Overcoming Learning Traps Through Social Learning

Brett K. Hayes (b.hayes@unsw.edu.au), Mathew E. Young (matty.young03@gmail.com), Yanjun Liu (yanjun.liu6@unsw.edu.au) & Ben R. Newell (ben.newell@unsw.edu.au)

School of Psychology, University of New South Wales (UNSW), Sydney, Sydney, NSW 2052, Australia

Abstract

People often fall into learning traps where a false belief about the structure of the environment leads to under-exploration of rewarding options. Two studies (N = 324) examined whether observation of the approach decisions of another learner facilitated escape from a trap. After an initial learning phase in a task where approach of different category members could lead to gains or losses, we identified whether participants had learned an optimal two-dimensional categorization rule or fallen into the trap of using a one-dimensional rule. Participants then observed the approach decisions of another learner using the same or a different category rule. Participants' categorization rules in a final round of category learning were then assessed. A substantial proportion of those who had initially fallen into the trap shifted to the optimal rule after observing use of an alternate rule. This effect was found following observation of both optimal (Experiment 1) and sub-optimal rules (Experiment 2). In contrast, those who learned the optimal rule in the initial learning phase were unaffected by social observation. The results show that social learning is a viable approach for facilitating escape from learning traps.

Keywords: Learning traps; social learning; decision-making; category learning

Introduction

Imagine that an investor wishes to shift their portfolio toward more ethical and sustainable investments. However, their first experience investing in a green energy start-up ends badly. After that, they avoid all investments involving green energy. This means that they avoid repeating their earlier loss – but it also means that they may miss out on some highly potentially profitable investments (Forbes, 2023). This is an example of *a learning trap* – where an early negative experience leads to a false or incomplete belief about the environment ("all ethical/green investments perform poorly"), which in turn leads to avoidance of rewarding choice options. In environments where outcome feedback is only available for chosen but not forgone options, this cycle persists because the learner receives no new information that would lead them to revise their false belief (Denrell & March, 2001; Rich & Gureckis, 2018).

In addition to reduced economic rewards (e.g., Li, Gureckis & Hayes, 2021; Rich & Gureckis, 2018), learning traps can lead to poor management decisions (Elwin, 2013), polarized beliefs (Perfors & Navarro, 2019), distorted first impressions and negative social stereotypes (Bai, Fiske & Griffiths, 2022; Denrell, 2005). Previous work has shown that traps can form relatively early in the learning process and that, once formed, can persist over time despite opportunities for further learning (Liu, Newell, Lee & Hayes, 2024). Moreover, those in a trap are often blind to important changes in the reward structure of their environment (Blanco & Sloutsky, 2019; Lee, Li, Lee & Hayes, 2024).

Given their profound negative effects, it is crucial that we gain a better understanding of the mechanisms that underlie trap formation and endeavor to use this information to find ways of preventing or escaping from traps. Unfortunately, to date, most attempts to remediate the effects of traps have been unsuccessful. Rich and Gureckis (2018), for example, trialed a number of interventions to reduce learning trap prevalence (e.g., individuating exemplars with unique features, adding stochasticity to outcomes, occluding some features). These manipulations failed to reduce trap frequency or sometimes led to general impairments in learning, so that learners failed to identify any of the cues that predicted rewards.

Learning Traps and Social Learning

Learning traps most often arise when a learner only receives feedback about chosen options but not those they have avoided. They are less likely to form when learners are provided with "full feedback" on the outcomes associated with both chosen and foregone options (Liu et al., 2024; Rich & Gureckis, 2018). Given that full-feedback is rare in learning environments outside the laboratory, one might wonder why traps are not even more widespread? One answer is that learners outside the laboratory often have access to other forms of feedback – including the behavior of other learners. Our green-energy shy investor may note that her colleagues are continuing to invest in companies that promote green energy with profitable results. Observing such behavior could lead the investor to re-examine green energy options, which could lead to escape from the trap.

In such examples, there are two types of information that could impact the learner's decisions – the observation of *which options other people choose to approach or avoid* and *the outcomes* of these choices. In the latter case, observing that someone else has approached a choice option that you are currently avoiding and that this produced a reward, could be seen as a form of full feedback. Hence, escape from the current learning trap seems likely. What is less clear is whether escape from a trap would be facilitated by merely *observing the approach and avoidance patterns of other learners, without outcome feedback*. The current studies aimed to answer this question.

Learners' exploration decisions are often influenced by observing the behavior of others faced with the same decision-making tasks (Rendell et al., 2011; Whalen, Griffiths & Buchsbaum, 2018; Xie & Hayes, 2022). Previous

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

work has shown that observing other people's decisions can improve individual performance in bandit tasks where the goal is to learn which of two or more response options yields the highest mean rewards (Toelch et al., 2013; Toyokawa, Kim & Kameda, 2014, Toyokawa, Whalen & Laland, 2019). For example, Toyokawa et al., (2014) tracked the performance of participants learning about the reward structure in a multi-armed bandit task. All participants could choose to explore options and receive the associated payoffs. Some learners could only do so on an individual basis while others had the opportunity to observe the options that other learners had approached. Those with access to this social information earned more rewards over the course of the task than those learning individually (also see Hawkins, Berdahl, Pentland, Tenenbaum & Kraft, 2023).

In many bandit tasks the only cues to guide future decisions are the relative frequency and magnitude of rewards associated with previous decisions. In learning traps tasks, however, features associated with each option can be used to predict the outcome if that option is approached. Many previous studies of learning traps (e.g., Lee et al., 2024, Rich & Gureckis, 2018) have used a feature structure such that approaching the members of categories based on different combinations of visual features will lead to either gains or losses respectively. Such a structure is illustrated in Figure 1A, where approaching three combinations of binary features leads to a small gain but approaching one combination leads to a large loss. Learning the optimal approach pattern requires integrating information across two dimensions. However, previous work (e.g., Lee et al., 2024, Li et al., 2021; Rich & Gureckis, 2018) has shown that many learners fall into the trap of using an overly simplistic single-dimension rule, as illustrated in Figure 1B (e.g., approach all stimuli with feature value 0 on dimension 1 and avoid those with feature value 1). This is analogous to our investor treating all green energy stocks as equivalent. In the learning traps task, this strategy avoids losses but also leads to avoidance of rewarding items.

To date, few studies have examined how observation of different decision patterns impacts the formation of learning traps (but see Budiono, Hartley & Gureckis, 2024). In one of the few exceptions, Liquin and Gopnik (2022) presented children (4-5 years old) and adults with a learning traps task similar to that used in the present studies. They found that young children tended to show more initial exploration of stimuli, and hence were less likely to fall into traps, than adults. In subsequent studies, however, they found that these general trends could be reversed if adults were exposed to "child-like" evidence (i.e., observing approach and avoid decisions from a child learner) or if children were exposed to "adult-like" evidence. Although the alternative forms of evidence were not explicitly social in nature, these results suggest that adults who have fallen into a trap may benefit from observing an alternative pattern of approach behavior. This conclusion must be treated as tentative, however, since both adults and children in each evidence condition were also given explicit hints about the relevant rule.

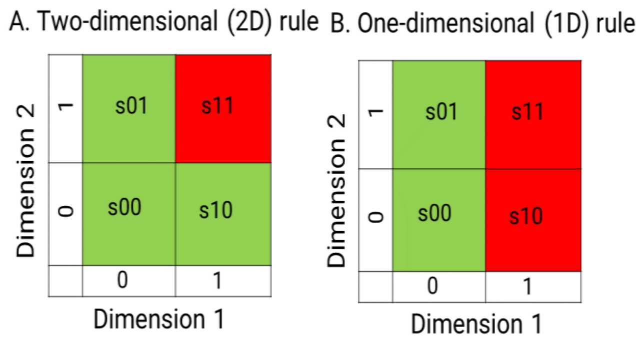

Figure 1. Stimulus structure used in the current studies showing A. optimal approach (green) and avoid (red) patterns, and B. approach and avoid patterns indicative of a learning trap

The Current Studies

Our studies tackled the novel question of whether observing the approach and avoidance patterns of another learner or "demonstrator" can facilitate escape from a one-dimensional learning trap. To escape from such a trap, the learner has to approach some stimuli that they are currently avoiding and discover that they yield rewards. Hence, in each study, after a period of individual learning where some learners fell into a one-dimensional trap, learners had the opportunity to observe the approach patterns of another (virtual) learner. In some cases, these approach patterns were consistent with those currently used by the learner, and in other cases they were inconsistent (see Figure 2). In the latter case, the demonstrator approached some instances that the learner was currently avoiding. Our overarching question was whether exposure to such rule-inconsistent information would assist learners in discovering the optimal two-dimensional rule in a final individual learning phase.

We also examined how learners responded to observing other learners' decisions, and how this influences learning performance. Previous work on social learning and cultural transmission (e.g., Lyons, Young & Keil, 2007; Morin, 2016; Wu, Vélez & Cushman, 2022) suggests two distinct ways that learners may respond after observing a decision-making pattern that differs from their own. They could *imitate* the observed behavior, replacing their current pattern of approach decisions with those of the demonstrator. In this case, there is no attempt to infer why the demonstrator exhibits this behavioral pattern – it is simply assumed that it is associated with a higher net reward. In contrast, learners may infer the strategy or beliefs that guide the demonstrator's behavior, and compare those to their current beliefs about the reward environment. In this case, the learner may *emulate* the beliefs or rules that are thought to underlie the demonstrator's behavior (Wu et al., 2022).

The current studies examined whether people generally imitate or emulate the patterns of other learners in a learning traps task. In each experiment, after the initial learning phase, we identified people who had already discovered the optimal two-dimensional rule (2D learners) as well as those who had fallen into a one-dimensional trap (1D learners). As shown in Figure 2, in subsequent social learning 2D learners observed the demonstrator using a 2D-rule (consistent condition) or a suboptimal 1D rule (inconsistent condition). In contrast, for initial 1D learners, consistent social learning involved seeing the demonstrator using the same sub-optimal rule, while those in the inconsistent condition saw a demonstrator use a different rule (in Experiment 1, this was the optimal 2D rule). If learners simply imitate, then in the final learning phase, those in the consistent conditions should shift their behavior towards the observed pattern. This may benefit initial 1D learners, who shift towards the optimal rule, but would harm 2D learners, who would shift towards a 1D rule. In contrast, if learners emulate, then we may see improvement in the performance of the 1D learners without decrement in the performance of the 2D learners, who may ignore the observed pattern if they infer it leads to fewer rewards.

Experiment 1

Method

Participants. The participants were 174 adults (M_{age} = 36.08 years, 82 males, 91 females, 1 other) recruited from UNSW undergraduates (n = 24) or the online platform Prolific (n = 150). Participants received a base reward of $1.00 or £1.00 (Prolific) for participation in the task, and of up to $2.00 or £2.00 (Prolific) based on task performance. No participants were excluded.

Design and Procedure. The experimental design is summarized in Figure 2 and consisted of three phases: an initial category learning task that allowed us to identify sub-groups of participants using different category rules (phase 1), a social learning phase where learners observed a virtual demonstrator's approach/avoid trials (phase 2), and a final learning phase (phase 3). The experimental paradigm was implemented in jsPsych (de Leeuw, 2015).

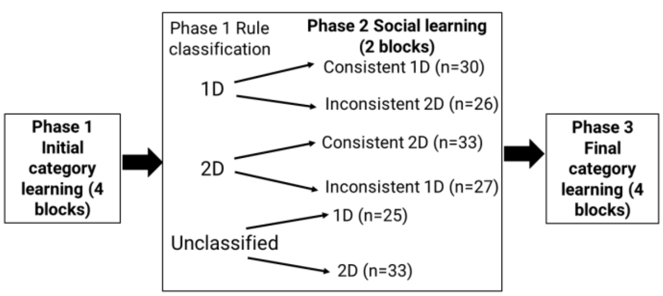

Figure 2. Summary of Experiment 1 design

Phase 1 – Initial category learning. Participants were told that they were "virtual beekeepers" tasked with collecting honey from beehives. Hives were inhabited by either "friendly" bees who produced honey (leading to a +1 points gain) or "dangerous" bees who would sting (leading to a -3 points loss). Bee stimuli were similar to those used by Rich and Gureckis (2018) and Lee et al. (2024). They varied on three binary-valued feature dimensions (body pattern: stripes or spots; wings: single or double; number of legs: 2 or 6) (stimuli can be viewed at: https://osf.io/c68e7/?view_only=1990207467dc48cd9648d5d5a4608d16).

Two feature dimensions were relevant to categorizing bees. As per Figure 1A, a conjunctive two-dimensional rule perfectly predicted category membership. Features on the remaining dimension were not predictive. The relevant dimensions and feature combinations that determined category membership were randomly assigned for each participant. Before commencing the learning phase, participants had to reach errorless performance on an instruction comprehension check consisting of four multiple-choice questions. If a participant made any errors, they were returned to the instruction screens.

Learning in phase 1 consisted of 4 blocks of 16 trials. On each trial, an exemplar was presented and the participant chose to approach ("harvest") or avoid the exemplar by pressing labelled on-screen buttons. Participants started with a balance of 50 points. Approaching a friendly bee led to a +1 points gain; approaching a dangerous hive led to a -3 points loss. If a bee was avoided, no feedback was provided and the points balance was unaffected. An on-screen counter tallied current earnings. In each block, the 8 unique exemplars were presented twice in random order. Within a block, a learner's rule use was classified as 2D or 1D if their approach/avoid pattern was consistent with the respective rule on at least 15 trials. If choices did not conform to either rule, the learner was said to use an "unclassified rule". Learners were classified into one of three sub-groups based on the category rule they used in the final block in this phase. Note, that the stimulus structure meant there were two possible 1D rules, focused on one of the relevant dimensions (i.e., 1D-dimension 1 or 1D-dimension 2 rule). Classification into sub-groups was done automatically by the experimental code.

Phase 2 – Social learning. Participants were introduced to an on-screen visual avatar of another player and told that the player had spent an equivalent amount of time playing the honey harvester game and watching their approach/avoid decisions may or may not be helpful for earning points. Avatar gender and a gender-neutral name was randomly assigned from a pool of such names. Although there was no actual player, participants were informed that the approach/avoid patterns shown for each avatar matched those of actual participants in previous studies in our lab.

Two blocks of 16 observation trials were then presented, where a training stimulus was shown, and participants were told whether or not the demonstrator approached it. As per Figure 2, for each of the 2D and 1D learner subgroups from phase 1, approximately half the participants were assigned to consistent or inconsistent social learning conditions. Those in the consistent conditions observed an approach/pattern that was consistent with the learner's phase 1 rule (2D, 1D-dimension 1 or 1D-dimension 2). Those in the inconsistent conditions observed a different rule: 2D learners observed a demonstrator making approach decisions based on a 1D-dimension 1 or 1D-dimension 2 rule; 1D learners observed a demonstrator using a 2D-rule. Those using an unclassified rule in phase 1 were randomly assigned to view a pattern following either a 1D or 2D rule. At the end of this phase, as a manipulation check, participants were asked to judge whether the observed pattern was similar or different to their existing rule.

Phase 3 – Final category learning. After social learning, participants returned to the honey harvester game and had a chance to earn more points. They completed four more blocks of 16 trials. The type of category rule that learners were using on the final block (1D, 2D, Unclassified) was the key dependent measure.

Results and Discussion

Manipulation check. All participants who initially learned a 1D or 2D rule and observed consistent approach/avoid patterns correctly judged them as following the same rule. A majority (70%) of those in the inconsistent condition correctly judged the demonstrator as using a different rule. This was taken as evidence that our social learning manipulation worked as intended.

Rule learning in Phases 1 and 3. The proportion of participants using different category rules in each learning block in phase 1 is shown in Figure 3. By the final block in this phase, 34.5% of participants had learned the optimal 2D category rule, 32.2% had fallen into the 1D learning trap and 33.3% were using an unclassified rule. These proportions are similar to those observed in previous studies using similar category stimuli and payoffs (e.g., Lee et al., 2024; Li et al., 2021; Liu et al., 2024).

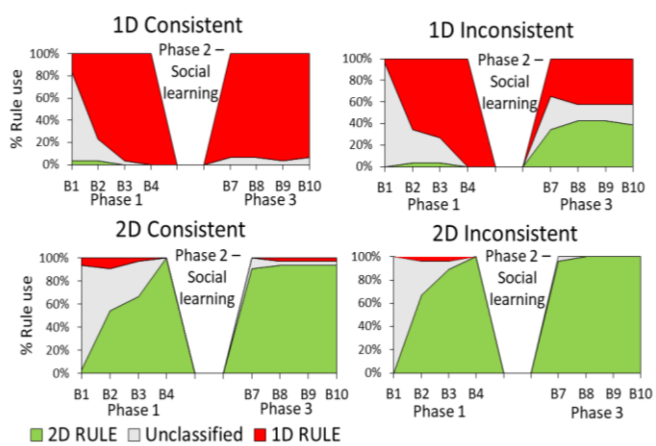

Figure 3. Exp. 1: % of participants using various category rules in Phases 1 (B1-B4) and 3 (B7-B10) in each social learning condition. Only those using a 1D or 2D rule in the final block of Phase 1 are shown.

The key question was whether, after phase 2 social learning, each sub-group continued to use the same rule in phase 3 or shifted to a different rule. Figure 3 shows that when participants who initially used a 1D or 2D rule observed approach/avoid patterns consistent with their initial rule, they continued to use that rule in the final phase. By the end of this phase nearly all (93.3%) of those who used a 1D rule in phase 1 and observed the same rule during social learning, continued to use that rule. For those who used a 2D rule in phase 1 and observed a consistent rule, the corresponding proportion was 100%.

Observing someone using an inconsistent rule in the social learning phase had different effects depending on whether learners initially acquired 1D or 2D rules. Figure 3 shows that for those who started with a 1D rule, observation of a demonstrator using a 2D rule led nearly 40% of this sub-group to discover and use the optimal rule in phase 3. In contrast, those who started with a 2D rule showed little change in rule use after observation of a demonstrator using a 1D rule. Those who used an unclassified rule in phase 1 (not shown in the Figure) often shifted in the direction of observed rule (48% of those who observed a 1D rule used this rule in the final phase; 49% of those who observed a 2D rule used this rule in the final phase). The remainder of this sub-group continued to use an unclassified rule at the end of Phase 3.

These trends were supported by multinomial regression analyses.[1] For those who used a 1D rule in Phase 1, prediction of final rule use was significantly increased by adding social learning condition (consistent/inconsistent) into the equation, $\chi^2(1, N=56) = 18.631$, $p < 0.001$, as compared to an intercept-only model. Regression parameters showed that 1D users who observed an optimal 2D rule were more than 19 times more likely to use a rule that was either 2D or unclassified rule in the final learning block compared to those who viewed a 1D rule (Odds Ratio $OR = 19.231$, $\beta = 2.949$, $SE = 0.833$, $z = 12.546$, $p < .001$). Adding social learning to the regression equation also improved prediction of final rule use in unclassified rule users, χ^2 (1, N=58) = 8.035, $p = .018$. Those in this sub-group who observed a 1D rule were over 4 times more likely to adopt a 1D rule over 2D in the final block compared to those who observed a 2D rule ($OR = 4.274$, $\beta = 1.451$, $SE = 0.676$, $z = 4.607$, $p = .032$).

Changes in rule use following social learning were presumed to be at least partially mediated by changes in levels of exploration of category exemplars. This was supported by an analysis of approach decisions for initial 1D or 2D learners who observed consistent or inconsistent rules. For those with an initial 1D rule, the mean proportion of exemplars approached in phase 3 was higher following observation of an inconsistent rule ($M = 0.638$) as compared to a consistent rule ($M = 0.498$), $F(1, 54) = 37.144$, $p < .001$, $\eta^2_p = 0.408$. In contrast, those with an initial 2D rule approached more stimuli throughout the task ($M = 0.756$) and showed no significant change in approach following consistent ($M = 0.749$) or inconsistent social learning ($M = 0.734$), $F(1, 58) = 1.620$, $p = .208$.

In sum, for those who were in a one-dimensional learning trap, observing a learner following an optimal two-dimensional rule led many to escape from the trap. Social learning led to higher rates of approach in the final phase

[1] Because all phase 1 2D users observing an inconsistent rule used a 2D rule at the end of phase 3, regression analysis could not be carried out for this subgroup. In the regression analysis for initial 1D users, those using a 2D or unclassified rule at the end of phase 3 were combined to eliminate empty cells.

which increased the learner's chance of discovering the optimal rule. In contrast, those who were already using an optimal rule were unaffected by observing someone else using a one-dimensional rule.

Experiment 2

The previous study found that those in a one-dimensional learning trap could benefit from observing the approach patterns of another learner using an optimal two-dimensional rule. This study examined whether a similar outcome could be achieved by observing another learner using a different but *sub-optimal* rule. The general design was similar to the previous study except that, in the "inconsistent" condition, one-dimensional learners observed approach patterns of a demonstrator using the complementary one-dimensional rule 1D' (e.g., if the learner initially learned a 1D rule that focused on dimension 1, they saw a demonstrator using a 1D rule that focused on dimension 2). Although a 1D' rule is suboptimal, it necessarily involves approaching some rewarding stimuli that a 1D learner would avoid (as well as avoiding some stimuli the 1D learner is currently approaching). This could prompt the learner to increase their exploration and discover the optimal rule. Based on the results of the previous study, we expected that those who initially learned a two-dimensional rule would be unaffected by observation of a demonstrator using either version of the 1D rule.

Method

Participants. The participants were 150 adults (M_{age} = 29.3 years, 79 males, 71 females), recruited from UNSW undergraduates (n = 65) or the online platform Prolific (n = 85). Monetary compensation for participation was unchanged from Experiment 1. No participants were excluded.

Design and Procedure. Experiment 2 followed the same three-phase design and procedure as the previous study except that, in phase 2, those who had initially acquired a one-dimensional rule were randomly assigned to observe a demonstrator whose approach/avoid patters followed the same 1D rule (consistent group) or the alternative 1D' rule (inconsistent group) – see Figure 4.

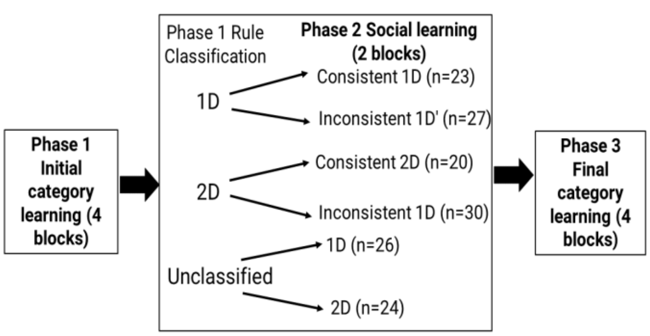

Figure 4. Summary of Experiment 2 design

Results and Discussion

The rules that participants in each sub-group used in phases 1 and 3 are shown in Figure 5. As in the previous experiment, most of those using a 2D rule in the initial phase continued to use this rule at end of the final phase, regardless of what was observed in the social learning phase. A multinomial logistic regression found that prediction of final rule use for this sub-group was unchanged by adding by social learning condition was added to the equation, $\chi^2(1, N = 50) = 0.085$, p = .770.

Crucially, for those who had fallen into the 1D trap in phase 1, the evidence observed during social learning had a marked effect on subsequent rule use. Observing rule-consistent evidence perpetuated use of the 1D rule but observing someone using an alternative 1D' rule led around 30% of participants to discover the optimal rule by the end of phase 3. Prediction of final rule use was improved in a multinomial logistic regression by adding social learning conditions to the equation, $\chi^2(1, N = 50) = 8.87$, p = .003. Those who observed a demonstrator using the alternative 1D' rule were over 12 times more likely to use either a 2D or unclassified rule (predominantly the former) in the final learning block as compared with those who observed the same 1D rule (OR = 12.99, β = 2.56, SE = 1.097, z = 5.544, p = .020.

We again observed a tendency for those whose rule use was unclassified in phase 1 to shift towards whichever rule was observed during social learning (42% of those who observed a 1D rule used a 1D rule in the final phase; 33% of those who observed a 2D rule used a 2D rule in the final phase). Unclassified participants shown 1D information in Phase 2 were over 29 times more likely to adopt a 1D rule by the end of the experiment compared to those shown 2D information (OR = 29.41, β = 3.379, SE = 1.25, z = 7.37, p = .007).

As in the previous study, for 2D learners the type of evidence observed did not affect the proportion of exemplars that were approached in phase 3 (consistent 2D rule: M = 0.746; inconsistent 1D/1D' rule: M = 0.752), $F(1, 48) = 0.970$, p = .330, η^2_p = 0.02. For 1D learners, observing the 1D' rule increased mean exploration in phase 3 (M = 0.576) as compared to those who observed the same rule (M = 0.506), $F(1, 48) = 8.125$, p = 0.006, η^2_p = 0.145.

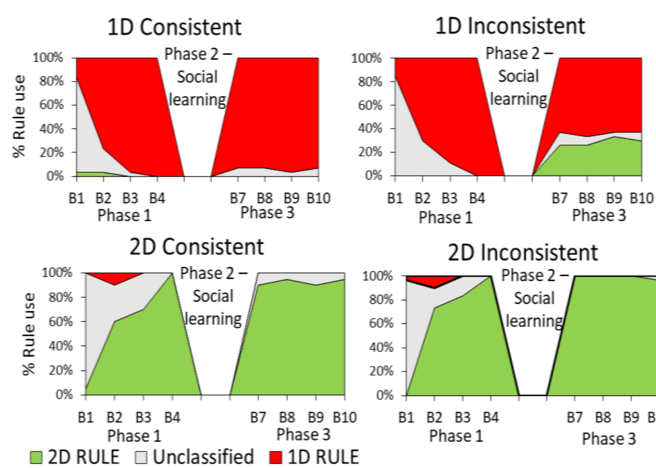

Figure 5. Exp. 2: % of participants using various category rules in Phases 1 (B1-B4) and 3 (B7-B10) in each social learning condition. Only those using a 1D or 2D rule in the final block of Phase 1 are shown.

General Discussion

Learning traps are pernicious cycles of decision-making that can lead to missed rewards and that are resistant to change. The current studies examined whether social learning, in the form of observation of approach and avoid decisions made by another learner, offers an escape route for those in a learning trap. We found that many of those who had fallen into the trap of using an overly simplistic one-dimensional rule, and who had the opportunity to observe approach and avoid patterns consistent with an optimal decision-rule, shifted towards that rule in their subsequent decision-making. A smaller but nevertheless positive shift was found when trapped learners observed another using an alternative suboptimal rule. In both cases, observing the alternative rule appeared to prompt exploration of previously avoided stimuli, disconfirming the false belief that approaching these stimuli would lead to a loss. In contrast, those who learned the optimal decision rule in the initial learning phase were unaffected by social learning. They continued to use the optimal rule to guide subsequent decision-making regardless of whether they observed another learner using the same or a different sub-optimal rule.

These results are noteworthy for a number of reasons. They document one of the first successful attempts to counteract the negative effects of learning traps. Previous attempts have either been unsuccessful (e.g., Rich & Gureckis, 2018) or have had modest success in preventing traps from occurring (Watson et al., 2024). In contrast, the current work shows that some forms of social learning can lead to *escape* from an established trap.

One recent study (Budiono et al., 2024) also examined the effects of social learning on learning traps. Their approach was broadly similar to ours, in that learners observed others using different decision-rules. One important difference was that learners observed their partner's decisions trial-by-trial, in tandem with making their own decisions (e.g., they saw an exemplar, decided whether to approach or avoid, and then saw what the partner's decision was). Social observation of an alternate rule was generally beneficial, but the observed positive effects were considerably smaller than those observed in the current studies. We speculate that this was due to our learners having the opportunity to observed blocks of decision-making trials in the social learning phase, which made it easier for learners to infer the demonstrator's underlying decision-rule.

The current results add to the growing body of evidence that social learning can benefit decision-making in complex environments (cf. Hawkins, et al., 2023; Toelch, 2013; Toyokawa et al., 2014, 2019; Wu et al., 2022). In particular, these results reaffirm that social observation of another's pattern of decisions can have a positive effect without the outcome of those decisions being revealed.

Our results suggest that learners were generally emulating rather than imitating their social partner. Learners were highly selective in whether they shifted towards the approach patterns observe in the social learning phase. They shifted when the observed pattern differed from their current rule, and held out the prospect of discovery of new information and rewards. However, when the observed pattern involved avoiding exemplars that learners had already approached and found to be rewarding, the observations were ignored. This selective responding also suggests that our social learning results cannot simply be explained as a demand effect (e.g., the learner assuming that the experimenter is providing helpful advice in phase 2 that they should follow).

Our results may reflect a "copy when uncertain" strategy, identified in previous studies of the impact of social learning on decision-making (e.g., Toelch et al., 2013). According to this approach, learners are most likely to shift towards decision patterns they observe in social learning when they are unsure about the optimality of their current decision rule. This notion could be tested directly in future work by asking leaners to rate the confidence in their decision rule at the end of phase 1. We predict that those using a 1D-rule will be generally less confident than those using a 2D rule. Within the sub-group of 1D rule users, those with the lowest confidence will be most likely to be affected by social observation of an alternative rule.

The current work has focused on the positive effects of social learning on boosting the learning of optimal decision rules and reducing the prevalence of learning traps. However, there are likely to be some circumstances where social learning can have the opposite effect – encouraging formation of incomplete or biased beliefs about reward structure and overly simplistic decision rules. In the current task it is possible that, for those who initially acquired a 1D rule, observation of a demonstrator using the same rule may have discouraged further stimulus exploration and reduced the chances of escaping from a trap. Although this cannot be ruled out, examination of previous data sets from similar tasks without social learning suggests that it is unlikely. Lee et al. (2024), for example, found that very few of those who fell into a one-dimensional trap in the first few learning blocks (<5%), subsequently escaped when given extensive additional learning opportunities.

Social learning may be more likely to have a negative impact on trap formation when it occurs before learning commences. Observation of a demonstrator making approach/avoid decisions consistent with a one-dimensional trap may lead to a subsequent reduction in stimulus exploration (cf. Spektor & Wulff, 2024). Such negative effects are likely to be accentuated if the demonstrator is perceived to have more expertise with the task than the learner (Gweon, 2021).

In sum, we have shown that observation of approach and avoidance patterns that are inconsistent with an existing suboptimal rule, can lead to escape from a learning trap and discovery of an optimal decision rule. Future work is needed to further examine the attentional and learning mechanisms that underlie such shifts. Another important question for the future is whether such positive shifts require social observation per se or whether similar results could be obtained with a non-social demonstrator (e.g., an algorithmic "advisor system", cf. Strickland et al., 2023).

References

Bai, X., Fiske, S. T., & Griffiths, T. L. (2022). Globally inaccurate stereotypes can result from locally adaptive exploration. *Psychological Science, 33*(5), 671-684.

Blanco, N. J., & Sloutsky, V. M. (2019). Adaptive flexibility in category learning? Young children exhibit smaller costs of selective attention than adults. *Developmental Psychology, 55*(10), 2060 - 2076.

Budiono, R., Hartley, C., & Gureckis, T. M. (2024). How does social learning affect stable false beliefs? In Samuelson, L.K., Frank, S.L., Toneva, M., Mackey, A. & Hazeltine, E. (eds.). *Proceedings of the Annual Meeting of the Cognitive Science Society, 46*.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods, 47*(1), 1-12. https://doi.org/10.3758/s13428-014-0458-y

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in iImpression formation. *Psychological Review, 112*(4), 951–978. https://doi.org/10.1037/0033-295X.112.4.951

Denrell, J., & March, J. G. (2001). Adaptation as Information Restriction: The Hot Stove Effect. *Organization Science, 12*(5), 523-538. https://doi.org/10.1287/orsc.12.5.523.10092

Elwin, E. (2013). Living and learning: Reproducing beliefs in selective experience. *Journal of Behavioral Decision Making, 26*(4), 327-337.

Forbes magazine (2023, June 9). *Investing in renewable energy for a sustainable future*. https://www.forbes.com/sites/investor-hub/article/investing-in-renewable-energy-for-a-sustainable-future/

Gweon, H. (2021). Inferential social learning: Cognitive foundations of human social learning and teaching. *Trends in Cognitive Sciences, 25*(10), 896-910. https://doi.org/10.1016/j.tics.2021.07.008

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

Lee, W. J., Li, A. X., Lee, J. E., & Hayes, B. K. (2024). Learning traps and change blindness in dynamic environments. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 50 (9) 1345-1360. https://doi.org/10.1037/xlm0001390

Li, A., Gureckis, T., & Hayes, B. K. (2021). Can losses attenuate learning traps? In T. Fitch, T. Fitch, C. Lamm, H. Leder & K. Tessmar-Raible (Ed.s) *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society*, pp. 1201-1207

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

Liu, Y., Newell, B., Lee, J. E., & Hayes, B. (2024). Examining the Relationship Between Selective Attention and the Formation of Learning Traps. In Samuelson, L.K., Frank, S.L., Toneva, M., Mackey, A. & Hazeltine, E. (eds.). *Proceedings of the 46th Annual Meeting of the Cognitive Science Society, 46*, pp. 2145-2151

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of over-imitation. *Proceedings of the National Academy of Sciences, 104*(50), 19751-19756. https://doi.org/10.1073/pnas.0704452104

Morin, O. (2016). *How traditions live and die*. Oxford: Oxford University Press.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., ... & Laland, K. N. (2010). Why copy others? Insights from the social learning strategies tournament. *Science, 328*(5975), 208-213. doi:10.1126/science.1184719

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

Spektor, M. S., & Wulff, D. U. (2024). Pre-decisional information search adaptively reduces three types of uncertainty. *Proceedings of the National Academy of Sciences, 121*(47), e2311714121.

Strickland, L., Boag, R. J., Heathcote, A., Bowden, V., & Loft, S. (2023). Automated decision aids: When are they advisors and when do they take control of human decision making? *Journal of Experimental Psychology: Applied, 29* (4), 849-868.

Toelch, U., Bruce, M. J., Newson, L., Richerson, P. J., & Reader, S. M. (2014). Individual consistency and flexibility in human social information use. *Proceedings of the Royal Society B: Biological Sciences, 281*:20132864. https://doi.org/10.1098/rspb.2013.2864

Toyokawa, W., Kim, H. R., & Kameda, T. (2014). Human collective intelligence under dual exploration-exploitation dilemmas. *PloS one, 9*(4), e95789. https://doi.org/10.1371/journal.pone.0095789

Toyokawa, W., Whalen, A. & Laland, K.N. (2019). Social learning strategies regulate the wisdom and madness of interactive crowds. *Nature Human Behavior, 3*, 183–193 https://doi.org/10.1038/s41562-018-0518-x

Watson. P., Lee, W. J., Lee, J. E., Liu, Y., Newell, B. R., & Hayes, B. K. (2024). *Trapped by selective attention: The role of attentional processes in the emergence and prevention of learning traps*. https://osf.io/preprints/psyarxiv/7fx4pprint

Whalen, A., Griffiths, T. L., & Buchsbaum, D. (2018). Sensitivity to shared information in social learning. *Cognitive Science, 42*(1), 168-187.

Wu, C. M., Vélez, N., & Cushman, F. A. (2022). Representational exchange in human social learning: Balancing efficiency and flexibility. In Wu, C. M., Schulz, E. & Dezza, I. C. (ed.s) *The drive for knowledge: The science of human information seeking*. Cambridge:

Cambridge University Press.
https://doi.org/10.31234/osf.io/rm52c
Xie, B. & Hayes, B. K. (2022). Sensitivity to evidential dependencies in judgments under uncertainty, *Cognitive Science*, *46*(5), e13144,