

## **UC Merced**

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

### **Title**

Some forms of uncertainty may suppress the evolution of social learning

### **Permalink**

<https://escholarship.org/uc/item/97q8w74p>

### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

### **Authors**

Turner, Matthew Adam

Moya, Cristina

Smaldino, Paul

et al.

### **Publication Date**

2022

### **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

# Some forms of uncertainty may suppress the evolution of social learning

Matthew A. Turner (maturner@stanford.edu)<sup>1</sup>, Cristina Moya (moya@ucdavis.edu)<sup>2,3</sup>,  
Paul E. Smaldino (psmaldino@ucmerced.edu)<sup>4,5</sup>, and James Holland Jones (jhj1@stanford.edu)<sup>1,5</sup>

<sup>1</sup>Department of Earth System Science, Stanford University, Stanford, CA 94305 USA

<sup>2</sup>Department of Anthropology, University of California at Davis, Davis, CA 95616 USA

<sup>3</sup>Centre for Culture and Evolution, Brunel University London, Uxbridge, UB8 3PH UK

<sup>4</sup>Cognitive and Information Sciences, University of California at Merced, Merced, CA 95340 USA

<sup>5</sup>Center for Advanced Study in the Behavioral Sciences, Stanford University, Stanford, CA 94305 USA

## Abstract

Social learning is essential to survival. It is likely to evolve when it is more efficient than asocial, trial-and-error learning. The consensus in cultural evolutionary theory holds that some amount of environmental variability and uncertainty about the best decisions are necessary for social learning to evolve. However, current models for the evolution of social learning tend to conflate forms of uncertainty, and rarely consider different ones in tandem. Moreover, many models are limited by considering only two possible behaviors and environmental states. Here we use evolutionary agent-based modeling to identify the complex ways in which different forms of uncertainty affect social learning. We model a time-varying environment with dozens of possible behaviors performed by agents engaging in individual and social learning. We show that ambiguous payoffs, larger possible decision sets, and shorter agent lifespans sometimes increase social learning prevalence, as expected. However we also find which concrete uncertainty conditions cause evolution to select against social learning.

**Keywords:** anthropology; evolution; learning; social cognition; agent-based modeling

## Introduction

Social learning is essential to human and other species' everyday life and survival. It allows individuals to solve problems when acquiring information from others is more efficient than learning on one's own (Laland, 2004). Theory predicts that social learning should be favored in contexts with greater uncertainty (Boyd & Richerson, 1985; Henrich & Boyd, 1998), and this prediction has received some empirical support (McElreath et al., 2005; Kendal et al., 2018). However, the meaning of the term "uncertainty" is not always clear, and often conflates environmental variability, spatial heterogeneity, and ambiguity or uncertainty about payoff structure. Moreover, most models of the evolution of social learning blackbox key cognitive learning processes that underlie it (Heyes, 2016).

In this paper we use agent-based modeling to compare the effect of different sources of uncertainty on social learning by un-blackboxing typically abstracted-out model components of environmental variability, payoff structures and agent life histories, and learning mechanisms. *Uncertainty* here means variability where the probabilistic structure is unknown. Uncertainty increases when payoffs are more similar across behaviors,

when environmental variability increases, when the number of possible behaviors increases, and when lifespan decreases. In this paper we show that more ambiguous payoff structures and shorter lifespans sometimes do lead to greater reliance on social learning—however, we also identify and explain cases where greater uncertainty leads to less social learning due to the possibility that social information is misleading. Here we show that different sources of uncertainty interact in complex ways to affect the evolution of social learning. We thus conclude that many predictions made by previous models of the evolution of social learning are likely overgeneralized.

## Social Learning

Social learning, as we consider it here, occurs whenever an individual acquires a behavior by observing another individual. This need not require explicit instruction, and is, in fact, widespread across a broad range of non-human taxa (Kendal et al., 2018; Allen, 2019). Importantly, social information can be inherited both from parents — i.e., via *vertical transmission* like genetic information — and from others in the same generation — i.e., via *horizontal transmission* (Cavalli-Sforza & Feldman, 1981). The joint action of vertical and horizontal transmission gives rise to qualitatively different evolutionary dynamics. For example, inter-generational environmental change will affect the adaptive value of genetic information and vertically-transmitted cultural information more than information that is horizontally transmitted. We include both horizontal and vertical transmission pathways in our model. For simplicity, we ignore oblique transmission in which non-parental members of the previous generation are observed.

Environmental variability has been seen as a key selective force in shaping social learning starting with the first formal models of cultural evolution (Cavalli-Sforza & Feldman, 1981; Boyd & Richerson, 1985). Totally stable environments will not favor learning mechanisms because information can become genetically hardwired, while extreme environmental instability will degrade the value of social learning as information becomes rapidly outdated (Feldman, Aoki, & Kumm, 1996). This suggests that an intermediate degree of environmental predictability will favor social learning. Strategies can also

evolve to mitigate the risks of relying on outdated social information by weighing more heavily information from others who more recently acquired it (Rendell et al., 2010).

Uncertainty has also been modelled as arising from other aspects of the environment. For example, Perreault, Moya, and Boyd (2012) vary the ambiguity of the environmental cue individuals get through individual learning about the state of the world. Perhaps not surprisingly, the more ambiguous the social information, the greater the selection for weighing social information heavily. Alternatively, uncertainty about the optimal behavior has been modelled by increasing the number of cultural traits to choose from (Nakahashi, Wakano, & Henrich, 2012; Muthukrishna, Morgan, & Henrich, 2016).

Empirical research supports some of these theoretical predictions. Organisms flexibly use social learning as a function of the ambiguity of the environmental cue and of other environmental features that are often subsumed under the rubric of uncertainty. While some studies explicitly impose a cost whenever participants use social information (Morgan, Rendell, Ehn, Hoppitt, & Laland, 2012; Atkisson, O’Brien, & Mesoudi, 2012), others allow the costs of each strategy to emerge as a function of task structure and assess its consequences for learning strategies. For example, when participants received equivocal private information about the best investment to make in a lab game, they were more likely to rely on social information to make their choice (Toelch, Bruce, Newson, Richerson, & Reader, 2013).

McElreath et al. (2005) developed a similar experiment where participants “pulled” virtual slot machine arms (often called “bandits”), each yielding stochastic payoffs. Participants relied more on social learning when the bandits had higher-variance payoffs, and when the highest-paying bandit changed more frequently. The number of options to choose between can also increase uncertainty about the optimal choice, and has been shown to increase participants’ reliance on social learning (Muthukrishna et al., 2016).

Thus existing theoretical work and the empirical evidence seem to support that various forms of uncertainty favor the evolution of social learning. However, uncertain outcomes are operationalized in different ways across models, and any given model tends to focus on only one or two forms of uncertainty at a time. Our agent-based modeling approach enables us to explicitly specify different forms of uncertainty independently in order to understand which of these environmental factors particularly favor the evolution of social learning. Simultaneous modelling also allows us to examine their interaction. We first attempt conceptual replications of previous models’ findings, and then examine where they diverge.

## Research overview

Computational agents in our model face a problem: every time step they perform one of several behaviors, with each behavior represented by a “bandit” that pays off 1 or 0 with some probability. One of these behaviors (the optimal behavior) pays off with a higher probability than all the others. Agents decide which behavior to perform based either on a success-biased observation of a peer’s behavior (social learning) or based on their own observations (asocial learning). Agents then update their memory of expected payoffs for each behavior when they receive a payoff from their chosen action. Within this framework we have four mechanisms by which we operationalize and vary uncertainty: (1) the expected payoffs of the optimal behavior and all the rest—when payoffs are nearly identical, uncertainty in the form of ambiguity increases; (2) the environmental variability, i.e., the probability that the optimal behavior changes between generations; (3) the number of possible behaviors the environment allows—which behavior is optimal is more uncertain when there are more possibilities; and (4) agent lifespan—agents experience fewer learning opportunities and die more uncertain about which behavior is optimal when their lifespan is shorter.

The primary outcome measure of our model is the average difference between the frequency of (horizontal) social learning and the frequency of asocial learning across all agents. If social learning is more prevalent than asocial learning this suggests that the optimal behavior is more likely found by copying peers than by trial-and-error search. Conversely, when asocial learning is more prevalent, this suggests social information is likely to be misleading. When social and asocial learning are equally prevalent this means there is no discernible advantage to either, i.e., the agents have weak priors on which channel provides more reliable information. Each agent has its own social learning frequency that it inherits from its parent (haploid reproduction) with mutation, so evolution selects for, or computes (Smaldino & Richerson, 2013), the optimal social learning frequency.

Using our model we found that increased uncertainty sometimes led to increased reliance on social learning, as expected from prior literature. However, we also find cases where increased uncertainty decreased agents’ reliance on social learning due to increased uncertainty that made social information less reliable, thereby increasing reliance on asocial learning.

## Model

We developed an agent-based model of a society of  $N$  individuals who each must decide which of  $B$  behaviors to perform at each time step. Each behavior is a bandit indexed by integer  $b$ , a common modelling and experimental approach for representing behaviors with probabilistic payoffs (Sutton & Barto, 2018; McElreath

et al., 2005; Rendell et al., 2010; Schulz & Gershman, 2019). Each behavior indexed by  $b$  is modeled here as a “pull” of bandit  $b$  (Figure 1, left). Payoffs from each behavior  $b$  are binomially distributed with mean  $\pi_b$ , and yield a payoff of 1 when the behavior is successful and 0 otherwise. Agents must decide which behavior to perform at each time step. To do this, each agent  $i$  employs either social information with probability  $s_i$  or asocial information with probability  $a_i = 1 - s_i$  (Figure 1).

We operationalized uncertainty in four different ways: (1) payoff ambiguity,  $A_\pi$ , which measures the difference between the optimal expected payoff behavior  $\pi_{\text{high}}$  and the expected payoff of the other behaviors,  $\pi_{\text{low}}$ ; (2) environmental variability,  $u$ , the probability the optimal behavior changes from one generation to another; (3) the number of possible behaviors,  $B$ ; and (4) the lifespan, or time steps per generation,  $L$ .

We ran the simulation for  $R$  rounds (i.e., generations of  $L$  time steps each). At the end of each round agents reproduce and then die off. Those selected to reproduce pass on their social learning frequency  $s_i$  with mutation, so that selection favors higher payoff strategies at reproduction. We developed a series of computational analyses where we systematically vary the uncertainty and risk variables and observe the social-asocial learning difference, our primary outcome measure,  $\langle s - a \rangle$ . In other words, we examine when *horizontal* social learning is more likely to evolve than asocial learning.

## Agents

In each time step,  $N$  agents—autonomous problem solvers—select which behavior to perform based on either social or asocial information (Figure 1). In either case the agent tracks the mean payoff of each behavior  $b$ , denoted  $\bar{\pi}_b$ , and a *count* of how many times it has performed each behavior, denoted  $c_b$ . Agents’ beliefs about mean payoffs are updated from  $\bar{\pi}_b$  to  $\bar{\pi}'_b$  using exponential weighted averaging,  $\bar{\pi}'_b = \bar{\pi}_b + \frac{\text{Bandit}_b(0,1) - \bar{\pi}_b}{c'_b}$ , where  $\text{Bandit}_b(0,1)$  is 0 or 1 depending on the result of the bandit draw for behavior  $b$ .  $\bar{\pi}_b$  is initialized to 0 for all  $b$  at model initialization, and initialized based on vertical transmission from each new agent’s parent at the start of each generation, after the first, described in more detail below (see Table 2 for a summary of agent-level variables).

Agent  $i$  chooses social information with probability  $s_i$ , which is itself an inherited trait (see below for details on evolutionary dynamics). When an agent engages in social learning it first selects  $N_T$  potential teachers at random from the other  $N - 1$  agents in the population ( $N_T = 3$  in the example at the bottom of Figure 1, and is set to  $N_T = 10$  for our analysis). By setting  $N_T = 10$ , which is less than the total number of agents  $N$ , we make a conservative estimate of the scope of social learning since in real-world settings individuals do not always

Table 1: **Agent variables.** Agents have social learning and vertical transmission traits  $v_i$  and  $s_i$ ; individual-level dynamic variables tracking agent observations,  $c_i$  and  $\pi_i$ , and global parameter  $\tau$ . The subscript  $i$  is included when each agent  $i$  has its own local value of that variable. See text for variable initializations.

Variable	Description
$s_i$	Frequency of horizontal social learning
$v_i$	Propensity for vertical social learning
$c_i$	Vector length $B$ counting number of times each behavior was performed
$\pi_i$	Vector length $B$ of observed mean payoffs for each behavior
$\tau$	Softmax parameter that tunes how often agents exploit the best observed behavior (lower $\tau$ ) versus explore alternatives (higher $\tau$ )

have access to the whole population as teachers at any given moment. The agent selects one of the teachers of the  $N_T$ , weighted by each potential teacher’s accumulated payoffs in the round (Figure 1, bottom right). The agent then performs the behavior done by its chosen teacher in the previous time step.

Following Luce’s choice axiom (Luce, 1959), an agent using asocial information selects a behavior at random, with the probability of selecting any particular behavior weighted by the softmax function applied to that behavior’s observed mean payoff relative to all mean payoffs:

$$\Pr(\text{choose behavior } b) = \frac{\exp(\bar{\pi}_b/\tau)}{\sum_{b=1}^B \exp(\bar{\pi}_b/\tau)}. \quad (1)$$

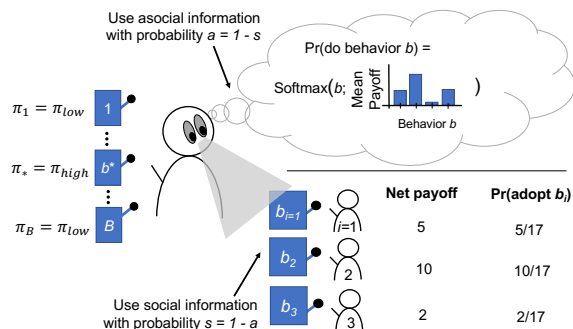
The softmax method is a biologically plausible model (Hill, Boorman, & Fried, 2016; Schulz & Gershman, 2019) that enables agents to explore alternative behaviors sometimes and exploit the best observed behavior other times, even for small sample sizes (Vul, Goodman, Griffiths, & Tenenbaum, 2014). The parameter  $\tau$  specifies how frequently alternative behaviors are explored—higher  $\tau$  means more frequent exploration. To mimic increased plasticity and exploration in the developmental period we use simulated annealing to decrease  $\tau$  over the course of a model generation (Gopnik, 2020), starting from  $\tau = 0.01$  and reducing  $\tau$  by  $9.999\text{e-}5$  each time step so that  $\tau = 1\text{e-}6$  after 100 time steps.

## Modeling uncertainty

In our model uncertainty is a tunable consequence of four environmental features:

1. We vary the latent expected payoffs yielded by the bandits. For simplicity, we assume that in any given environmental state, there is one optimal behavior

Figure 1: **Behavior selection.** Agents select a behavior to do at each timestep either using asocial information (top row) with probability  $a$  or social information with probability  $s = 1 - a$  (bottom row).



that yields an expected payoff of  $\pi_{\text{high}}$ , while all other behaviors yield a payoff of  $\pi_{\text{low}} < \pi_{\text{high}}$ . The difference between these quantities is the *payoff ambiguity*,

$$A_\pi = 1 - (\pi_{\text{high}} - \pi_{\text{low}}), \quad (2)$$

which is maximized when the expected payoffs are close to equal and minimized when they differ greatly.

2. We vary the frequency at which the optimal behavior changes, which we call *environmental variability*,  $u$ . At the start of each generation, with probability  $u$ , a new behavior is assigned a payoff of  $\pi_{\text{high}}$  while all other behaviors are assigned a payoff of  $\pi_{\text{low}}$ . Otherwise, the same behavior remains optimal across generations.
3. We vary the total *number of available behaviors*,  $B$ , which is a source of uncertainty since agents are less likely to know which behavior yields high payoffs.
4. Finally, we vary number of behavioral events per generation,  $L$ . This can be viewed as the *effective lifespan* of an agent. Decreasing this lifespan effectively increases the importance of each event for acquiring payoffs. When lifespans are short ( $L$  is small), agents experience greater uncertainty about which behavior is optimal given the fewer learning opportunities.

## Dynamics and evolution

Model dynamics occur on two timescales. The shorter scale is the individual behavioral event, in which horizontal social learning occurs and payoffs are accrued. There are  $L$  such events per generation. The longer scale is the intergenerational one. At the end of each generation's  $L$  time steps, agents reproduce before dying.  $N$  new offspring are then created, and assigned to a parent with probability proportional to the potential parents' accumulated payoffs. Each of these offspring inherits two traits from its parent  $i$ —the propensity for horizontal social learning frequency,  $s_i$ , and the magnitude of vertical transmission,  $v_i$ .

Table 2: **Uncertainty variables.** We model uncertainty with these four variables, explained in the text.

Variable	Description
$A_\pi$	Payoff ambiguity $1 - (\pi_{\text{high}} - \pi_{\text{low}})$ , which is maximal when $\pi_{\text{high}} = \pi_{\text{low}}$
$u$	Environmental variability, i.e., probability the optimal behavior switches between generations
$B$	Number of possible behaviors; only one is the optimal behavior
$L$	Agent life span, i.e., number of behaviors performed before reproduction and die-off

An offspring  $j$  inherits traits related to social learning from its parent,  $i$ , with mutation, such that  $s_j = s_i + \epsilon$  and  $v_j = v_i + \epsilon$ , where  $\epsilon$  is a random draw from a normal distribution with a mean of zero and a standard deviation of  $\sigma$  (we used  $\sigma = 0.05$  for all simulations reported here). All mutated traits were truncated in  $[0, 1]$ .

This propensity for vertical social learning affects how much offspring weight the information they learn about their parents' behavioral repertoire. In the absence of vertical transmission, offspring initially assign their expected payoffs for each behavior  $b$  to the mean value over all payoffs in the parent's memory,  $\bar{\pi}$ . With vertical transmission, offspring  $j$  learns about each behavior's expected payoff from its parent:  $\bar{\pi}'_b = \bar{\pi} + v_i(\bar{\pi}_b - \bar{\pi})$ . Here  $\bar{\pi}'_b$  is the mean payoff for behavior  $b$  that is transmitted to the child,  $\bar{\pi}_b$  is the expected payoff for behavior  $b$  in the parent's memory, and  $\bar{\pi}$  is the mean payoff across all behaviors observed by the parent. The number of observations of behavior  $b$  known to the child, call it  $c'_b$ , is set conservatively based on its parent  $p$ 's behavior observation count  $c_b$  as follows under vertical transmission  $c'_b = \min(\lfloor v_p c_b \rfloor, 1)$ . This means that if  $\lfloor v_p c_b \rfloor = 0$ , then  $c'_b$  is set to 0. If  $\lfloor v_p c_b \rfloor \geq 1$ , then  $c_b$  is set to 1. This keeps children from inheriting a high  $c_b$  that would cause new information to be weighted very little.

## Computational analyses

We manipulated environmental uncertainty parameters described above,  $A_\pi$  (via  $\pi_{\text{low}}$  and  $\pi_{\text{high}}$ ),  $u$ ,  $B$ , and  $L$ , to examine their effects on our main outcome, the difference between the probabilities of social and asocial learning  $\langle s - a \rangle$ . For each parameter setting in our analysis we calculated the average value of  $\langle s - a \rangle$  over the final 20,000 of  $T = 100,000$  behavioral events (i.e., time steps) for each simulation, across 100 runs for each combination of parameter values.  $\langle s - a \rangle$  is dynamically stable over this time frame for all analyses, which we demonstrate in time series plots in a supplemental Jupyter notebook included in the GitHub repository for this project (<https://github.com/mt-digital/UncMod/blob/CogSci2022-CameraReady/notebooks/>

Final%20CogSci%20results.ipynb).

In this paper, we highlight the results from four computational analyses. In the first analysis, we manipulated the base payoff  $\pi_{\text{low}} \in \{0.1, 0.2, \dots, 0.8\}$  and the optimal payoff  $\pi_{\text{high}} \in \{0.2, 0.3, \dots, 0.9\}$ . In the second analysis, we manipulated the environmental variability alone with  $u \in \{0.0, 0.1, 0.25, 0.5, 0.75, 0.9\}$ . In the third analysis, we manipulated the number of possible behaviors,  $B \in \{10, 20, \dots, 100\}$ , along with environmental variability  $u \in \{0.1, 0.9\}$ . In the final analysis, we manipulated the lifespan of agents, the number of steps per generation,  $L \in \{5, 10, 20, 50, 100\}$  and the environmental variability  $u \in \{0.1, 0.9\}$ .

The model was implemented in Julia (Bezanson, Edelman, Karpinski, & Shah, 2017) using the Agents.jl package for model development (Datseris, Vahdati, & DuBois, 2022) and the Gadfly package for figures (Jones, Arthur, Nagy, Matriks, & Contributors, 2021). Simulations were run on the Sherlock supercomputing cluster at Stanford University. Model and analysis code, and links to output datasets used for this paper’s figures, can be found on GitHub at <https://github.com/mt-digital/UncMod/tree/v0.1-CogSci2022Sub>.

## Analysis

### Payoff ambiguity $A_\pi$ (via $\pi_{\text{low}}$ and $\pi_{\text{high}}$ )

Payoff ambiguity,  $A_\pi$ , is the first form of uncertainty we analyze. To do this, we manipulated the base payoffs  $\pi_{\text{low}}$  and optimal payoff  $\pi_{\text{high}}$ , and by consequence varied payoff ambiguity,  $A_\pi$  (Equation 2). We found some cases supporting the claim that greater uncertainty in the form of payoff ambiguity leads to more social learning—or its logical equivalent, that less ambiguity leads to less social learning. However we also found several counterexamples where increased (decreased) uncertainty of this form suppresses (selects for) social learning.

First, we found cases that support the claim that greater uncertainty leads to more social learning. For example, consider the case where  $u = 0.1$  (Figure 2a) and  $\pi_{\text{low}} \approx 0.1$ . In this case social learning becomes less prevalent as  $A_\pi$  decreases when moving from parameter setting  $\pi_{\text{high}} \approx 0.6$  to  $\pi_{\text{high}} = 0.9$ , as classically expected.

On the other hand, we also found cases where social learning is selected for as ambiguity decreases, contrary to expectations. For example, consider the low environmental variability ( $u = 0.1$ ) case with  $\pi_{\text{high}} = 0.9$ . In this case  $\langle s - a \rangle$  decreases while  $\pi_{\text{low}}$  increases to  $\approx 0.5$ , which correspondingly increases  $A_\pi$ . When  $\pi_{\text{low}}$  continues to increase from  $\approx 0.5$  to  $0.8$ , the trend switches back to where social learning prevalence increases with payoff ambiguity.

Social learning sometimes increases when ambiguity decreases because horizontally-acquired information is more reliable than information either asocially acquired through trial and error or vertically transmitted to off-

Figure 2: Social-asocial learning difference  $\langle s - a \rangle$  over different payoff structures defined by  $\pi_{\text{low}}$  and  $\pi_{\text{high}}$  for low environmental availability,  $u = 0.1$ , in (a) and high environmental variability,  $u = 0.9$ , in (b).  $B = 100$ ,  $L = 20$ .

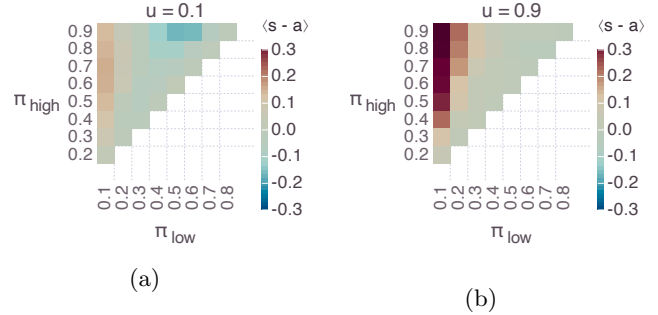
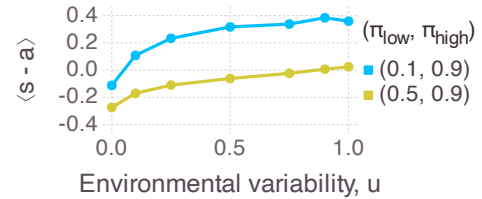


Figure 3: Social-asocial learning difference  $\langle s - a \rangle$  increases with environmental variability  $u$ .  $B = 100$ ,  $L = 20$ .



spring. When intergenerational environmental variability is high ( $u = 0.9$ ; Figure 2b) it is unlikely that the average individual knows which of the 100 possible behaviors is optimal (i.e. when  $B = 100$ ), at least in the first time steps of the generation. However, when  $A_\pi$  is low, such as when  $\pi_{\text{low}} = 0.1$  and  $\pi_{\text{high}} = 0.9$ , it is likely that at least one out of the other ten randomly selected individuals is doing the optimal behavior, and success-biased social learners can reap the benefits of this pooled information.

### Environmental variability $u$

We find that that social learning frequency broadly increases with environmental variability Figure 3 for two of the most sensitive payoff structures identified in the previous analysis. The lower  $\langle s - a \rangle$  for more ambiguous payoffs in this figure illustrates that the evolution of social learning responds differently to different types of uncertainty (Figure 3). Note that environmental changes do not particularly degrade social information in this model because  $s$  only involves horizontal transmission.

### Number of possible behaviors $B$

Next we manipulated the number of possible behaviors  $B$  (i.e. the number of bandits) agents could perform, which increases the difficulty of finding the optimal behavior. In the case of low environmental variability,  $B$  has minimal effect on the extent of social learning (see solid lines in Figure 4a). This is likely because of the relatively low need to learn from members of the same

generation when the environment doesn't vary. In contrast, with high environmental variability (dotted lines), we sometimes observe a decreased reliance on social learning as the number of behaviors increases, particularly for the high-ambiguity and low-information case of  $\pi_{\text{low}} = 0.1, \pi_{\text{high}} = 0.2$  (Figure 4a).

This appears to contradict the claim that individuals should rely on social learning for more difficult or uncertain tasks. In this computational experiment, agents only live for  $L = 20$  time steps, so that when  $B$  is smaller ( $B \approx 10\text{-}20$ ) it is more likely that one out of ten randomly selected agents is doing the optimal behavior. When  $\pi_{\text{low}} = 0.1$  and  $\pi_{\text{high}} = 0.2$ , the optimal behavior is twice as likely to pay off. When  $B$  becomes larger than 20, it is increasingly unlikely any agent knows what the best behavior is, and so social and asocial learning drift to  $s = a = 0.5$ .

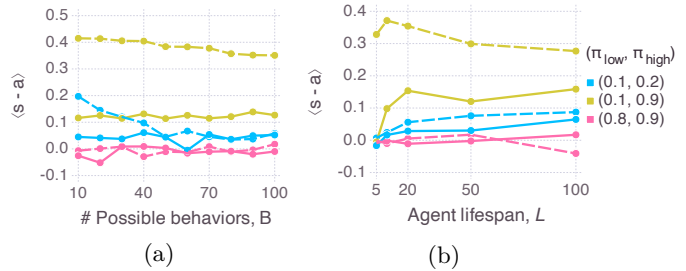
### Agent lifespan $L$

Shorter lifespans (i.e., lower  $L$ ) increase agents' uncertainty as it gives them fewer opportunities to learn the optimal behaviors. We observe a non-monotonic effect of  $L$  on  $\langle s - a \rangle$ , though most variation over  $L$  occurs in the form of a sharp increase in  $\langle s - a \rangle$  as  $L$  increases from 5 to 20 time steps for some levels of payoff ambiguity (Figure 4b). This result is particularly acute for the low-ambiguity case of  $\pi_{\text{low}} = 0.1$  and  $\pi_{\text{high}} = 0.9$  for both  $u = 0.1$  and  $u = 0.9$ . As in the previous analyses, this appears to be due to the availability of reliable teachers for horizontal social learning. When  $L$  is very small, no agents have been able to discern the optimal behavior, so social learning confers little benefit over asocial learning. However, once agents live long enough some inevitably begin doing the optimal behavior. It then becomes advantageous to pursue social learning more frequently than asocial learning. When  $u = 0.9$  it appears that increasing  $L$  slightly decreases average reliance on social learning. This is likely due to the fact that softmax will always eventually converge to the optimal behavior with enough samples. If individuals are able to eventually identify the optimal behavior on their own, there should be a slight correction to lower levels of horizontal social learning.

### Vertical transmission magnitude evolution

Vertical transmission is an important component of this model, as can be seen by the effect of intergenerational environmental variability on  $\langle s - a \rangle$ . However, in our analyses, the vertical transmission magnitude,  $v$ , evolves to be around 0.5 across our computational experiments after being initialized to 0.1. In supplemental information (Jupyter notebook included in our GitHub repository) we provide equivalent visualizations for  $v$  as we have presented here for  $s$ . We will further analyze the evolution of vertical social transmission in future work given current space limitations.

Figure 4: Social-asocial learning difference  $\langle s - a \rangle$  over (a) number of behaviors  $B$  ( $L = 20$ ) and (b) agent life span  $L$  ( $B = 100$ ) for environmental variabilities  $u = 0.1$  (solid lines) and  $u = 0.9$  (dashed lines).



## Discussion

In this paper we re-examined the relationship between uncertainty and social learning. While uncertainty is generally thought to favor social learning, we show that increased payoff ambiguity, increased number of possible behaviors, and decreased lifespan sometimes *suppress* the evolution of social learning. Our results derive from an agent-based model where individuals solve the problem of choosing among many possible behaviors to find the one with the highest payoff using social learning from peers, asocial learning by trial-and-error, and vertically-transmitted information from parents. Our model could be extended in a number of ways, including to consider questions about who learns from whom under different forms and levels of uncertainty (Muthukrishna et al., 2016).

In general this work highlights the importance of precision in characterizing sources of risk and uncertainty in understanding social phenomena. While social learning may be favored under one form of uncertainty, other forms can select against it. The time scales under which those uncertainties operate also matter. This work thus answers calls for greater precision in theory-building and hypothesis generation. More practically, this work could be adapted to guide reinforcement learning AI agents to optimize their use of social information (Jaques et al., 2019).

## Acknowledgements

This work was supported by a grant from the Stanford Medicine Center for Innovation in Global Health and by NSF BCS-2028160. Some of the computing for this project was performed on the Sherlock cluster. We thank the Stanford Research Computing Center for providing these computational resources.

## References

Allen, J. A. (2019). Community through Culture: From Insects to Whales: How Social Learning and Culture Manifest across Diverse Animal Communities. *BioEssays*, 41(11), 1–8. doi: 10.1002/bies.201900060



- Atkisson, C., O'Brien, M. J., & Mesoudi, A. (2012). Adult learners in a novel environment use prestige-biased social learning. *Evolutionary Psychology*, *10*(3), 519–537. doi: 10.1177/147470491201000309
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, *59*(1), 65–98. doi: 10.1137/141000671
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton: Princeton University Press.
- Datseris, G., Vahdati, A. R., & DuBois, T. C. (2022). Agents.jl: A performant and feature-full agent based modelling software of minimal code complexity. *Simulation*, 1–13. doi: 10.1177/00375497211068820
- Feldman, M. W., Aoki, K., & Kumm, J. (1996). Individual versus social learning: evolutionary analysis in a fluctuating environment. *Anthropological Science*, *104*(3), 209–232.
- Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1803). doi: 10.1098/rstb.2019.0502
- Henrich, J., & Boyd, R. (1998). The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behavior*, *19*(4), 215–241. doi: 10.1016/S1090-5138(98)00018-X
- Heyes, C. (2016). Blackboxing: Social learning strategies and cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693). doi: 10.1098/rstb.2015.0369
- Hill, M. R., Boorman, E. D., & Fried, I. (2016). Observational learning computations in neurons of the human anterior cingulate cortex. *Nature Communications*, *7*. doi: 10.1038/ncomms12722
- Jaques, N., Lazaridou, A., Hughes, E., Gulcehre, C., Ortega, P. A., Strouse, D., ... de Freitas, N. (2019). Social Influence as Intrinsic Motivation for Multi-Agent Deep Reinforcement Learning. In *Proceedings of the 36th international conference on machine learning* (p. 10).
- Jones, D. C., Arthur, B., Nagy, T., Mattriks, S. G., & Contributors, O. (2021). *Gadfly.jl*. *GiovineItalia*. doi: <https://doi.org/10.5281/zenodo.5559613>
- 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. doi: 10.1016/j.tics.2018.04.003
- Laland, K. N. (2004). Social Learning Strategies. *Learning and Behavior*, *32*(1), 4–14.
- Luce, R. D. (1959). *Individual choice behavior*. John Wiley.
- McElreath, R., Lubell, M., Richerson, P. J., Waring, T. M., Baum, W., Edsten, E., ... Paciotti, B. (2005). Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior*, *26*(6), 483–508. doi: 10.1016/j.evolhumbehav.2005.04.003
- Morgan, T. J., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1729), 653–662. doi: 10.1098/rspb.2011.1172
- Muthukrishna, M., Morgan, T. J., & Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior*, *37*(1), 10–20. doi: 10.1016/j.evolhumbehav.2015.05.004
- Nakahashi, W., Wakano, J. Y., & Henrich, J. (2012). Adaptive Social Learning Strategies in Temporally and Spatially Varying Environments: How Temporal vs. Spatial Variation, Number of Cultural Traits, and Costs of Learning Influence the Evolution of Conformist-Biased Transmission, Payoff-Biased Transmissio. *Human Nature*, *23*(4), 386–418. doi: 10.1007/s12110-012-9151-y
- Perreault, C., Moya, C., & Boyd, R. (2012). A Bayesian approach to the evolution of social learning. *Evolution and Human Behavior*, *33*(5), 449–459. doi: 10.1016/j.evolhumbehav.2011.12.007
- 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
- Schulz, E., & Gershman, S. J. (2019). The algorithmic architecture of exploration in the human brain. *Current Opinion in Neurobiology*, *55*, 7–14. doi: 10.1016/j.conb.2018.11.003
- Smaldino, P. E., & Richerson, P. J. (2013). Human Cumulative Cultural Evolution as a Form of Distributed Computation. In P. Michelucci (Ed.), *Handbook of human computation* (pp. 979–992). New York: Springer Science+Business Media. doi: 10.1007/978-1-4614-8806-4
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning : An Introduction* (2nd ed.). Cambridge, MA: MIT Press.
- Toelch, U., Bruce, M. J., Newson, L., Richerson, P. J., & Reader, S. M. (2013). Individual consistency and flexibility in human social information use. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1776). doi: 10.1098/rspb.2013.2864
- Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science*, *38*(4), 599–637. doi: 10.1111/cogs.12101