How does social learning affect stable false beliefs?

Rheza Budiono (rheza@nyu.edu)
Catherine A. Hartley (cate@nyu.edu)
Todd M. Gureckis (todd.gureckis@nyu.edu)
Department of Psychology, 6 Washington Pl
New York, NY 10003 USA

Abstract

Learning traps are false beliefs that entrench themselves by discouraging the exploration required to correct them. In previous lab experiments, these learning traps have proven remarkably difficult to prevent. Here, we investigate whether learning traps remain stable in contexts in which both individual and social learning are possible. In two of our three experiments, we found that learners trapped by a false belief were significantly more likely to escape a learning trap when they were able to observe another decision-maker's choices (without observing their outcomes). However, social learning was not a panacea. Social learning was constrained by the challenge of inferring others' beliefs, and trapped learners struggled to learn from partners with sub-optimal decision rules, even when their partner's choices were informative. Collectively, these results suggest that while social learning can help overcome the limits of individual learning, learning from others comes with its own challenges and limitations.

Keywords: learning traps; social learning; observational learning; exploration; selective attention; rule inference

Introduction

False beliefs are difficult to correct when they prevent the exploration needed to correct them. For example, a person might try an Indonesian restaurant for the first time and have a bad experience, leading them to hold the belief that they dislike all Indonesian restaurants. In turn, this belief leads them to avoid Indonesian food, which prevents subsequent updates to their belief. If there exists an Indonesian restaurant that they would really enjoy, their current (false) belief prevents them from trying it. In such cases, we describe the learner as "trapped" because their false belief causes them to avoid potentially corrective experiences (March, 1991; Denrell & March, 2001; Erev, 2014).

Several recent experimental studies have explored the situations under which such "learning traps" emerge, establishing them as robust phenomena in individual learners (Rich & Gureckis, 2018; Li, Gureckis, & Hayes, 2021; Allidina & Cunningham, 2021; Liquin & Gopnik, 2022; Blanco, Turner, & Sloutsky, 2023; Bai, Griffiths, & Fiske, 2024). In these studies, learning traps arise from the links between an individual's beliefs, choices, and experiences. However, in real-world environments, people can learn from others' choices as well as their own. Indeed, prior work has shown that people can learn from others' choices – even without observing the outcomes of those choices (Toyokawa, Kim, & Kameda, 2014; Toyokawa, Whalen, & Laland, 2019; Hawkins et al., 2023). Moreover, people may become more exploratory after observing others' exploratory choices, thus preventing them from falling into learning traps (Ortmann & Luhmann, 2023).

Here, we investigate whether learners trapped by a false belief remain trapped when both individual and social learning are possible. Social contexts may provide additional information that help people "escape" learning traps and form more accurate beliefs about the world. A trapped learner might observe others acting in ways that seem to contradict their false belief, causing them to subsequently explore. For example, after observing a friend visit a Indonesian restaurant, one might reconsider an avoidance policy for this cuisine. Alternatively, an individual trapped with a false belief might "explain away" social information that is inconsistent with their own belief. For example, a learner might disregard the line outside an Indonesian restaurant because they deem the people in line to be misinformed or different from themselves. Finally, when an individual and their social partner share the same false belief, this might reinforce the learning trap, increasing confidence that their false belief is correct through both confirmation and conformity.

Learned false-belief task

Our task builds upon an existing approach-or-avoid task in which participants often learned a stable false belief (Rich & Gureckis, 2018). In that task, individual participants were presented with objects composed of several discrete features (e.g., insects that vary along features like the number of wings or the number of stripes). On each trial, the subject was shown one object and asked if they would like to approach or avoid it. Approaching gives information about the quality of the object (in terms of gaining or losing money). Avoiding gives no information. Critically, the task was structured such that only certain combinations of the feature values of the objects were bad (see Fig. 1). For instance, in Fig. 1A only objects that have feature value 1 on dimension 2 and feature value 1 on dimension 2 are bad, the rest of the objects are good. The optimal decision rule that maximizes the reward while avoiding punishment is shown in Fig. 1B, where the learner approaches all the good items and avoids the bad ones (this is a "2D rule" since it requires attention to dimensions 1 and 2).

Rich & Gureckis found that subjects in these tasks tended to instead adopt one of two possible "false beliefs" about the

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

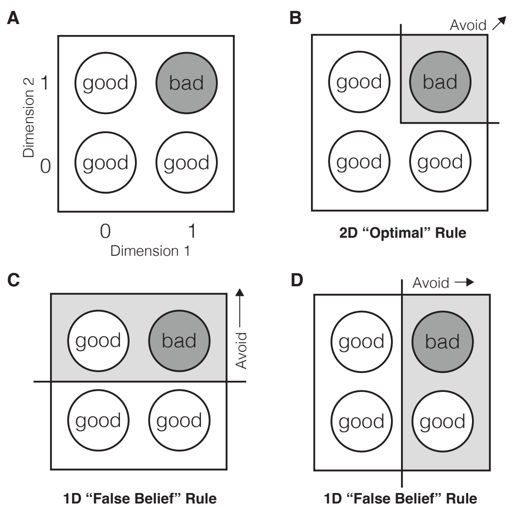

Figure 1: **The structure of the learned false belief task. A.** Objects vary on several dimensions with binary values (0 or 1). Only objects with feature value 1 on both dimension 1 and 2 are “bad”; the rest are “good” **B.** The optimal “2D” decision rule, which approaches all good things and avoids all bad things. **C & D**. Two sub-optimal decision rules. The policies here would avoid everything that has a value 1 on dimension 2 (panel C) or dimension 1 (panel D).

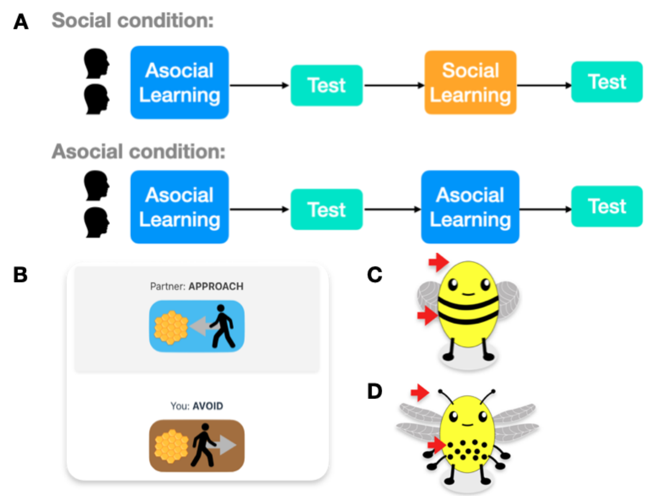

Figure 2: **Experiment details. A.** Schematic of the experimental design. **B.** Two example cartoon bees from the stimulus set. Cartoon bees varied along four binary feature dimensions: antennae/none, double/single wings, spots/stripes, and six/two legs. The red arrows indicate the two relevant features for a possible game instance. Dangerous bees are indicated by a conjunction of two features, such as antennae and spots (right). **C.** An example of a choice-sharing screen from the social learning phase. Partners observed each other’s choices, but not their associated outcomes.

task shown in Fig. 1C&D. In these cases, the subject either learned to avoid all objects with value 1 on dimension 2, or all objects with value 2 on dimension 1. Both of these rules are too general, as the subject would also avoid some rewarding items. However, subjects following either rule would never encounter negative feedback and were typically convinced that they had mastered the task. This false belief persisted because they had fallen into a learning trap — their early mistaken beliefs had become self-reinforcing. Rich and Gureckis (2018) found that only approximately 20% of participants learned the optimal 2D decision rule, whereas approximately 45% of participants settled for a 1D decision rule where they attended to only one of the two relevant features (Fig. 3).

Learning from another person’s choices

In this current work, we leverage the structure of this task design and ask whether someone stuck in a learning trap is more likely to escape after observing another person’s choices (but not outcomes). In our experiments, subjects first completed an individual learning session in which they faced a task similar to the one above. Then in a second phase, they continued to learn, but could also see the choices of a partner learning the same task (see Fig. 2A).

The initial individual learning phase allowed us to first establish the baseline belief pattern of each individual, but also to consider how learning in the second phase was influenced by the concordance of the two social partner’s beliefs. For example, learners trapped by a sub-optimal 1D decision rule (e.g., Fig. 1C&D) can be paired with a partner with the same belief, a partner using the optimal 2D decision rule (Fig. 1B), or a partner using a different sub-optimal 1D decision rule. In the latter two cases, the trapped learner sees their partner consistently approaching the bees they are mistakenly avoiding – thus receiving the evidence required to escape their learning trap.

Experiment 1

Experiment 1 was a live dyadic online experiment in which human participants played an approach-or-avoid game designed by Rich and Gureckis (2018), with an additional social learning component.[1]

Method

Participants We collected data from 176 participants (88 dyads) recruited from Prolific. There were 84 participants (42 dyads) in the social condition, and 92 participants (46 dyads) in the asocial control condition. Participants received \$11.25 for participation (45 minutes at a rate of \$15 per hour) and received a performance-based bonus that ranged up to \$4.

Stimuli and Design Participants encountered various cartoon bees and had to decide whether to approach or avoid

[1]See the project website for a demo of our task, additional visualizations, etc.: `https://gureckislab.org/papers/#/ref/budiono2024socialtrap`

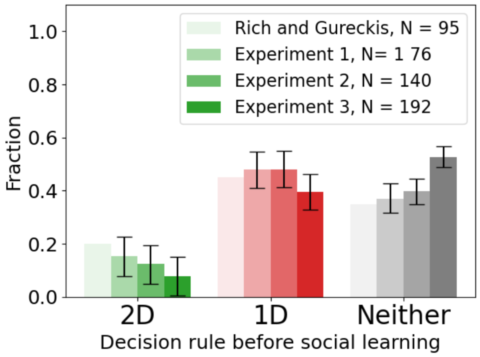

Figure 3: **Fraction of participants following each decision rule in the first test phase,** before social learning. We show data from all three of our experiments, as well as data from Rich and Gureckis (2018) for comparison.

them. Bees could be friendly or dangerous – if you approached a friendly bee, you would harvest honey (+1 point); if you approached a dangerous bee, you would get stung (-5 points).

Cartoon bees were designed that varied along four salient binary feature dimensions. They could have antennae or no antennae, they could have single wings or double wings, they could have stripes or spots, and they could have 2 legs or 6 legs. There were 16 unique bees in total. Unbeknownst to the participant, two of these four features could be used to perfectly predict which bees were friendly and which were dangerous, and a unique conjunction of these two features determined whether a bee was dangerous (Fig. 1A). For example, the two relevant features might be stripes/spots and antennae/none, and bees were dangerous if and only if they had both spots and antennae (Fig. 2D). In this case, the optimal 2D decision rule would be to avoid a bee if and only if it had both spots and antennae. Crucially, however, one could also avoid all punishment (while missing some rewards) using a sub-optimal 1D decision rule that attends to only one of the two relevant features (e.g. "avoid a bee if it has spots").

Individual and Social Learning Phases Participants were assigned to one of two conditions: a "social learning" condition, or an "asocial control" condition. In both conditions, participants were paired with a partner, completed an asocial learning phase, and completed a test phase during which they made choices without observing choice outcomes (Fig. 2A). Then, participants in the social learning condition would enter a social learning phase during which partners could observe each other's choices at the end of each trial (Fig. 2A). Participants in the asocial control condition instead completed a second asocial learning phase. Finally, all participants completed another test phase. In both conditions, there were 192 trials split over the four phases. Each learning phase consisted of 64 trials, and each test phase consisted of 32 trials.

There are a few details about the social learning phase that are important to note. First, participants could not copy their partner's choice on any given trial, because choices were shared after each participant in the dyad had made their choice. Instead, a partner's choice was an additional piece of feedback that could be used for learning. Second, choice outcomes were not shared (Fig. 2B). This means that if a participant decided to avoid a bee and their partner approached it, they did not get to see whether the bee was friendly or dangerous.

Finally, note that participants in the asocial control condition were still paired with a partner and waited to start each trial together. Moreover, they were not told ahead of time whether they would do a social learning phase or a second asocial learning phase. This was done in order to control for possible confounds between the two conditions, such as differing inter-trial interval distributions and social motivations.

Results

In each test phase, we classified participants' behavior as following either the optimal 2D decision rule, a sub-optimal 1D decision rule (a learning trap), or neither. Following Rich and Gureckis (2018), a participant was classified as following a decision rule if 30 out of 32 decisions in the test phase were consistent with the decision rule. There are 16 unique bees, each shown twice during the test phase, so this criterion allows for two deviations from the rule over the two full passes of the stimulus set.

Replicating Rich and Gureckis (2018), we found that the learning trap was prevalent in the first test phase (before social learning). Approximately half of all participants displayed a 1D decision rule, selectively attending to only one out of two relevant features (Fig. 3). Furthermore, we found that the learning trap was robust to additional individual learning trials. A second individual learning phase did not significantly reduce the number of asocial control participants who were trapped (Fig. 4A).

After confirming the stability of the learning trap in the purely individual learning setting, we investigated whether learners who were trapped by a false belief could break free by observing the choices of another decision-maker. As previously discussed, a trapped learner could in principle escape by learning from either (1) a learner following the optimal 2D decision rule, or (2) a learner trapped by a different sub-optimal 1D decision rule. In both cases, their partner's choices would provide evidence against the learning trap because their partner would approach the bees the trapped observer is mistakenly avoiding.

Ultimately, our analyses were limited by the number of naturally occurring dyads of interest. There were only 7 instances of a trapped learner paired with a partner following the optimal 2D decision rule, and only $1/7$ subsequently escaped their learning trap. Although this fraction is larger than the fraction of trapped learners who escaped their learning trap in the asocial control group, the difference was not sta-

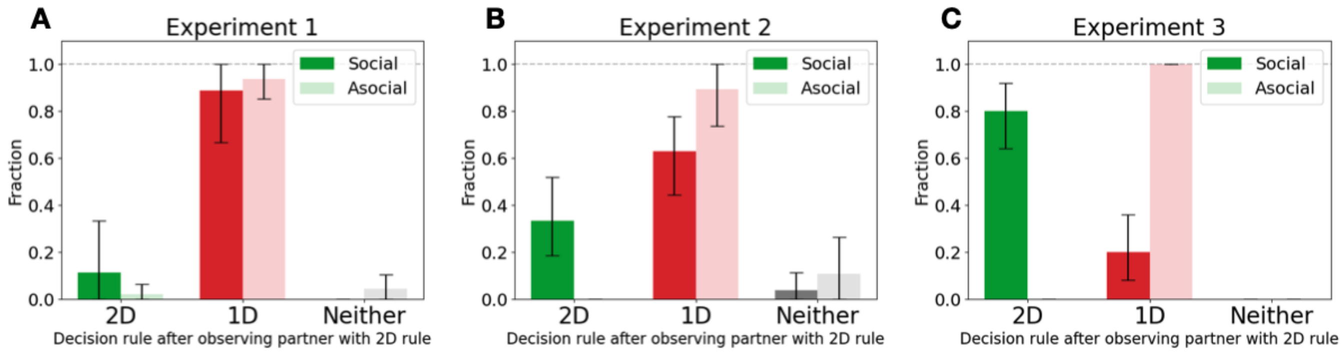

Figure 4: **Effect on trapped learners of observing a partner following the optimal 2D decision rule.** These plots show the fraction of selected participants following each decision rule in the second test phase (with 95% CIs, bootstrapped with 1,000 resamples). In the social condition, we selected participants who (1) displayed a sub-optimal 1D decision rule in the first test phase, and (2) observed an optimal 2D decision-maker during the social learning phase. In the asocial condition, we selected participants who (1) displayed a sub-optimal 1D decision rule in the first test phase.

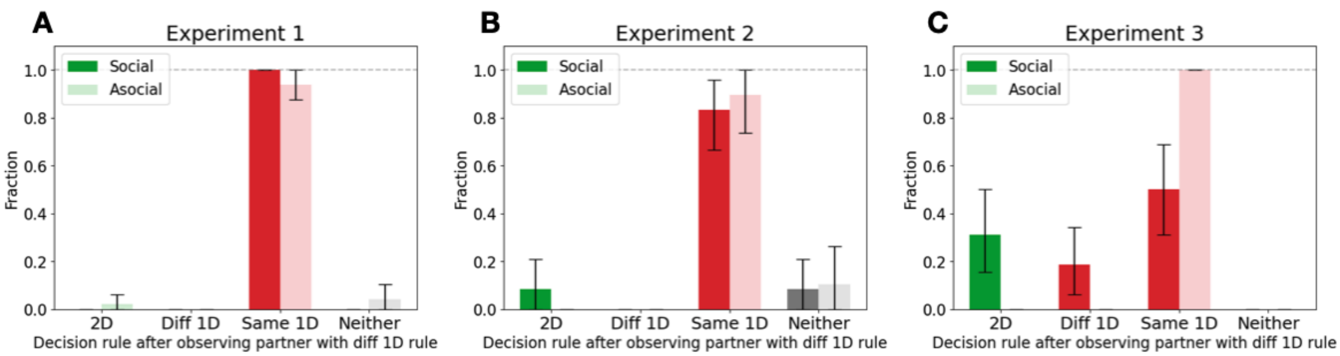

Figure 5: **Effect on trapped learners of observing a partner following a different 1D decision rule.** These plots show the fraction of selected participants following each decision rule in the second test phase (with 95% CIs, bootstrapped with 1,000 resamples). In the social condition, we selected participants who (1) displayed a sub-optimal 1D decision rule in the first test phase, and (2) observed a decision-maker who followed a different sub-optimal 1D decision rule during the social learning phase. In the asocial condition, we selected participants who (1) displayed a sub-optimal 1D decision rule in the first test phase.

tistically significant (Fig. 4A; $P > 0.3$[2]; $N_{\text{social}} = 7$, $N_{\text{asocial}} = 48$). There was also no significant effect on trapped learners of observing a partner following a different sub-optimal 1D decision rule (Fig. 5A; $P > 0.3$, $N_{\text{social}} = 8$, $N_{\text{asocial}} = 48$).

Experiment 2

In Experiment 1, we were limited by the number of naturally occurring dyads of interest. To overcome this data limitation in Experiment 2, we paired human participants with bots who made choices according to programmed decision rules.

Participants

We collected data from 184 participants recruited from Prolific. There were 140 participants in the social condition, of which 69 observed a 2D decision rule during the social learning phase and 71 observed a 1D decision rule. There were 44 participants in the asocial control condition. Participants received \$10 for participation (40 minutes at a rate of \$15 per hour) and received a performance-based bonus that ranged up to \$4.

[2]All significance tests are one-sided bootstrap tests (1, 000 resamples).

Method

Experiment 2 was structurally identical to Experiment 1 (Fig. 2A). However, instead of observing another human decision-maker during the social learning phase, participants observed choices generated by a computer program which implemented one of three decision rules: it could follow a 2D decision rule (i.e., it avoided a bee if the bee displayed both relevant features), or it could follow one of two 1D rules (i.e., it avoided a bee if it displayed just one of the two relevant features). On a given trial, the bot's response time was drawn from a gamma distribution fit to the distribution of participant response times from Experiment 1. In the consent form, participants were told that they would be either paired with a human or a bot. All subsequent mentions of their partner referred to them as their "partner," without specifying whether their partner was a human or a bot.

We also changed the graphic design of the outcome feedback screen in Experiment 2. Due to a coding error, the new outcome screens also showed each participant their total score during the test trials. This meant that participants could have inferred the outcomes of the test trials by noting their scores before and after. However, the vast majority of asocial control participants who displayed a 1D decision rule in the first intended test phase remained trapped by a 1D decision rule in the second test phase (Fig. 4B). Indeed, no such participants were able to transition to the optimal 2D decision rule, and the percentage of participants persisting with a 1D decision rule was not significantly different than 100%. This shows that in the absence of a social learning intervention, participants' beliefs had likely reached an equilibrium by the end of the first learning phase.

The final change was that we added attention checks to the social learning phase. On each attention check, participants had to report their partner's choice on the last trial. There were 12 attention checks distributed randomly throughout the social learning phase. We verified that participants in the social condition were paying attention to their partner's choices. Over 90% of participants missed one or zero attention checks, and 74% of participants correctly answered all attention checks.

Results

First, we confirmed the stability of the learning trap in the asocial control condition. The learning trap was remarkably stable, as evidenced by the fact that no trapped learners in the asocial control condition were able to learn the optimal 2D decision rule after a second individual learning phase (Fig. 4B).

Next, we tested whether trapped learners were more likely to learn the optimal 2D rule after observing either (1) a learner following the optimal 2D decision rule, or (2) a learner trapped by a different sub-optimal 1D decision rule. Recall that in principle, a trapped learner can learn from both kinds of partners.

Unlike in Experiment 1, there was a significant effect of observing a partner following an optimal 2D decision rule (Fig. 4B; $P < 1/1{,}000$, $N_{\text{social}} = 25$, $N_{\text{asocial}} = 19$). A third of the previously trapped learners learned the optimal 2D rule after observing a partner following the 2D rule. This shows that after observing a partner following the optimal 2D decision rule, some trapped learners were able to infer that they were mistakenly avoiding rewarding bees after observing their partner approach those bees.

Although a couple trapped learners were able to escape after observing a partner who followed a different 1D decision rule, this was not a statistically significant improvement compared to the asocial control condition (Fig. 5B; $P > 0.1$, $N_{\text{social}} = 24$, $N_{\text{asocial}} = 19$). Trapped learners were either less willing or less able to learn from a partner following a different sub-optimal 1D decision rule, despite the fact that such a partner would approach the bees that the trapped learner mistakenly avoids. In principle, this is the evidence required for the trapped learner to escape.

Experiment 3

In Experiment 2, social information was only moderately successful in helping trapped learners escape. Two thirds of trapped learners remained trapped after observing an optimal 2D partner, and over 90% remained trapped after observing a partner following a different 1D decision rule. We wondered why there were not more trapped learners who learned from their partners, and why trapped learners were much more likely to learn from observing choices generated by a 2D decision rule versus a 1D decision rule.

One possible contributing factor is the difficulty of inferring another decision-maker's decision rule. In our task, decision-rule inference seemed to be a crucial first step towards learning from a partner. Although it is theoretically possible for participants to recall their partner's response to each stimulus, it seems prohibitively difficult due to the relative complexity of our stimuli. The challenge of decision-rule inference could explain why many trapped learners generally failed to learn from their partners in the previous experiments, as well as why trapped learners were much more likely to learn from an optimal 2D partner than a partner following a different 1D decision rule. First, a 1D decision rule may have been inherently harder to infer for someone with a different 1D decision rule, because it disagrees more with the other 1D rule compared to a 2D rule. Additionally, trapped learners may have deemed a partner following a different 1D decision rule to be incompetent (as they incorrectly avoided half of the rewarding stimuli the trapped learner knows about), thus ignoring their partner's choices and choosing not to infer their partner's decision rule.

In Experiment 3, we tested the hypothesis that the challenge of decision-rule inference prevented participants from learning from their partners. In particular, we removed the need for decision rule inference by directly providing a description of the rule at the start of the social learning phase. We reasoned that removing the need for decision-rule inference would allow us to better understand the extent to which it limited trapped learners' ability to learn from their partners.

Participants

We collected data from 192 participants recruited from Prolific. There were 139 participants in the social condition, of which 49 observed a 2D decision rule during the social learning phase and 90 observed a 1D decision rule. There were 53 participants in the asocial control condition. Participants received \$7.50 for participation (30 minutes at a rate of \$15 per hour) and received a performance-based bonus that ranged up to \$4.

Method

Experiment 3 was identical to Experiment 2, except in two respects. First, we fixed the mistake in the outcome screen so that the total score was correctly omitted during test trials. Second, as previously mentioned, participants in the social condition were told their partner's decision rule from the first test phase (before social learning commenced). Before being told their partner's decision rule, participants were first asked to describe their own decision rule. They were asked "Which bees should you avoid?" and responded by inputting natural language into a text box. They were told that their answer would be shared with their partner, and that their partner's response would be shared with them. An example description of a 2D decision rule reads: "Avoid bees with both antennae and dots on their body." Participants in the asocial control condition were also asked to report their own decision rule in a "mid-game questionnaire."

Results

Compared to previous experiments, trapped learners were overall much more likely to escape the learning trap. It seems that trapped learners benefited greatly from being able to read their partner's decision rule in addition to observing their partner's choices.

Eighty per cent of trapped learners paired with an optimal 2D partner were able to learn the optimal 2D decision rule, compared to none in the asocial control condition (Fig. 4C; $P < 1/1{,}000$, $N_{\text{social}} = 25$, $N_{\text{asocial}} = 19$). For comparison, in Experiment 2, only a third of trapped learners paired with an optimal 2D partner were able to learn the optimal 2D decision rule (Fig. 4B).

Moreover, over 30% of trapped learners paired with a partner following a different 1D decision rule were able to learn the optimal 2D decision rule, compared to none in the asocial control condition (Fig. 5C; $P = 2/1{,}000$, $N_{\text{social}} = 32$, $N_{\text{asocial}} = 19$). For comparison, in Experiment 2, under 10% of trapped learners paired with a partner following a different 1D decision rule were able to learn the optimal 2D decision rule (Fig. 5B).

Altogether, these results suggest that participants were previously limited by the challenge of decision-rule inference – both when they had to infer a 2D decision rule and when they had to infer a different 1D decision rule. Notably, many trapped learners still failed to learn from their partners even after reading their partner's informative decision rule. It was not enough for a trapped learner to know their partner's decision rule, a trapped learner also had to (1) infer that they were likely avoiding some rewarding bees, and (2) take the risk of exploration. Moreover, a trapped learner was still less likely to learn from the choices of a sub-optimal 1D decision-maker versus the choices of an optimal 2D decision-maker. The challenge of decision-rule inference did not wholly account for the disparity in Experiment 2. Thus, trapped learners observing a different sub-optimal 1D decision rule were less likely to do either (1) or (2).

Discussion

Here, we found that people could escape learning traps by learning from the choices of a partner. However, a majority of trapped learners still remained trapped after observing a partner's potentially informative choices, and trapped learners were much more likely to learn from the choices of a partner following an optimal decision rule versus a different sub-optimal one that was nonetheless informative. Telling trapped learners their partner's decision rule (thereby removing the need for decision-rule inference) made trapped learners much more likely to learn from their partners, although it did not change the fact that they were much more likely to learn from a partner following an optimal decision rule.

We identified two possible explanations (which are not mutually exclusive) for why trapped learners were less likely to learn from a partner following a different sub-optimal decision rule. First, some trapped learners could have been employing a limited social-learning strategy, such that they could only learn from their partner by copying their decision rule verbatim. A trapped learner employing a "copy or not" strategy could escape their learning trap by copying a 2D decision rule, but not another 1D decision rule. The existence of such limited social learners is evidenced by the substantial fraction (over 15%) of trapped learners who copied a different 1D decision rule after it was explicitly described in Experiment 3 (Fig. 5C), jumping from one learning trap to another, despite having collected the experiences necessary to act optimally.

Another possible explanation is that some participants' willingness to learn from their partners was modulated by their judgment of their partner's competence. For instance, suppose that Alice follows a 1D decision rule and observes the choices of Bob, who follows a different 1D decision rule. From Alice's perspective, Bob appears incompetent because he approaches only half of the rewarding bees that Alice knows about(Fig. 1C,D). As such, Alice may decide not to pay attention when Bob approaches bees that Alice avoids. On the other hand, if Bob employs a 2D decision rule, then Bob appears competent because he approaches all the rewarding bees that Alice knows about. Subsequently, Alice might be more likely to pay attention when Bob approaches a bee that Alice avoids. Such an effect of choice agreement on social influence was found by Najar, Bonnet, Bahrami, and Palminteri (2020).

More broadly, the present work suggests that social learning can be an effective intervention against learning traps. Future work elucidating the contexts in which social information is most influential can help us better understand the dynamics of collective beliefs (Toyokawa et al., 2019; Witt, Toyokawa, Gaissmaier, Lala, & Wu, 2024; Wisdom, Song, & Goldstone, 2013), and inform how we might design social networks in order to mitigate learning traps (Erev, 2014; Hardy, Thompson, Krafft, & Griffiths, 2023).

Acknowledgments

We thank Emily Liquin, Pat Intara (formerly Pat Little), and Solim LeGris for helpful feedback on the manuscript. We also thank the anonymous reviewers for their valuable feedback. This work was supported by NSF BCS 2021060 to TMG.

References

Allidina, S., & Cunningham, W. A. (2021). Avoidance begets avoidance: A computational account of negative stereotype persistence. *Journal of Experimental Psychology: General*, *150*(10), 2078–2099. doi: 10.1037/xge0001037

Bai, X., Griffiths, T., & Fiske, S. (2024, January). Multidimensional Stereotypes Emerge Spontaneously When Exploration is Costly.
doi: 10.31234/osf.io/mbuhv

Blanco, N. J., Turner, B. M., & Sloutsky, V. M. (2023, February). The benefits of immature cognitive control: How distributed attention guards against learning traps. *Journal of Experimental Child Psychology*, *226*, 105548. doi: 10.1016/j.jecp.2022.105548

Denrell, J., & March, J. G. (2001). Adaptation as Information Restriction: The Hot Stove Effect. *Organization Science*, *12*(5), 523–538.

Erev, I. (2014, October). Recommender systems and learning traps. In *Proceedings of the first international workshop on decision making and recommender systems* (pp. 38–41). Bolzano, Italy: Free University of Bozen-Bolzano.

Hardy, M. D., Thompson, B. D., Krafft, P. M., & Griffiths, T. L. (2023, December). Resampling reduces bias amplification in experimental social networks. *Nature Human Behaviour*, *7*(12), 2084–2098. doi: 10.1038/s41562-023-01715-5

Hawkins, R. D., Berdahl, A. M., Pentland, A. S., Tenenbaum, J. B., Goodman, N. D., & Krafft, P. M. (2023, October). Flexible social inference facilitates targeted social learning when rewards are not observable. *Nature Human Behaviour*, *7*(10), 1767–1776. doi: 10.1038/s41562-023-01682-x

Li, A. X., Gureckis, T. M., & Hayes, B. K. (2021). Can losses help attenuate learning traps? *Cognitive Science Society*.

Liquin, E. G., & Gopnik, A. (2022, January). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, *218*, 104940. doi: 10.1016/j.cognition.2021.104940

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, *2*(1), 71–87.

Najar, A., Bonnet, E., Bahrami, B., & Palminteri, S. (2020, December). The actions of others act as a pseudo-reward to drive imitation in the context of social reinforcement learning. *PLOS Biology*, *18*(12), e3001028. doi: 10.1371/journal.pbio.3001028

Ortmann, A., & Luhmann, C. (2023). Social Learning from Incomplete Information in a Dynamic Decision-Making Task. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*(45).

Rich, A. S., & Gureckis, T. M. (2018, September). The limits of learning: Exploration, generalization, and the development of learning traps. *Journal of Experimental Psychology: General*, *147*(11), 1553. doi: 10.1037/xge0000466

Toyokawa, W., Kim, H.-r., & Kameda, T. (2014, April). Human Collective Intelligence under Dual Exploration-Exploitation Dilemmas. *PLOS ONE*, *9*(4), e95789. doi: 10.1371/journal.pone.0095789

Toyokawa, W., Whalen, A., & Laland, K. N. (2019, February). Social learning strategies regulate the wisdom and madness of interactive crowds. *Nat Hum Behav*, *3*(2), 183–193.

Wisdom, T. N., Song, X., & Goldstone, R. L. (2013). Social learning strategies in networked groups. *Cognitive Science*, *37*(8), 1383–1425. doi: 10.1111/cogs.12052

Witt, A., Toyokawa, W., Gaissmaier, W., Lala, K., & Wu, C. M. (2024, February). Social learning with a grain of salt.
doi: 10.31234/osf.io/c3fuq