcome of one trial was randomly selected for a bonus payment (mean bonus was $5). Our study protocol was approved by Massachusetts Institute of Technology's Institutional Review Board.

ASE Game. Participants completed 136 trials of the ASE game. On each trial, participants were shown a pot, represented as a horizontal bar, split into two portions, option A (beige) and option B (gray). The bar was always divided into unequal portions to create two uneven offers, a high offer and a low offer (e.g., $8-$4, see Fig. 1A). Participants were paired with one of four possible stimuli, three of which were "online partners", actually preprogrammed agents, and one explicitly randomized computer trial. The payoff matrix of the game was a two-by-two design in which each player could make a binary choice to select the high or low option through selecting option A or B. If both players selected the same bid, the round would result in a $0 payout (Fig. 1B).

Partner-stimuli. The partner algorithms were designed to emulate different kinds of strategies. Two of the partners played fixed policies throughout the task. One partner always selected the high bid, implementing a greedy strategy. Another partner always selected the low bid, resembling a deferential strategy. We designed another algorithm that attempts to implement an equitable strategy by alternating between selecting the high and low offer (high, low, high, low), similar to how people may attempt to equalize resources over time[1]. As a nonsocial and random control, participants also played computer trials which were explicitly randomized. To simulate online game play, participants were asked to generate a username consisting of the first two letters of their first name, the first two letters of their last name, and two digits. Partner identities were randomly assigned to colored avatars (red, green, and yellow) and usernames ("AnPr55", "LuTe61", "GrBa49"). Participants encountered each stimulus once every four trials to reduce memory load between exposures to the stimulus, in random interleaved order (Fig. 1C).

Trial structure. At the onset of each round, participants were paired with a partner and saw the available offers. Participants were given a 4-second decision deadline to indicate their choice (option A or B). After a brief 2-second interstimulus interval (ITI), participants were shown their outcome and their partner's outcome. After feedback was presented for 3 seconds, participants experienced a randomly distributed 2-4 second delay while the next round was being "setup". If participants did not respond within the 4-second response window, both players earned $0 on that round. Participants completed 34 rounds with each stimulus (136 rounds total).

[1] Given the additional complexity of evaluating performance with alternator, we save discussion and analysis of performance with the alternator for future work.

Experiment setup. All participants were instructed about the payoff matrix at the beginning of the experiment. To be admitted into the study, participants needed to answer all three questions of a comprehension quiz correctly. To estimate social anxiety levels we used the Social Interaction Anxiety Scale (SIAS; Peters, 2000). All participants were debriefed about the use of deception at the end of the experiment. Prior to online data collection all intended statistical analyses and predicted behavioral effects were preregistered on the Open Science Framework (OSF) repository. The study preregistration can be found here: https://osf.io/9u32k.

Results

Coordination. We expected coordination rates, the portion of trials in which participant's enacted the complementary policy, to differ across high and low bidding agents as participants could be more averse to accepting lower payouts. We observed exactly this, participants coordinated at significantly lower rates with the high bidding agent compared to the low bidding agent ($t = 41.40$, $p < .001$; Fig. 1D). Further, participants substantially varied in how they managed coordination strategies with the high bidding agent (Fig. 1E) with some participants accepting lower payouts while others resisted this convention.

Social anxiety measures. We next examined the impact of social anxiety levels on coordination, with the prediction that social anxiety would increase compliance with disadvantageous inequality. Those with increased social anxiety were indeed more likely to coordinate with the high bidding agent and thus comply with disadvantageous inequality ($t = 4.76$, $p < .001$; Fig. 1F). Further, this effect was selective, there was no effect of social anxiety on coordination with the low bidding agent. Our behavioral results reveal two key findings. First, people spontaneously coordinate with different kinds of partners that play different strategies, however, coercive strategies that force an individual to accept lower payouts drastically reduce coordination. However, individuals substantially vary in how they respond to disadvantageous conventions. Specifically, social anxiety was associated with increased compliance with disadvantageous inequality.

Computational Modeling

B-RL. Differences in coordination rates with high and low bidding strategies could indicate asymmetries in learning when the high versus low offer results in a payout. In contrast, differences in coordination rates could be independent of learning and instead reflect utility-based social preferences—one might perfectly learn when a low bidding strategy maximizes monetary returns but refuse to accept this convention. To tease out the individual contributions of learning versus social preferences, we used a B-RL model to dissociate the relative influence of these components on choice. Our base model, M1, was adapted from a similar B-RL model developed in prior work to model trust choices (Lamba et al., 2020). We assume approximately Bayesian beliefs about selecting the high versus low offer, represented probabilistically and updated with observed trial outcomes.

$$p\left(\theta | outcome_{t1} \ldots outcome_{tn}\right) = \frac{p\left(outcome \ldots outcome_{tn} | \theta\right) \times p(\theta)}{p(outcome_{t1} \ldots outcome_{tn})}$$

Given that the likelihood of any observation (the partner agent will select the high or low bid) is Bernoulli, we modeled the belief about this probability, $p(\theta)$, as the conjugate, a beta distribution. Each time the high bid strategy resulted in coordination, alpha was incremented +1, and beta was incremented +1 for the low bid strategy. Alpha and beta values were used to update the mean and variance of the posterior distribution at the end of each trial.

$$\mu_t = \left(\frac{\alpha}{\alpha+\beta}\right), \sigma_t^2 = \left(\frac{\alpha * \beta}{(\alpha+\beta)^2 * (\alpha+\beta+1)}\right)$$

Each agent was individually modeled in separate distributions and all priors were initialized using the beta distribution conjugate prior (beta 1,1; i.e. uniform). Individual choice probabilities were modeled through a softmax logistic function, comparing the mean value of the posterior distribution where ζ and ψ are inverse temperature and choice bias parameters, respectively.

$$p\left(high\ bid\right) = \frac{e^{\zeta * \mu_{jt}}}{e^{\zeta * \mu_{jt}} + e^{\zeta * \psi}}$$
$$p\left(low\ bid\right) = 1 - p\left(high\ bid\right)$$

The effective learning rate from trial outcome history was captured through the decay parameter, γ. When observed outcomes are not decayed ($\gamma = 1$), the beta distribution is maximally updated with each observation through the alpha and beta hyperparameters. Lower decay parameters ($\gamma = 0$) inject noise into the posterior distribution, mixing the posterior distribution with the uniform prior and "forgetting" previously observed outcomes. We implemented two decay parameters to capture potential asymmetries in learning from low versus high bid strategies. Here, $\gamma 0_{high}$, can reinforce or decay the action of selecting the lower offer when it results in successful coordination, whereas $\gamma 0_{low}$, can reinforce or decay selecting the higher offer.

Social Utility B-RL. The social utility-based Bayesian RL model, M2, augments M1 with additional utility parameters that allow choice policies (i.e., strategies) to be further modulated by social preferences. Based on social utility equations proposed by Fehr and Schmidt (1999), we calculated the advantageous and disadvantageous inequality in payoffs on each trial, u_{adv} and u_{disadv}, as the difference between participant and partner earnings. We then estimated the extent to which advantageous and disadvantageous inequality aversion (AIA

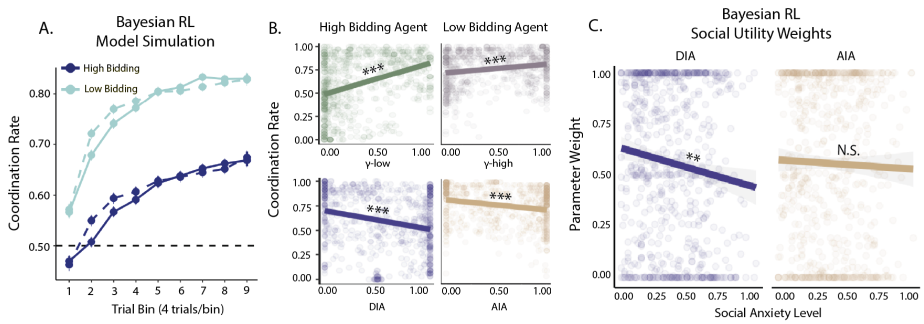

Figure 2: **A.** MLE simulations from the B-RL model. Group aggregated model simulations from the MLE-optimized parameters are plotted as the dotted lines. Empirical data is plotted as solid lines. **B.** Model validation. Optimized model parameters show the predicted association with task performance. **C.** Selective effect of social anxiety on DIA. Social anxiety is associated with reduced expression of DIA but not with AIA, a pattern which results in increased compliance with disadvantageous inequality.

and DIA, respectively) influence posterior beliefs by decaying the complementary policy. Through γ and social utility parameters, participants can form the complementary coordination strategy with their partner or resists inequality by miscoordinating.

$$\gamma_{high} = \gamma 0_{high} - AIA * u_{adv}\ ,\ \ \gamma_{low} = \gamma 0_{low} - DIA * u_{disadv}$$
$$\alpha_{t+1} = \alpha_t * \gamma_{high}\ ,\quad \beta_{t+1} = \beta_t * \gamma_{low}$$

Model performance. Our baseline B-RL model (M1), and social utility-based B-RL (M2) model fit the data approximately equally in terms of mean AIC, though Bayesian model comparison favored M1 (pxp > .99) which had 2 fewer parameters. However, the model fit advantage of M1 vs. M2 depended on whether individual AIA ($t = 5.67$, $p < .001$) or DIA ($t = 5.62$, $p < .001$) weights were greater than 0. Maximum likelihood estimation (MLE) simulations from the social utility B-RL model recapitulates mean coordination rates with each partner agent (Fig. 2A).

Effect of decay and social utility on coordination. The decay terms in the model, $\gamma 0_{low}$ and $\gamma 0_{high}$, function as the effective learning rate, governing how much posterior beliefs are updated from observing the partner agent's actions. As expected, higher γ parameters (i.e., increased updating from each observation) were associated with increased coordination with the relevant agent identity: $\gamma 0_{low}$ increased coordination with the high bidding agent ($t = 8.71$, $p < .001$) and $\gamma 0_{high}$ increased coordination with the low bidding agent ($t = 3.71$, $p < .001$; Fig. 2B). We also observed a significant difference in γ for each strategy, such that participants learned to take the higher offer more quickly than they learned to take the lower offer (mean $\gamma 0_{high} = 0.51$, mean $\gamma 0_{low} = 0.39$, $t = 7.74$, $p < .001$). These results indicate that learning differs across strategies, favoring learning to take higher payouts over lower payouts.

Next, we tested the extent to which social utility weights, AIA and DIA, affected coordination rates. As expected, greater weights on DIA decreased coordination with the high bidding agent ($t = -8.90$, $p < .001$), whereas greater weights on AIA decreased coordination with the low bidding agent ($t = -7.28$; $p < .001$; Fig. 2B). At the group level, AIA and DIA weights did not significantly differ. These findings suggest that updating biases towards the high and low policy, and utility-based preferences both distinctly drive coordination strategies.

Trait-based individual differences. The B-RL model also allowed us to test whether the effect of social anxiety on coordination was related to learning or expression of social preferences. Participants with increased social anxiety showed lower weights on DIA ($t = -3.00$, $p < .01$), but social anxiety was not related to AIA. Social anxiety was not significantly related to high or low decay parameters. These findings reveal that social anxiety increases compliance with disadvantageous outcomes through reduced expression of social preferences, rather than through learning differences.

Bayesian Causal Model. The Bayesian causal model predicts a participant's choice on every trial by emulating how the participant reasons about the expected utility associated with each option. Expected utility is a function of the simulated player's preferences, which are constant across trials, and beliefs, which update over time based on the player's experience. The utility of an outcome is $U(a, a_{other}; \boldsymbol{\omega}) = \omega^{\$} Money - \omega^{\mathrm{DI}} DI - \omega^{\mathrm{AI}} AI$, where a and a_{other} are the actions of the player and the partner, respectively, $\boldsymbol{\omega}$ is a vector of the player's preferences[2], and $Money$, DI, AI are the objective features of the outcome (how much money the player received, how much less they received than their partner, and how much more they received than their partner, respectively), which are functions of a and a_{other}. To predict a partner's choice, simulated players reason about the partner's expected utility-based decisions (Baker et al., 2009, 2017; Jern et al., 2017). Integrating over their beliefs about their partner's

[2]Preference weights are constrained to the interval $(0,1)^3$ by a logit-normal prior.

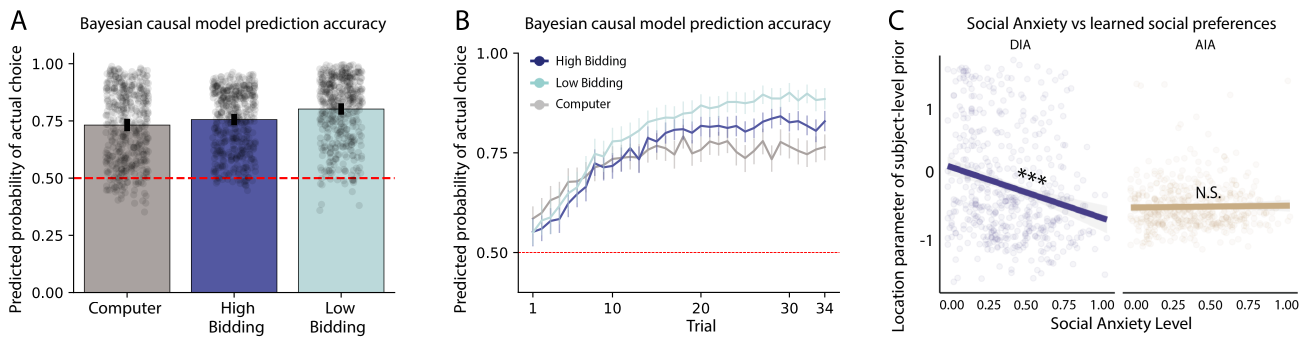

Figure 3: Bayesian causal model. **A.** Predicted probability of participants' actual choices, stratified by partner. Points give the average probability for each participant across all trials. Bars give the average across all participants and trials. Error bars are 95% bs CI. Chance level prediction shown in red. **B.** Predicted probability of participants' actual choices by trial, averaged across participants. **C.** Association between self-reported social anxiety and social preferences learned by the Bayesian causal model.

mental contents and likely decisions, players estimate the expected utility of the available actions. A simulated player's decisions follow probabilistically from the softmax of the expected utilities. When the partner's action is revealed, the simulated player inverts their mental model of the partner's mind to infer what latent mental states led to the observed action and updates their beliefs given this new evidence.

Simulated players' beliefs are hierarchically structured; players represent the preferences and rationality of partners at multiple levels of abstraction. For instance, a simulated player's prior belief about partner j's disadvantageous inequality aversion is represented as the distribution $\omega_j^{DI} \sim \text{LogitNormal}(\mu^{DI}, \tau^{DI})$. The hyperparameters μ^{DI} and τ^{DI} are themselves given hyperpriors, representing a player's beliefs about the population of partners. Beliefs are updated according to Bayes' rule. Based on the previous interactions with partner j, the player's posterior belief about the partner's preferences is given by

$$\mathrm{P}\big(\boldsymbol{\omega}_j, \beta_j \mid \mathbf{D}_j, \boldsymbol{\mu}, \boldsymbol{\tau}\big) \propto \Big[\prod_{t=1}^{T_j} \mathrm{P}\big(a_{j,t} \mid \boldsymbol{\omega}_j, \beta_j\big)\Big]\, \mathrm{P}\big(\boldsymbol{\omega}_j, \beta_j \mid \boldsymbol{\mu}, \boldsymbol{\tau}\big),$$

where $\mathbf{D}_j$ is the observed actions of partner j up to the present trial, T_j. Based on the history of interactions with all partners ($\mathbf{D} = \bigcup \mathbf{D}_j$), the posterior belief about the hyperparameters is given by

$$\mathrm{P}\big(\boldsymbol{\mu}, \boldsymbol{\tau} \mid \mathbf{D}\big) \propto \mathrm{P}\big(\boldsymbol{\mu}, \boldsymbol{\tau}\big) \prod_{j=1}^{J} \int \mathrm{P}\big(\mathbf{D}_j \mid \boldsymbol{\omega}_j, \beta_j\big)\, \mathrm{P}\big(\boldsymbol{\omega}_j, \beta_j \mid \boldsymbol{\mu}, \boldsymbol{\tau}\big)\; d\boldsymbol{\omega}_j\, d\beta_j$$

This hierarchical structure offers a cognitively natural way to model how participants form beliefs about multiple partners, simultaneously learning about individuals and making inductive inferences about people's decision-making in general (see Gelman, 2014; Tenenbaum et al., 2011; Zhi-Xuan et al., 2022). The hierarchical Bayesian model uses every interaction to update beliefs about all partners according to the strength of the evidence, with the hyperpriors modulating transfer learning.

Bayesian causal model results. The model predicts the probability of a participant's choice in every trial. Comparing these predictions to participants' actual behavior shows that the model reliably predicts the empirically observed choice to be more likely than the alternative action (i.e., the predicted probability of the observed choice is > 0.5). Fig. 3A shows the predicted probability of each participant's actual choices, averaged across all trials. In early rounds, when there is little information about each partner's decision process, the model performs near chance at predicting participants' choices (Fig. 3B). As evidence accumulates, the model emulates how participants learn about their partners, and how they incorporate this information, along with their own preferences and rationality, to choose actions that noisily maximize their expected utilities. We fit individual priors over each participant's preferences and rationality via gradient descent. In convergence with the results of the B-RL model, comparing the inferred preferences to participants' self-reported social anxiety reveals that participants' inferred aversion to disadvantageous inequality is negatively correlated with social anxiety ($t = -5.18$, $p < .001$), while advantageous inequality aversion is not ($t = 1.84$, $p > .05$, n.s.) (Fig. 3C)[3]. The model also shows positive associations between social anxiety and participants' preference for monetary reward (irrespective of equity; $t = 4.54$, $p < .001$) and their rationality ($t = 2.14$, $p < .05$).

Conclusions and Future Direction. In sum, our work advances knowledge on how humans navigate the complexities of asymmetric coordination. We show that learning dynamics and social utility jointly inform how people respond to advantageous and disadvantageous conventions. Further, social anxiety increases compliance with disadvantageous roles, which our computational models link to reduced expressed of social preferences rather than learning biases. In future work, we aim to investigate how neural representations of partner strategies and individual social preferences shape strategic coordination.

[3] This plot shows the location parameter of the logit-normal prior. Unlike the preference weight, the location parameter is unbounded.

Acknowledgements

We would like to thank the anonymous reviewers for helpful feedback and comments. We would also like to thank Kartik Chandra for adapting the probabilistic programming language `memo` (Chandra et al., 2025) to make it possible to implement hierarchical belief updating. AL is supported by a fellowship from the Simons Center for the Social Brain at MIT. SDH is supported by a fellowship from the William H. Neukom Institute for Computational Science.

References

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 598. https://doi.org/10.1038/s41562-017-0064

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

Carleton, R. N., Collimore, K. C., & Asmundson, G. J. (2007). Social anxiety and fear of negative evaluation: Construct validity of the BFNE-II. *Journal of Anxiety Disorders*, *21*(1), 131–141. https://doi.org/10.1016/j.janxdis.2006.03.010

Chandra, K., Chen, T., Tenenbaum, J., & Ragan-Kelley, J. (2025, January 9). *A Domain-Specific Probabilistic Programming Language for Reasoning About Reasoning (or: A memo on memo)*. https://doi.org/10.31234/osf.io/pt863

Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(4), 429–453. https://doi.org/10.3758/CABN.8.4.429

Fehr, E., Fischbacher, U., & Kosfeld, M. (2005). Neuroeconomic Foundations of Trust and Social Preferences: Initial Evidence. *American Economic Review*, *95*(2), 346–351. https://doi.org/10.1257/000282805774669736

Fehr, E., & Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Gelman, A. (2014). *Bayesian data analysis* (Third edition). CRC Press.

Houlihan, S. D., Kleiman-Weiner, M., Hewitt, L. B., Tenenbaum, J. B., & Saxe, R. (2023). Emotion prediction as computation over a generative theory of mind. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *381*(2251), 20220047. https://doi.org/10.1098/rsta.2022.0047

Jern, A., Lucas, C. G., & Kemp, C. (2017). People learn other people's preferences through inverse decision-making. *Cognition*, *168*, 46–64. https://doi.org/10.1016/j.cognition.2017.06.017

Kleiman-Weiner, M., Ho, M. K., Austerweil, J. L., Littman, M. L., & Tenenbaum, J. B. (2016). Coordinate to cooperate or compete: Abstract goals and joint intentions in social interaction. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *38*, 1679–1684.

Kleiman-Weiner, M., Tenenbaum, J. B., & Zhou, P. (2018). Non-parametric Bayesian inference of strategies in repeated games. *The Econometrics Journal*, *21*(3), 298–315. https://doi.org/10.1111/ectj.12112

Kocovski, N. L., & Endler, N. S. (2000). Social anxiety, self-regulation, and fear of negative evaluation. *European Journal of Personality*, *14*(4), 347–358. https://doi.org/10.1002/1099-0984(200007/08)14:4<347::AID-PER381>3.0.CO;2-7

Lamba, A., Frank, M. J., & FeldmanHall, O. (2020). Anxiety Impedes Adaptive Social Learning Under Uncertainty. *Psychological Science*, *31*(5), 592–603. https://doi.org/10.1177/0956797620910993

Lamba, A., Frank, M. J., & FeldmanHall, O. (2024). Keeping an Eye Out for Change: Anxiety Disrupts Adaptive Resolution of Policy Uncertainty. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *9*(11), 1188–1198. https://doi.org/10.1016/j.bpsc.2024.07.015

Lamba, A., Nassar, M. R., & FeldmanHall, O. (2023). Prefrontal cortex state representations shape human credit assignment (T. Kahnt, C. Büchel, & R. P. Costa, Eds.). *eLife*, *12*, e84888. https://doi.org/10.7554/eLife.84888

Leitmann, G. (1978). On generalized Stackelberg strategies. *Journal of Optimization Theory and Applications*, *26*(4), 637–643. https://doi.org/10.1007/BF00933155

Peters, L. (2000). Discriminant validity of the Social Phobia and Anxiety Inventory (SPAI), the Social Phobia Scale (SPS) and the Social Interaction Anxiety Scale (SIAS). *Behaviour Research and Therapy*, *38*(9), 943–950. https://doi.org/10.1016/S0005-7967(99)00131-X

Sákovics, J., & Steiner, J. (2012). Who Matters in Coordination Problems? *American Economic Review*, *102*(7), 3439–3461. https://doi.org/10.1257/aer.102.7.3439

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. *Science*, *331*(6022), 1279–1285. https://doi.org/10.1126/science.1192788

Yu, D., & Thompson, B. (2024, June 2). *People balance joint reward, fairness and complexity to develop social norms in a two-player game*. https://doi.org/10.31234/osf.io/qe6bp

Zhi-Xuan, T., Gothoskar, N., Pollok, F., Gutfreund, D., Tenenbaum, J. B., & Mansinghka, V. K. (2022, August 4). *Solving the Baby Intuitions Benchmark with a Hierarchically Bayesian Theory of Mind* (1). arXiv: 2208.02914 [cs]. https://doi.org/10.48550/ARXIV.2208.02914