# UC Merced

## Proceedings of the Annual Meeting of the Cognitive Science Society

### Title

Joint Action and Reward-Seeking in a Social Probability-Learning Task

### Permalink

https://escholarship.org/uc/item/9584n9vp

### Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

### Authors

Yosef, Bar

Deak, Gedeon O

### Publication Date

2025

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

# Joint Action and Reward-Seeking in a Social Probability-Learning Task

**Bar Yosef (baryos@ucla.edu)**
Department of Neurology,
University of California, Los Angeles

**Gedeon O. Deák (gdeak@ucsd.edu)**
Department of Cognitive Science,
University of California, San Diego

**Abstract**

Despite the prevalence of social learning in humans, the cognitive mechanisms underlying social transmission of behavior are not fully understood. We refine and expand a recently-developed paradigm, the Social Multi-Armed Bandit (SMAB) task, to systematically manipulate social information and measure its effects on human decision-making during predictive learning. We compared a dyadic task, in which social influence is maladaptive, to a control task in which participants receive no task-related social information. We found that misleading social information resulted in more maladaptive choices in the Dyadic than the Control condition, confirming findings from a prior study (Adrian et al., 2019). Although the maladaptive Dyadic social effect attenuated across trials and between two "games" (100 trials each), correlations between partners' higher-order response patterns persisted across games. These correlated response patterns between individuals within dyads suggest a tendency to emulate higher-order patterns (i.e., heuristics or strategies). The results imply that adults sometimes emulate decision strategies even when the outcome is disadvantageous. They also suggest that social learning may be reflected in higher-order response patterns even after people learn that imitating specific actions is maladaptive.

**Keywords:** imitation; misinformation; probability learning; reward learning; social learning; decision-making

## Introduction

Decision-making and action selection by humans often occurs in a social context (Rilling & Sanfey, 2011). Our decisions are guided by predictions informed by contingencies learned from prior experience (Schultz et al., 1997). A general question is how experiences alter predictive representations and action decisions. These experiences typically include social interactions (Whiten & van de Waal, 2018). Social information can boost learning if individuals represent others' actions as they plan their own actions (Iacoboni, 2005; Heyes, 2012). Moreover, the values of both social and nonsocial outcomes might be encoded in common neural representations (Nicolle et al., 2012; Ruff & Fehr, 2014). However, the mechanisms by which social observations guide prediction and action selection have yet to be fully characterized due to the inherent complexity of social dynamics, as well as limited integration across disciplines (Olsson et al., 2020).

Effects of social information, including contingency patterns, on learning and decision-making can be seen even in infancy and childhood (e.g., Chang & Deák, 2020; Plate et al., 2019; Sebanz et al., 2006). In children and adults, joint action involves social information sharing during goal-directed activity (e.g., Kourtis et al., 2019).

Social information can influence decision-making under uncertainty. This can be explored by modifying probabilistic learning paradigms such as the multi-armed bandit (MAB; Daw et al., 2006). MAB tasks present multiple options ("arms") with reward probabilities unknown to the decision-maker. Participants choose actions that either explore the environment, or exploit (i.e., repeat) previously-rewarded actions. In stable environments with discernible optimal choices, individuals typically converge toward an optimal, or maximizing strategy, favoring the historically most-rewarding arm (Cohen et al., 2007; Speekenbrink & Konstantinidis, 2015).

We refined and expanded a recently-developed paradigm, the Social Multi-Armed Bandit (SMAB; Adrian et al., 2019), that systematically introduces social information and quantifies its effects on decisions. In a Dyadic condition, two participants alternate their choices, with outcomes determined by different hidden reward contingencies that are manipulated to make direct imitation (i.e., repetition of a partner's last response, or "copying the form of an action"; Whiten et al., 2009) unfavorable. In this paradigm, the other participant's actions become a potential source of interference. This allows us to study how potentially misleading social information affects decision-making.

In probabilistic learning in social environments, early exploration may entail copying another person's actions, ostensibly as a learning strategy (Kendal et al., 2004). Unexplored strategies might be gleaned from observing others' actions (Cohen et al., 2007). Previous social probability learning studies show that socially gleaned information can improve individual performance. For example, Schulze and Newell (2016) found that disadvantageous matching responses were generally avoided when participants made decisions collaboratively, compared to a single-player condition. Similarly, Witt et al. (2024) showed that participants flexibly integrated social information to guide decisions under uncertainty, leading to enhanced outcomes despite individual differences in reward valuation. Social influence can, however, be either adaptive or maladaptive, depending on factors such as the appropriateness of the other person's actions, the goal of the task, task constraints, and participants' respective valuations of outcomes. In the current SMAB task, it was



In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

disadvantageous for participants to imitate their partner because the partner had incompatible reward contingencies. This potentially conflict-inducing paradigm was designed (Adrian et al., 2019) to reveal the limits of human social learning as participants accrue evidence that imitation is maladaptive.

A prior study (Adrian et al., 2019) found reduced maximizing in the Dyadic condition compared to a single-player control version of the task. This was presumably due to the conflicting feedback from the player's own decision-outcome contingencies and their partner's action-outcome contingencies. Specifically, participants' highest-reward-probability option was their partner's lowest-reward option, and vice versa. In this situation, any persistent copying of the partner's previous action might reflect habitual imitation (Heyes, 2011). A third, intermediate-reward-probability option was identical for both partners, making the "switched" reward contingencies less obvious. The present study expands on this paradigm with several changes: First, an improved Control condition was similar to the Dyadic condition except that a confederate (rather than a naive partner) made an invariant, non-contingent action in between the participant's decisions. This controlled for the presence of another person (Golland et al., 2015), the pacing of trials, and seeing another hand touch the screen in between trials. Thus the Control condition is similar to the Dyadic condition except that the "partner" makes no game-relevant decisions and receives no outcomes. The present study also uses a between-subjects design, improved stimuli, a second block of trials, and analyses of decision sequences over all trials, including similarity of sequences within dyads.

We hypothesized that suboptimal performance would be greatest early in the task (Schulze & Newell, 2016). Given the uncertainty of the paradigm, participants might initially rely on observing their partner's actions and outcomes, leading to fewer rewards (Sebanz et al., 2006). As participants learn their own underlying response contingencies, they might increasingly ignore social information to optimize outcomes. Aside from directly imitating a partner's choices, partners might demonstrate social influence by adopting parallel response distributions. Similarities in higher-order patterns could suggest emulation: reproducing another's strategies or policies without copying specific actions (Tomasello, 1993; Whiten et al., 2009). In reward learning paradigms like SMAB, emulation would entail modeling the partner's action-contingency patterns, and selecting similarly-distributed actions for a similar outcome pattern (Ng & Russell, 2000). Whereas imitation in the SMAB would entail choosing the identical arm as the partner, emulation would entail making similar distributions of choices as the partner, across non-identical arms that yield commensurate reward probabilities. Thus, we measure emulation by examining within-dyad similarity in strategy use. Emulation is not inherently adaptive or maladaptive in the SMAB—that depends on the partner's strategy (Whiten et al., 2009). We hypothesized that emulation would occur to some degree within dyads.

In sum, we manipulated social information and measured its effects on decisions in a social MAB task wherein social information would be misleading. We compared the dyadic task to an individual control task. We assessed learning within and between games by examining decisions across trials. We also examined dyadic participants' tendency to match their partner's specific decisions ("imitation") or to converge on similar strategies ("emulation").

## Materials and Methods

### Participants

Participants were 61 right-handed, English-fluent undergraduates (Dyadic condition: n = 32, Control: n = 29) at UC-San Diego. Dyadic participants (18 f, 13 m, 1 nonbinary) averaged 20.7 years (range: 17.7-26.6), and self-identified as 17 Asian, 2 Black, 4 Hispanic/Latinx, 4 White, 4 multi-racial, and 1 "other." Control participants (17 f, 12 m) averaged 19.7 years old (range: 17.5-24.1) and self-identified as 13 Asian, 8 Hispanic/Latinx, 4 White, 3 multi-racial, and 1 "other." Three control participants were excluded from Game 1 and three (1 control; 1 dyad) from Game 2 due to data recording problems. Each participant received course credit plus a small payment (i.e., $0.05 USD/reward, mean total = $7.07/participant).

### Materials

The Social Multi Arm-Bandit (SMAB) task was displayed on an 58.5 x 32.5 cm ELO touchscreen monitor mounted in a table. The game was programmed in Python (tkinter; Lundh, 1999). The three arms were represented as uniquely shaped "ice holes" (Figure 1).

### Procedure

The task was presented as an "ice fishing game." Three "ice holes" were assigned hidden reward probabilities of 30% (low arm), 60% (medium arm), and 90% (high arm). Participants were told to earn as many points as possible by selecting one arm per trial and that they would earn $0.05 for each point, with no further information. In each trial participants chose an arm by touching an ice hole, prompting either a reward of one point (green fish) or non-reward (red "X": Figure 1A-B). Each player's cumulative reward count was displayed on the monitor as their respective scores. Each participant completed two rounds of 100 trials (i.e., Games 1 and 2).

In the Dyadic condition, participants were tested in pairs in a sound-attenuated, isolated testing room seated across from each other on either side of a touchscreen table. Participants alternated with their partner in choosing ice holes (Figure 1A). The arms with reward probabilities of 30% and 90% were reversed between paired participants (Figure 1C), thereby imposing potentially maladaptive social interference. Control condition sessions occurred in the same setting and were designed to control for the presence of another person (Golland et al., 2015), the pacing of trials, and seeing another

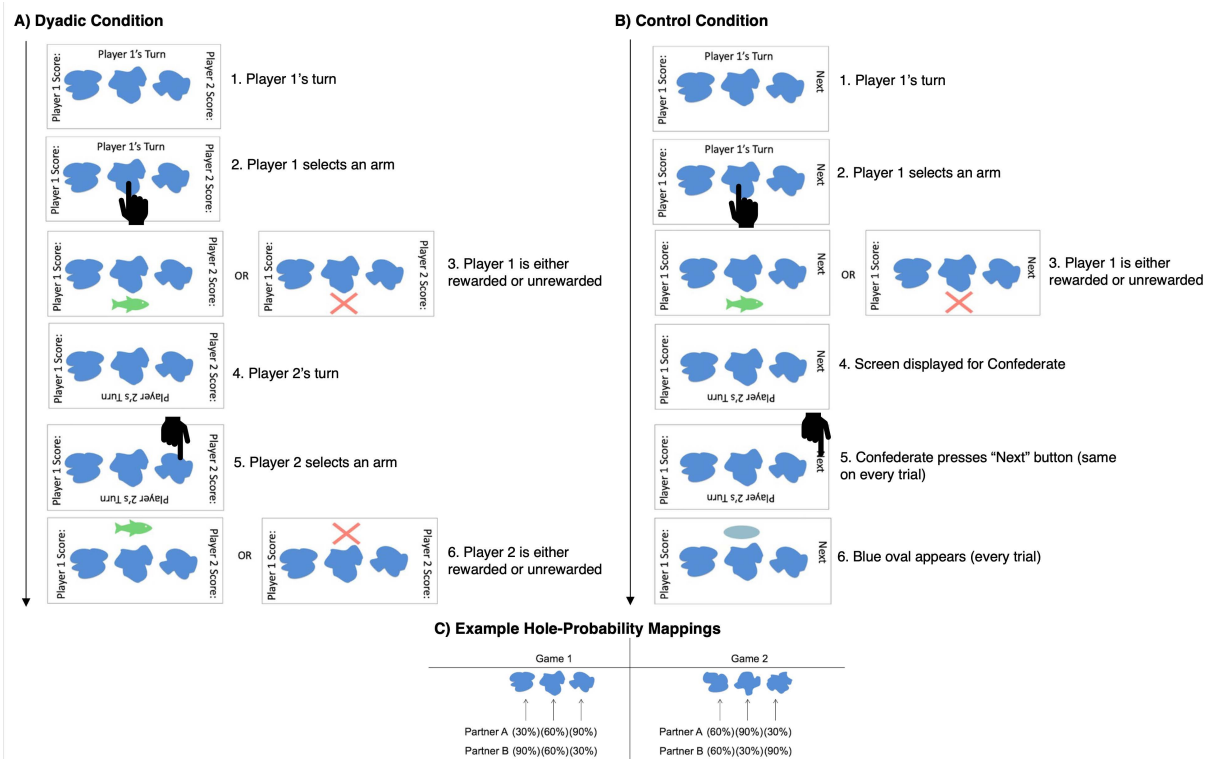



Figure 1: Ice fishing task trials. In both conditions a quasi-random trial sequence (A, B) repeats 100 times per game (total=200 trials). A green fish indicates a reward (one point); a red "X" indicates no reward. *Dyadic (Social) Condition (A):* Two players alternate selections. *Control Condition (B):* A confederate presses the Next button between trials, prompting a blue oval. *Arm Shapes and Probability Mappings (C):* Shapes are randomly assigned to "arms" (left, middle, right) with different shapes in Games 1 and 2. Reward probabilities are quasi-randomly assigned in Game 1; these remain the same in Game 2 ("Same" condition) or are re-assigned ("Different" condition). In the Dyadic condition, 30% and 90% arm-assignments are reversed for Players 1 and 2.

person's hand touch the screen in between trials. Participants completed trials identical to the Dyadic condition, but in between each trial a confederate (seated across from the participant, like Dyadic partners) pressed a "Next" button, and a blue oval appeared (Figure 1B). That is, unlike the Dyadic condition, the confederate's actions were game-irrelevant and had no differential outcomes (i.e., they always pressed the "Next" button).

The shapes and reward probability of ice holes were randomized. The shapes differed in Games 1 and 2. The reward probabilities changed locations in Game 2 for half of participants—the "Different" condition—but not for participants assigned to the other, "Same" condition (Figure 1C). This manipulation (which yielded no significant differences) was intended to examine the effect of changing reward contingencies (Cohen et al., 2007).

All procedures were approved by the UCSD IRB.

## Data Analysis

The course of choices within and across games was analyzed using growth curve analysis (Mirman, 2014) of cumulative proportions of high (90%) reward-probability and of low (30%) reward-probability choices. The former suggests a maximizing strategy, and the latter suggests maladaptive imitation of social information. Linear mixed-effects (LME) models (suitable for repeated measures [RM], data) were fitted using lme4 1.1.29 (Bates et al., 2015) in R 3.6.3.

The model included fixed effects for trial and game (both within-subjects), condition (between: Dyadic; Control), mapping (between: Same; Different), and their interactions, as shown in equation (1). The models were fitted using REML and a maximal random effects structure (Barr et al., 2013; Lüdecke, 2021; Luke, 2017) with a random intercept for participant, a linear effect of participant over trials (i.e., random slope), and random participant-by-game effects. To handle non-convergence, the trial variable was rescaled and the random effects structure was simplified. Fixed effect coefficients' were evaluated using Satterthwaite's (1946) approximated dfs.

1 Cumulative % of choices to arm (30/90) = trial_scaled * condition * game * mapping + (trial_scaled * game|participant )

Because the mapping effect and all interactions with mapping were non-significant in both models, the "Same"

and "Different" conditions were pooled for subsequent analyses.

In addition to examining choice behavior across trials, we examined choice behavior summarized over entire games using overall maximization rate. This was calculated for each game for a given participant as the proportion of choices of the high arm across all test trials. We examined the similarity of strategy use between Dyadic condition partners (i.e., emulation) by correlating maximization rates between partners within dyads.

## Results

We compared choice trajectories between Dyadic and Control conditions across two games. A positive trajectory of high-reward-arm choices and negative trajectory of low-reward-arm-choices across trials would indicate adaptive learning. A higher overall proportion of low-arm choices in the Dyadic condition than the Control condition would indicate a maladaptive social influence. Growth curve models of high- and low-arm choices largely confirm both results: a learning trend, and a maladaptive effect of misleading social information which, as explained below, was transitory.

High-arm choices increased and low-arm choices decreased over trials in both conditions (Figure 2, Table 1), confirming the predicted learning effect: participants shifted towards adaptive choices. Specifically, the low-arm initial mean choices averaged 21.7% (95% CI [17.9, 25.5], $p < .001$, $df = 59$), with a significant decrease of 2.52% per trial (95% CI [-4.2, -0.9], $p = .003$, $df = 57$). Conversely, high probability (90%) arm initial mean choices averaged 47.7% (95% CI [41.0, 54.4], $p < .001$, $df = 59$), with a significant increase of 3.50% per trial (95% CI [1.0, 6.0], $p = .006$, $df = 57$).

Trajectories were not equivalent between groups, however. Dyadic participants made significantly (8.18%) more low-arm choices, than Control participants (95% CI [3.0, 13.4], $p = .003$, $df = 58$), and marginally fewer (-8.95%) high-arm choices (95% CI [-18.0, 0.1], $p = .053$, $df = 58$).

There was a significant condition-by-game interaction in low-arm choices due to a larger (11.16%) decline from Game 1 to Game 2 in the Dyadic than the Control conditions (95% CI [-20.5, -1.8], $p = .020$, $df = 59$). This underscores the impact of condition over time: Dyadic participants chose the low arm more especially in the initial game (i.e., Dyadic 1 trend in Fig. 2).

We also examined choice behavior summarized over entire games. The similarity of Dyadic partners' strategy-use (a possible index of emulation) was quantified as correlations in maximization rates (Figure 3C,D). These were moderate to high: Game 1 $r = .778$, $p < .001$; Game 2 $r = .613$, $p = .015$. Notably, because partners' responses were independent and non-contingent, we could have found null or even negative correlations. See the supplement for additional analyses: https://osf.io/pdgbs/?view_only=334bc22ab8ba437380b080ef4e9aee80

## Discussion

We used the Social Multi-Armed Bandit (SMAB), to manipulate social information and measure its effects on predictive learning for decision-making. Results show increased maladaptive behavior in a dyadic task when the partner's actions and outcomes are misleading, indicated by more low-reward arm selections compared to an individual control condition. This was presumably due to seeing the partner's frequent rewards from choosing that arm. This finding generally replicates Adrian et al. (2019)'s behavioral findings of more maximization in a single-player block of SMAB than in a dyadic block (in a within-subjects design).

The current results show that this effect (i.e., more dyadic low-arm selections) attenuated with experience, decreasing in Game 2. This suggests that participants eventually learned to reduce maladaptive imitation of misleading observed actions. A social influence was also seen in the above-chance similarity of strategies adopted by individuals within a dyad; that is, an "emulation" effect.

The attenuation of maladaptive social influence between games supports characterizations of human social learning as flexible and multidimensional (Toyokawa et al., 2019). This study adds to a growing body of recent findings that shed light on the high degree of flexibility present in human social learning (Charpentier et al., 2020; Witt et al., 2024). Although social learning is endemic, it is not always beneficial, and selectivity of social learning is important for

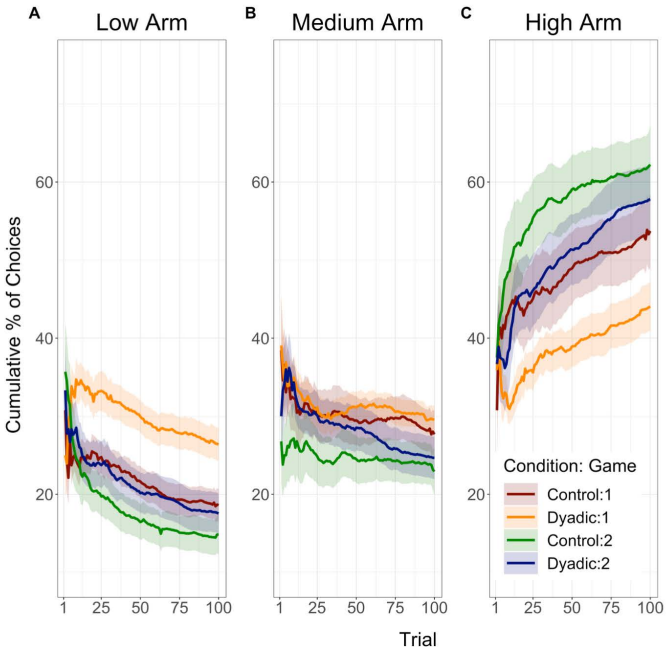



Figure 2: Dyadic vs. control group cumulative choices of each arm. Mean cumulative response percentages for each arm: (A) low (30% reward-probability) arm, (B) medium (60%) arm, and (C) high (90%) arm (i.e., maximization rate). Cumulative choices for each arm are shown separately for Games 1 and 2 in each condition. Error clouds: SEmean.

Table 1. Results of Growth Curve Models for the Cumulative Percentage of Low Probability Arm and High Probability Arm Choices

| Variable | Low Probability (30%) Arm | | | | High Probability (90%) Arm | | | |
|---|---|---|---|---|---|---|---|---|
| | *Estimates* | *CI* | *p* | *df* | *Estimates* | *CI* | *p* | *df* |
| (Intercept) | 21.69 | [17.9, 25.5] | **< .001** | 59 | 47.67 | [41.0, 54.4] | **< .001** | 59 |
| Trial | -2.52 | [-4.2, -0.9] | **.003** | 57 | 3.50 | [1.0, 6.0] | **.006** | 57 |
| Condition | 8.18 | [3.0, 13.4] | **.003** | 58 | -8.95 | [-18.0, 0.1] | .053 † | 58 |
| Game | 0.83 | [-6.4, -8.1] | .818 | 60 | 7.47 | [-4.1, 19.1] | .203 | 60 |
| Mapping | -6.69 | [-15.7, 2.3] | .142 | 56 | 2.83 | [-11.6, 17.2] | .695 | 55 |
| Trial * Condition | 0.55 | [-1.7, 2.8] | .619 | 57 | -0.30 | [-3.6, 3.0] | .856 | 57 |
| Trial * Game | -3.76 | [-7.3, -0.2] | **.039** | 64 | 2.15 | [-3.3, 7.6] | .430 | 70 |
| Condition * Game | -11.16 | [-20.5, -1.8] | **.020** | 59 | 5.54 | [-9.5, 20.5] | .463 | 59 |
| Trial * Mapping | 3.90 | [-0.3, 8.1] | .068 † | 54 | -1.98 | [-8.0, 4.0] | .513 | 54 |
| Condition * Mapping | 8.88 | [-3.4, 21.2] | .154 | 55 | -3.23 | [-22.9, 16.4] | .742 | 54 |
| (Trial * Condition) * Game | 1.19 | [-3.5, 5.8] | .611 | 62 | 1.80 | [-5.3, 8.9] | .614 | 69 |
| (Trial * Condition) * Mapping | -1.12 | [-6.9, 4.7] | .700 | 54 | -1.13 | [-9.4, 7.2] | .785 | 54 |
| ICC | 0.76 | | | | 0.88 | | | |
| $N_{Participant}$ | 61 | | | | 61 | | | |
| Observations | 11570 | | | | 11570 | | | |
| Marginal $R^2$ /Conditional $R^2$ | 0.132 / 0.795 | | | | 0.109/ 0.891 | | | |

adaptive learning (Kendal et al., 2018). The social learning strategies (SLSs) concept defines the actors ('who'), behaviors ('what'), and circumstances ('when') that bias an individual's social learning tendencies (Kendal et al., 2018)—for example, strategies like 'copy when uncertain,' 'copy the majority,' and 'copy if better' (Laland, 2004). Due to discrepant reward contingencies between partners in SMAB, a copy when uncertain strategy might have biased participants to make suboptimal choices in Game 1. As participants learned the reward probabilities for their own choices, they adaptively reduced their use of social information later in the task.

The results also imply flexibility of imitation versus emulation. Although imitative responses attenuated across games, partners used social information by adopting similar response strategies (i.e., emulation; Tomasello, 1993)—specifically, choosing respective arms with analogous reward probabilities at similar rates, which entailed reversing their partner's selection-rates for two of the arms.

Emulation allows agents to modify behaviors to approximate another person's goals or strategies, whereas imitation entails transmission of specific actions (Whiten et al., 2009). Situational factors might make either emulation or imitation more appropriate (Williamson & Meltzoff, 2011). To our knowledge, this is among the first studies to find evidence of adaptive, probabilistic emulation of choice distributions in a predictive learning task where imitation is maladaptive. This suggests a tendency to adaptively utilize social information at multiple levels from affiliated individuals. That tendency might reduce potentially costly exploration (Witt et al., 2024), lower cognitive load when adopting strategies (Hinsz et al., 1997) and flexibly modulate individuals' responses to observed choice outcomes (Li et al., 2010). These findings echo recent work by Charpentier et al. (2020) who showed that individuals dynamically arbitrate between imitating observed actions and emulating goal-directed policies based on task structure and contextual demands. More broadly, the results suggest that social probability-learning tasks can illuminate social adaptation during learning at multiple levels—i.e., individual decisions, and strategic policies across multiple decisions.

Few studies have examined the dynamics of "live" dyadic social interactions in probability learning tasks, and some of those were virtual or "distance" studies. Such methods,

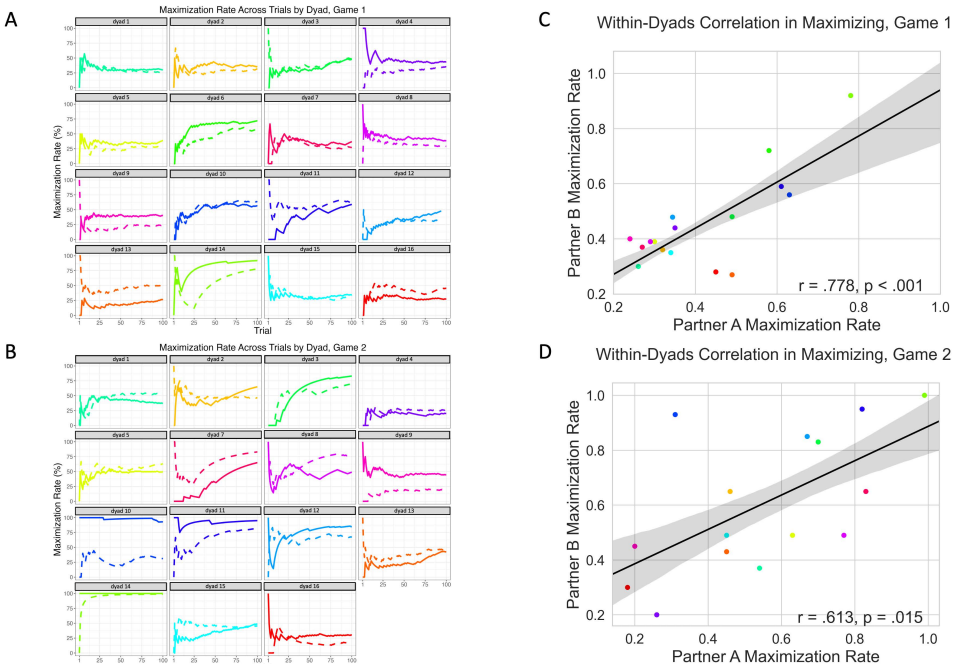



Figure 3: Trends and correlations in dyadic condition games. (A, B) Maximization changes across trials in Games 1 and 2 for each dyad, with solid/dotted lines for each participant. (C, D) Correlation of within-dyad maximization rates in Games 1 and 2, with r values as Pearson coefficients. Lines show linear regression; error clouds represent SEmean.

though convenient, limit the types of embodied social cues that can be studied (e.g., facial and vocal expression and other nonverbal behaviors; Walther et al., 2005). More generally, such methods preclude information carried by real-time, spatially contiguous interactions in shared physical environments (Ickes & Gonzalez, 1994). Future research could explore how physical versus virtual presence of others influences the use of social information in decision-making under uncertainty.

A limitation of this study is that we cannot specify the mechanisms driving the observed social effects. For example, we cannot pinpoint the exact mechanism of the transitory, maladaptive presumably-imitative effect. Similarly, strategy emulation may perhaps be due to a generalized, developmentally canalized orientation to social cues (Whiten et al., 2009). Prosociality bias has been reported in some contexts (Zaki & Mitchell, 2013), and motor synchrony, or simply shared attention, can enhance prosociality and action coordination (Valdesolo & DeSteno, 2011; Wolf et al., 2016). Further, vicarious prediction learning engages similar brain areas and computations as first-hand learning (Burke et al., 2010; Cooper et al., 2012; Ruff & Fehr, 2014). This suggests that we might automatically incorporate observed rewards into our representations of reward contingencies. Prosociality might also drive emulation if dyadic participants temper or increase maximization with a bias to maintain parity (i.e., reward equity) with their partner (e.g., Ruffle, 1998; Ulber et al., 2017). Alternatively, strategy emulation might be driven by participants conditionalizing their partner's rewarded actions versus their own rewarded actions, consistent with a copy-if-better strategic heuristic (Laland, 2004).

Future studies could modify the SMAB paradigm to specify the nature of emulation. For example, participants could be paired with confederates who adhere to either maximizing or to a less optimal (e.g., matching) strategy. If emulation is independent of the partner's strategy, it may reflect a robust cross-situational tendency; if emulation varies by strategy efficacy, it would suggest a conditional tendency, consistent with a "emulate-if-better" heuristic (Laland, 2004). SMAB could also be used to examine socially-informed predictive learning in specific populations, such as individuals with neuroatypicalities (e.g., Harlé et al., 2017) or children of different ages. For example, we could examine whether children's tendency to over-imitate in some contexts leads to prolonged imitation of a misleading social partner (Lyons et al., 2007). More generally, it would be useful to study the conditions that facilitate or discourage either specific action imitation, policy emulation, or both.

A strength of the current method is that the Control condition allowed us to more precisely specify the social influence on probabilistic learning. By using a confederate, we limited the difference between conditions to seeing a partner's game-relevant choices and outcomes.

In conclusion, we compared individual and dyadic versions of a probabilistic learning task in which imitation of a partner's rewarded actions was maladaptive. This social information temporarily increased maladaptive choices. Although this effect attenuated over time, within-dyad correlations in decision patterns, suggesting emulation, persisted. These results suggest that social influence on predictive learning is reflected in a flexible tendency to imitate as well as a tendency to emulate choice patterns.

## Acknowledgements

We thank Julia Adrian, Simone Froley, Dalin Guo, Zoe He, Siddharth, and Tzyy-Ping Jung for their insights and contributions, and Aihang Jin, Selena Lugosi, Jia-Hui Ma, Kexin Quan, Hannah Williams, Binghong Yu, Yisha Zhang, and Zhiyi Li for assistance. This work was supported by the UCSD Academic Senate. Finally, we thank the late Eduardo Macagno for his invaluable feedback on this project and the indelible impact he left on the authors.

## References

Adrian, J. A., Siddharth, S., Baquar, S. Z., Jung, T. & Deak, G. (2019). Decision-making in a social multi-armed bandit: Behavior, electrophysiology and pupillometry. Proceedings of the Annual Cognitive Science Society, 41, 1-5.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48.

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. Proceedings of the National Academy of Sciences, 107(32), 14431-14436.

Chang, L. M., & Deák, G. O. (2020). Adjacent and non-adjacent word contexts both predict age of acquisition of English words: A distributional corpus analysis of child-directed speech. Cognitive Science, 44(11), e12899.

Charpentier, C. J., Iigaya, K., & O'Doherty, J. P. (2020). A neuro-computational account of arbitration between choice imitation and goal emulation during human observational learning. Neuron, 106(4), 687-699.

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362(1481), 933–942.

Cooper, J. C., Dunne, S., Furey, T., & O'Doherty, J. P. (2012). Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. Journal of Cognitive Neuroscience, 24(1), 106–118. https://doi.org/10.1162/jocn_a_00114

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. Nature, 441(7095), 876–879. https://doi.org/10.1038/nature04766

Golland, Y., Arzouan, Y., Levit-Binnun, N. (2015). The mere co-presence: Synchronization of autonomic signals and emotional responses across co-present individuals not engaged in direct interaction. PLoS ONE, 10(5): e0125804.

Harlé, K. M., Guo, D., Zhang, S., Paulus, M. P., & Yu, A. J. (2017). Anhedonia and anxiety underlying depressive symptomatology have distinct effects on reward-based decision-making. PloS ONE, 12(10), e0186473.

Heyes, C. (2011). Automatic imitation. Psychological Bulletin, 137(3), 463-483.

Heyes C. (2012). What's social about social learning? Journal of Comparative Psychology, 126(2), 193–202. doi.org/10.1037/a0025180

Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. Psychological Bulletin, 121(1), 43.

Iacoboni, M. (2005). Neural mechanisms of imitation. Current Opinion in Neurobiology, 15(6), 632-637.

Ickes, W., & Gonzalez, R. (1994). " Social" cognition and social cognition: From the subjective to the intersubjective. Small Group Research, 25(2), 294-315.

Kendal, R.L., Coolen, I., Laland, K.N. (2004). The role of conformity in foraging when personal and social information conflict. Behavioral Ecology, 15(2), 269–277.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018). Social learning strategies: Bridge-building between fields. Trends in Cognitive Sciences, 22(7), 651–665.

Kourtis, D., Woźniak, M., Sebanz, N., & Knoblich, G. (2019). Evidence for we-representations during joint action planning. Neuropsychologia, 131, 73–83. doi.org/10.1016/j.neuropsychologia.2019.05.029

Laland, K. N. (2004). Social learning strategies. Learning and Behavior, 32(1), 4-14.

Li, P., Jia, S., Feng, T., Liu, Q., Suo, T., & Li, H. (2010). The influence of the diffusion of responsibility effect on outcome evaluations: electrophysiological evidence from an ERP study. Neuroimage, 52(4), 1727-1733.

Lüdecke, D. (2021). sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.9, https://CRAN.R-project.org/package=sjPlot.

Luke, S.G. (2017). Evaluating significance in linear mixed-effects models in R. Behavior Research Methods, 49, 1494–1502.

Lundh, F. (1999). An introduction to tkinter. URL: Www. Pythonware. Com/Library/Tkinter/Introduction/Index. Htm.

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

Mirman, D. (2014). Growth Curve Analysis and Visualization Using R (1st ed.). Chapman and Hall/CRC.

Ng, A. Y., & Russell, S. (2000, June). Algorithms for inverse reinforcement learning. In ICML (Vol. 1, p. 2).

Nicolle, A., Klein-Flügge, M. C., Hunt, L. T., Vlaev, I., Dolan, R. J., & Behrens, T. E. (2012). An agent independent axis for executed and modeled choice in medial prefrontal cortex. Neuron, 75(6), 1114-1121.

Olsson, A., Knapska, E., & Lindström, B. (2020). The neural and computational systems of social learning. Nature Reviews Neuroscience, 21(4), 197-212.

Plate, R. C., Wood, A., Woodard, K & Pollak, S. D. (2019). Probabilistic learning of emotion categories. Journal of Experimental Psychology: General, 148(10), 1814-1827.

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. Annual Review of Psychology, 62, 23-48.

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. Nature Reviews Neuroscience, 15(8), Article 8. https://doi.org/10.1038/nrn3776

Ruffle, B. J. (1998). More is better, but fair is fair: Tipping in dictator and ultimatum games. Games and Economic Behavior, 23(2), 247-265.

Satterthwaite, F.E. (1946). An approximate distribution of estimates of variance components. Biometrics Bulletin, 2, 110-114.

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. Topics in Cognitive Science, 7(2), 351-367.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. Science, 275(5306), 1593–1599.

Schulze, C. & Newell, B. R. (2016). More heads choose better than one: Group decision making can eliminate probability matching. Psychonomics Bulletin and Review, 23, 907–914.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. Trends in Cognitive Sciences, 10(2), 70–76.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. Behavioral and Brain Sciences, 16(3), 495–511. https://doi.org/10.1017/S0140525X0003123X

Toyokawa, W., Whalen, A., & Laland, K.N. (2019). Social learning strategies regulate the wisdom and madness of interactive crowds. Nature Human Behaviour, 3(2), 183-193.

Ulber, J., Hamann, K., & Tomasello, M. (2017). Young children, but not chimpanzees, are averse to disadvantageous and advantageous inequities. Journal of Experimental Child Psychology, 155, 48-66.

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. Emotion, 11(2), 262–266.

Walther, J. B., Loh, T., & Granka, L. (2005). Let me count the ways: The interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. Journal of language and social psychology, 24(1), 36-65.

Witt, A., Toyokawa, W., Lala, K. N., Gaissmaier, W., & Wu, C. M. (2024). Humans flexibly integrate social information despite interindividual differences in reward. Proceedings of the National Academy of Sciences, 121(39), e2404928121.

Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 364(1528), 2417–2428.

Williamson, R. A., & Meltzoff, A. N. (2011). Own and others' prior experiences influence children's imitation of causal acts. Cognitive Development, 26(3), 260-268. doi: 10.1016/j.cogdev.2011.04.002

Wolf, W., Launay, J., & Dunbar, R. I. (2016). Joint attention, shared goals, and social bonding. British Journal of Psychology, 107(2), 322–337. doi.org/10.1111/bjop.12144

Zaki, J., & Mitchell, J. P. (2013). Intuitive prosociality. Current Directions in Psychological Science, 22(6), 466–470.