Uncertainty in social learning evolution

May 21, 2022

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 improve on these shortcomings. 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 that, under some conditions, these forms of uncertainty can select against social learning.

1 Introduction

1.1 The history of "copy when uncertain"

It has become a near truism that "theory predicts that it is adaptive to learn when uncertain." This claim has been motivated by empirical work (such as Asch conformism studies) & theoretical models (Boyd Richerson 85, 88) that often do not mention the word uncertainty at all. However, they do show related effects of social learning increasing when individual learning is difficult (i.e., ambiguous lines in Asch), and intermediate levels of temporal uncertainty favoring social learning (i.e., too stable and you get genetic canalization of a trait, too much temporal variation and there's selection for individual learning to ensure the most up to date information).

Galef 2009 https://www.sciencedirect.com/science/article/pii/S006534540939004X#bb0059 cites B & R 85/88 & Laland 2004 https://link.springer.com/content/pdf/10.3758/BF03196002.pdf Laland 2004 cites B& R 88 https://books.google.com/books?hl=en&lr=&id=mqFcAgAAQBAJ&oi=fnd&pg=PA29&dq=Boyd+and+Richerson%E2%80%99s+(1988&ots=IfIbB_kx_a&sig=g8CYH5bVfwfXhTRGot0q761f9m
v=onepage&q&f=false

Other examples of empirical work: https://www.sciencedirect.com/science/article/pii/S0022096516300625

Henrich & Boyd 88 do have something like B that is the difficulty of getting right answer (or rho probability of doing so) https://www.sciencedirect.com/science/article/pii/S109051389800018X#FIG5

Lifespan not modelled

1.2 earlier structure

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 (?, ?). Theory predicts that social learning should be favored in contexts with greater uncertainty (?, ?, ?), and this prediction has received some empirical support (?, ?, ?). 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 (?, ?).

In this paper we use agent-based modeling to compare the effect of different sources of uncertainty on social learning by "unblackboxing" typically abstracted-out model components of environmental variability, payoff structures and agent life histories, and learning mechanisms. "Uncertainty" 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. We thus conclude that many predictions made by previous models of the evolution of social learning are likely overgeneralized.

1.3 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 nonhuman taxa (?, ?, ?). 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* (?, ?). 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 (?, ?, ?). 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 (?, ?). 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 (?, ?).

Uncertainty has also been modelled as arising from other aspects of the environment. For example, ? (?) 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 asocial 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.

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 asocial information (?, ?, ?), 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 (?, ?). ? (?) 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 (?, ?).

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.

1.4 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 (?, ?), 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.

2 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", a common modelling and experimental approach for representing behaviors with probabilistic payoffs (?, ?, ?, ?, ?). Each behavior b yields Bernoulli payoffs: payoff of 1 with probability π_b and zero payoff with probability $1 - \pi_b$; π_b is therefore also the expected payoff of behavior b. Agents must decide which behavior to perform at each time step. To do this, agents use an explore-exploit strategy to sometimes try the most profitable behavior they know about, and other times try alternatives that may pay off more reliably.

We operationalized uncertainty in four different ways: (1) payoff ambiguity, A_π, which measures the difference between the optimal expected payoff behavior π_{high} and the expected payoff of the other behaviors, π_{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 until agents evolved to be all social or all asocial learners, or until 20k time steps was reached. Only 44 out of 396k trials did not reach fixation (all social or all asocial). At the end of each generation agents reproduce and then die off. Those selected to reproduce pass on their boolean social learning trait, s, without mutation, which specifes whether agents learn socially. We developed a series of computational analyses where we systematically vary the uncertainty variables and observe the average prevalence of social learning trait s across agents and trials, which is our primary outcome measure, denoted $\langle s \rangle$.

2.1 Agents and their attributes

In each time step, N agents—autonomous problem solvers—select which behavior to perform based on their running tally of mean payoffs for each available behavior. Agent i tracks the mean payoff of each behavior b, denoted $\bar{\pi}_{ib}$, and a *count* of how many times it has performed each behavior, denoted c_{ib}. $\bar{\pi}_{ib}$ is initialized to 0 for all b at model initialization for all agents, and after each generation for individual learner agents. Social learning agents' $\bar{\pi}_{ib}$ is initialized as that of its teacher selected through performance-biased oblique learning. Agent i's accumulated payoffs from performing several behaviors over their lifetime of L time steps is denoted π_i.

Table 1: Dynamic agent-level variables. Each has an implicit time dependence.

Attribute	Description	Initial value
s_i	Social learner trait: 1 if agent i is social learner; 0 otherwise	0 or 1 w/ equal prob.
$\bar{\pi}_{ib}$	"Ledger": mean payoffs acquired via behavior b by i	B-vector of floating-point zeros
c_{ib}	Count of how many times agent i performed b	B-vector of integer zeros
π_i	Net payoffs accumulated by i within generation	0.0

2.2 Modeling uncertainty

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

1. 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 π_{high} while all other behaviors are assigned a payoff of π_{low}. Otherwise, the same behavior remains optimal across generations.

2. 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 that yields an expected payoff of π_{high}, while all other behaviors yield a payoff of $\pi_{\text{low}} < \pi_{\text{high}}$. As π_{low} increases towards π_{high}, uncertainty in the form of *payoff ambiguity* also increases, making it more difficult, and less rewarding, for agents to determine which behavior is the optimal one.

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.

2.3 Dynamics and evolution

Model dynamics can be split into three parts: initialization, *intra*generational behavior and payoff accumulation, and *inter*generational transmission of the social learning trait and oblique learning (Figure 1). The model runs that we analyze here were initialized with $N = 100$ agents to randomly have the social learner trait or not, with no accumulated payoffs, and blank "ledgers". At each time step, agents select one of B possible behaviors to perform, accumulating a payoff of 1 if the bandit pays off, and 0 if it does not—a bandit pays off with probability π_{high} if it is the unique optimal behavior, or with probability π_{low} if it is one of $B-1$ non-optimal behaviors. Agents accumulate payoffs in this way for L time steps in each generation. At the end of each generation, $N = 100$ agents are randomly selected with replacement to reproduce (asexual, haploid reproduction), weighted by their net payoffs over the generation, i.e., *performance-biased* reproduction. Child agents then learn from one teacher, selected again via a form of performance bias: a child agent chooses its teacher by first selecting a fully random subset of agents from the parent generation as potential teachers, then selecting the one with the greatest net payoffs among the subset. This process continues until agents have become all social or all asocial learners, or when $20k$ time steps have elapsed.

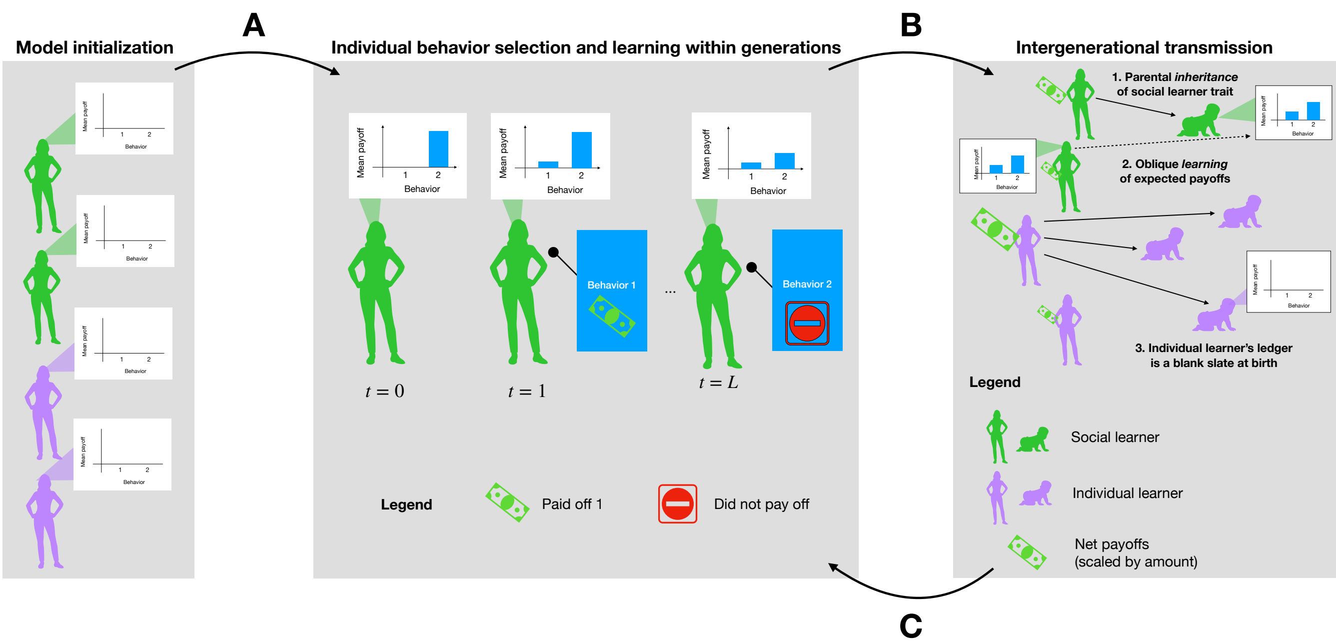

Figure 1: **Model dynamics.** The first generation of agents are initialized with empty ledgers, then perform behaviors on each subsequent time step (A). At the end of each generation, after all agents have performed L behaviors chosen using softmax selection, the intergenerational tramsission stage begins (B). During this stage agents reproduce, passing on their social learning trait without mutation (B1). Child agents learn from the previous generation if they are social learners (B2), or are initialized with blank ledgers if they are individual learners. All agents from the previous generation die off, and new child agents begin within-generation dynamics anew.

2.3.1 Initialization

Model initialization includes agent and environment initializations. Agents are randomly initialized to have the social learning trait or not, with equal probability. Agent "ledgers" and behavior counts are uniformly initialized to the no-knowledge, blank-slate case where all $\pi_{ib} = 0$ and $c_{ib} = 0$. Both the ledger and behavior count vectors have B entries, one for each behavior the environment affords. Net payoffs for the generation are also initialized to zero. All agents are initialized with the same softmax temperature, τ, that guides their behavior selection to be either more exploratory (greater τ) or more exploitative of what agents believe to be the optimal behavior(s) (lesser τ).

The environment is initialized to have B behaviors with one of them selected at random to be the optimal behavior, denoted b^*. Behavior b^* is initialized to have a probability π_{high} of paying off 1. All other behaviors are initialized to have probability π_{low} of paying off 1. The modeler must specify the environmental variability, u, and the number of time steps per generation, L.

Table 2: Global environmental and cognitive initialization parameters. (MT: TODO: Test more N_{T} for sensitivity analysis)

Symbol	Description	Values tested (bold=default)
B	Number of possible behaviors (represented by "bandits")	2, 4, 10
π_{high}	Probability that the unique optimal behavior pays off 1	**0.9**
π_{low}	Probability one of $B-1$ non-optimal behaviors pays off 1	0.1, 0.45, 0.8
τ	Softmax temp.; $\uparrow$=more exploration, $\downarrow$=more exploitation	0.01, **0.1**, 1.0
u	Probability optimal behavior changes between generations	0.0, 0.1, ..., 1.0
L	Number of time steps per generation	1, $B/2$, B, $2B$, $4B$
N_T	Number of teachers to pool, from which best selected	**5**

2.3.2 Intra-generational behaviors and payoffs

Each generation begins with agents initialized either by the $t = 0$ initialization outlined above, or initialized via inter-generational reproduction and learning, described in detail below. Within generations, there is no social learning. At each time step within a generation, agents select a behavior to perform using the softmax algorithm, then update their ledgers $\bar{\pi}_{ib}$ and behavior counts c_{ib}. If the chosen bandit pays off for an agent, its net payoff is incremented by one, $\pi'_i \leftarrow \pi_i + 1$. This process continues for L time steps, at which point reproduction, learning, and die-off occur, which re-initializes the next generation for performing L behaviors selected via this same process.

At each time step, agent i chooses behavior b at random, with each behavior weighted by the softmax function applied to that behavior's observed mean payoff relative to all mean payoffs in the ledger,

$$\Pr(i \text{ chooses behavior } b) = \frac{\exp(\bar{\pi}_{ib}/\tau)}{\sum_{b=1}^{B} \exp(\bar{\pi}_{ib}/\tau)}. \tag{1}$$

Softmax behavior selection is a biologically plausible model of behavior (?, ?) that enables agents to explore alternative behaviors sometimes and exploit the best observed behavior other times, in accordance with Luce's choice axiom (?, ?),. The parameter τ specifies how frequently alternative behaviors are explored versus how frequently the best observed behaviors are exploited. Greater τ means more frequent exploration, lesser τ means more frequent exploitation. We set $\tau = 0.1$ for all computational analyses presented in the main paper. We performed sensitivity analyses and found a weak dependence on τ that does not affect our main conclusions (MT: have run τ sensitivity checks, need to analyze them).

When agent i performs behavior b, i's behavior count is incremented by 1, $c'_{ib} \leftarrow c_{ib} + 1$. Agent i's ledger of mean payoffs are updated using exponential weighted averaging,

$$\bar{\pi}'_{ib} = \bar{\pi}_{ib} + \frac{\mathrm{Bandit_b}(0,1) - \bar{\pi}_{ib}}{c'_{ib}}, \tag{2}$$

where $\mathrm{Bandit}_b(0,1)$ is 0 or 1 depending on the result of the bandit draw for behavior b.

2.3.3 Inter-generational inheritance and social learning

Every L time steps, agents from the current generation reproduce, teach, and die, while the next generation to perform the next L time steps learns from one teacher from the previous generation if they inherited the social learner trait. Inter-generational dynamics thus depend on two payoff-biased selection mechanisms: reproducer selection and teacher selection.

N reproducers are sampled from the population with replacement weighted by net payoffs, i.e., at each of N draws for a parent,

$$\Pr(\text{Agent } i \text{ chosen for reproduction}) = \frac{\pi_i}{\sum_{i=1}^{N} \pi_i}. \tag{3}$$

We sample with replacement to allow for more successful agents to have multiple offspring. Child agents inherit their parents' social learning trait, so all social learner parents spawn social learner offspring, and parents lacking the trait spawn offspring lacking the trait.

Child agents with the social learning trait then must select and learn from a teacher from their parent's generation, including possibly their parent. Child agents give no preference to whether potential teachers are social learners. A child selects a teacher by first selecting a pool of $N_T = 5$ potential teachers. Among this pool, each child selects the teacher with the greatest net payoffs. In case of a tie the teacher is chosen at random from those with the shared maximum net payoff.

Social learner child agents each acquire their teacher's ledger, and all social learner children behavior counts are reset to 1 to limit ledger values to be between 0 and 1. Child agents without the social learner trait have all ledger and behavior count values re-initialized to 0. At this point the re-initialized child agents engage in behavior selection and payoff accumulation described above.

2.4 Computational analyses and outcome measures

We manipulated environmental uncertainty parameters described above, u, π_{low}, B, and L, to examine their effects on our main outcome measure, the mean prevalence of social learning $\langle s \rangle$, and supplemental outcome measures $\langle \pi \rangle / L$ and $\langle T \rangle / L$. $\langle \pi \rangle / L$ is the mean net payoffs accumulated by agents at the end of each generation for the final generation of agents, normalized by lifespan L. $\langle T \rangle / L$ is the mean time steps to convergence normalized by the lifespan of agents, i.e., the mean number of generations until convergence was reached. Mean values were calucu-lated across 100 model runs for each combination of uncertainty parameter values. To test the effect of the uncertainty parameters, we initialized several model runs with systematically varied uncertainty parameters, u, π_{low}, B, and L. We varied $u \in \{0.0, 0.1, \ldots, 1.0\}$ for each combination of the following parameters: $\pi_{\text{low}} \in \{0.1, 0.5, 0.8\}$, $B \in \{2, 4, 10\}$, and $L \in \{1, B, 2B, 4B\}$ for $B = 2$ and $L \in \{1, B/2, B, 2B\}$ for $B = 4, 10$.

Table 3: Outcome variables

Symbol	Description	Values
$\langle s \rangle$	Mean social learning prevalence over agents and trials	$\in [0.0, 1.0]$
$\langle \pi \rangle / L$	Mean payoffs accumulated in a generation normalized by lifespan	$\in [0.0, 1.0]$
$\langle T \rangle / L$	Mean number of generations to convergence	Bounded above by 20k / 8

2.5 A note on relationship with standard RL formalism and models

We presented our model as an agent-based model because that is how we originally developed it. Howevever, we believe the model could be translated in a straightforward way where each action at time t, typically denoted a_t, is to "pull" one of the B bandits, and the policy π would involve using the softmax function to weight random behavior selection (?, ?). Because of the complication of two relevant time intervals, intervals between actions and generational intervals, and space limitations we do not try to put the entire model in standard RL terms. (MT: Expand/improve.)

3 Analysis

To understand the effect of uncertainty parameters (1) environmental variability, u; (2) payoff ambiguity that increases with π_{low}; (3) number of behavioral options, B; and (4) generation length/agent lifespan, L, on the evolution of social learning, we analyze the results of model runs across systematically varied parameter combinations. We observed complex outcomes where the effects of uncertainty parameters were highly dependent on one another as measured by the mean prevalence of social learning, $\langle s \rangle$, across trials for each parameter setting. Increased environmental variability, u, tended to suppress social learning as expected from standard models of social learning evolution. However, the shape of the decrease in $\langle s \rangle$ was not uniform (Figure **??**). Depending on the parameter settings, the critical value of u at which social learning suppression began varied, as did the steepness of the decrease. We found that the location and width of the decrease in $\langle s \rangle$ is correlated with the difference between the expected net payoffs of a population of all social learners, $\langle \pi_{\text{S}} \rangle$, and the expected net payoffs of an all-asocial learner population, $\langle \pi_{\text{A}} \rangle$ (Figure **??**). evolution does not always choose the theoretically highest payoff strategy, and confirmed that evolution is more uncertain (i.e. selection is weaker (?, ?, p. 103)) when $\langle \pi_{\mathrm{S}} \rangle - \langle \pi_{\mathrm{A}} \rangle$ is small, where evolutionary uncertainty is measured by the average steps to model convergence, $\langle T \rangle$ (Figure **??**).

3.1 Evolution of social learning under uncertainty

We observed that different uncertainty contexts lead to different patterns in the evolution and suppression of social learning. To understand the evolution and suppression of social learning we inspected the social learning prevalence, $\langle s \rangle$, which is the average of the frequency of social learning across a model population over all trial model runs for a 4-tuple of uncertainty values, $(u, \pi_{\mathrm{low}}, B, L)$. We plot $\langle s \rangle$ on the y-axis over $u = 0.0, 0.1, \ldots, 1.0$ on the x-axis in each plot, with nine plots total over $\pi_{\mathrm{low}} = 0.1, 0.45, 0.8$ and $B = 2, 4, 10$, broken out by $L = 1, B/2, B, 2B$ within each plot (Figure **??**). Across uncertainty scenarios, $\langle s \rangle$ monotonically decreases u increases, in accordance with classic cultural evolutionary predictions (?, ?, ?, ?). Therefore we compare differences between uncertainty contexts in terms of the *suppression of social learning* over (increasing) u.

Differences emerged in the suppression of social learning on two dimensions. First there were systematic differences in the value of u when social learning first begins to be suppressed, i.e., where $\langle s \rangle$ first decreases. We show in the following subsection that this occurs when $\langle \pi_{\mathrm{S}} \rangle$ approaches $\langle \pi_{\mathrm{A}} \rangle$. The second dimension of analysis is the width of the transition from maximal to minimal $\langle s \rangle$. In some cases, this width was less than 1, but in many other cases the width was equal to 1, meaning that $\langle s \rangle$ decreased over all u. This indicated weak selection pressure since less than 0.001% of all trials ended with all agents either social or asocial learners. This is confirmed below by examining the average number of time steps it took to reach fixation (agents all social or all asocial).

For now we will examine differences in the location and width of social learning suppression. (MT: In outline form while I work out each point)

- First, note that as π_{low} increases towards $\pi_{\mathrm{high}} = 0.9$, social learning suppression widens, becoming linearly decreasing over u with no $\langle s \rangle$ reaching 1 or 0. This is due to weakened selection as the difference between social and asocial learning payoffs goes to zero.

- Second, as B increases, note that social learning suppression begins at increasingly high u. This is because it is increasingly worth the risk that a socially-learned ledger biases agents towards non-optimal behaviors when the number of choices increases. Furthermore, when B

is larger, one large ledger entry is not as misleading as when B is smaller since there is more probability mass in the softmax vector's non-maximal behavior entries when B is larger, and so alternatives to the socially learned "best" behavior will be explored more frequently.

- Finally, note that decreased lifespan both increases the location of social learning suppression and the width of social learning suppression, with the extreme case of $L = 1$ becoming linear and eventually flat compared to other values of L as π_{low} and B increase.

To support these explanations we need to analyze the average payoffs obtained by agents in these various uncertainty contexts and compare them with the expected payoffs when all agents are individual learners, and with the expected payoffs when all agents are social learners.

3.2 Relative benefit of social learning

Social learning begins to be suppressed when the expected payoffs from a population of all social learners, $\langle \pi_{\text{S}} \rangle$ (Figure **??**; small dashed line with diamond markers), decreased to the point where they are nearly the same as expected payoffs for a population of all individual learners, $\langle \pi_{\text{A}} \rangle$ (Figure **??**; horizontal long-dash lines). We see that social learning becomes completely suppressed when $\langle \pi_{\text{S}} \rangle < \langle \pi_{\text{A}} \rangle$. Mean payoffs realized by model agents approximately track expected payoffs of whichever strategy is optimal, social or asocial learning, but with some important variation.

Similarly, sometimes populations that start out mixed outperform all social learner populations when $\langle \pi_{\text{S}} \rangle > \langle \pi_{\text{A}} \rangle$, indicating that an initial diversity of learner types helps with the long-term prosperity of the group, even though the asocial learner type tends to die out. (MT: I suspect this may be sensitive to N_T, so my next analysis will be the effect of N_T.)

Second, notice that asocial learning may be the safe fallback option when evolution is unsure which strategy is optimal, but a population of all social learners would do better than what evolution tended to select for (Figure **??**, e.g., when $\pi_{\text{low}} = 0.45$, $B = 2$, and $L = 8$ in the middle left plot). (MT: I suspect this will also be sensitive to N_T.)

3.3 Selection strength and evolutionary uncertainty

If selection is truly weaker when $\langle \pi_{\text{S}} \rangle \approx \langle \pi_{\text{A}} \rangle$, then we should not only see a more perfect bimodality of outcomes, with roughly half of outcomes going to both $\langle s \rangle = 1$ and $\langle s \rangle = 0$, but we should also