Article
Evolution of Social Learning with Payoff and Content Bias

Charles Perreault * and Robert Boyd

Institute of Human Origins, School of Human Evolution and Social Change, Arizona State University, Tempe, AZ 874102, USA; robert.t.boyd@gmail.com
* Correspondence: perreault.c@gmail.com

Abstract: There has been much theoretical work aimed at understanding the evolution of social learning; and in most of it, individual and social learning are treated as distinct processes. A number of authors have argued that this approach is faulty because the same psychological mechanisms underpin social and individual learning. In previous work, we analyzed a simple model in which both individual and social learning are the result of a single learning process. Here, we extend this approach by showing how payoff and content biases evolve. We show that payoff bias leads to higher average fitness when environments are noisy and change rapidly. Content bias always evolves when the expected fitness benefits of alternative traits differ.

Keywords: cultural evolution; individual learning; payoff bias; content bias

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

There has been a substantial amount of theoretical work focused on the evolution of social learning [1–14]. Investigators asked what conditions favor individuals who imitate others, rather than learn on their own, and how selection shapes the process of imitation. In most of this work, individual and social learning are treated as distinct processes. Individual learning occurs when individuals use environmental cues to adjust their behavior to local conditions. Social learning is a separate transmission process in which the determinants of behavior are transmitted socially from one individual to another. This transmission process may be subject to errors, biases and systematic transformations, but most work assumes that social learning leads to reasonably accurate copying. Then, to build models of cultural evolution, investigators modify mathematical models drawn from population biology to account for the novel structure of social learning. Finally, to model the long-run evolution of the cultural capacities, researchers assume that the parameters that govern the cultural transmission process are genetically heritable, and ask how natural selection shapes the relative importance of social learning. This work has been widely influential, transforming the idea of cultural evolution from a vague analogy to an vibrant area of both theoretical and empirical research [15].

This approach has it critics. Some have complained that social and individual learning are not psychologically distinct processes [16]. Indeed, both individual and social learning involve cue-based inferences about what is the best behavior in the organism's environment. Others point out that this work assumes that social and individual learning are alternatives competing for determination of phenotype when in fact they are usually complementary processes that lead individuals in the same direction [17].

To assess the importance of these critiques, Perreault et al. [18] analyzed a model in which individual and social learning result from a single learning process. They assumed that individuals attempted to infer current state of a variable environment using two sources of information, the behavior of the individuals from the previous generation and a non-social cue that provides information about the current state of the environment, and that natural selection shaped the learning process over the long run so that it maximized the expected fitness. This model has qualitatively similar behavior to previous work that

assumed social and individual learning were distinct processes and so suggests that this assumption is not crucial. However, the model only treats one kind of directional process, termed guided variation [2], and does not deal with biased cultural transmission which occurs when social learning is biased in favor of some cultural models or some trait values.

Here, we demonstrate that this Bayesian framework is flexible and powerful enough to incorporate other cultural transmission processes in the same single learning process. More specifically, we extend the model to two forms of biased transmission. First, we suppose that learners can observe both behavior and the payoffs of a number of individuals and ask how selection should modify the learning psychology to make use of the payoff information. We derive the optimal payoff bias rule, and study the conditions under which it is adaptive. Second, in the original model competing traits had the same expected fitness. Now, we suppose that one trait has a fitness advantage over the long term and show how this gives rise to content bias in favor of that trait. In both cases, qualitative results are similar to those derived assuming individual and social learning are distinct processes. The current analysis also provides more specific predictions about the form of payoff and content biases.

2. Methods

To model the evolution of the learning process, we assume individuals belong to a large population that lives in a environment that switches between two states with a specified probability. There are two behaviors, labeled *state 1* and *state 2*. One behavior has higher fitness when the environment is in state 1 and the alternative behavior has higher fitness when the environment is in state 2. The adaptive problem is to infer the current state of the variable environment using two kinds of information: the behavior and, in the payoff bias model, the fitness of a random sample of n individuals from the previous generation (social cues) and a non-social cue that provides information about the current state of the environment. We derive an analytical expression for the optimal learning rule by modeling learning as Bayesian inference, a framework that has been widely used to study learning and cognitive development [19].

Each individual's decision depends on her cues. This means that given the frequency of behaviors and payoffs in the previous generation we can calculate, the distribution of behaviors and payoffs in the present generation. However, the optimal learning rule depends on parameters that in turn depend on the long term average state of the population and accuracy of the non-social cue, facts that individuals do not know. Thus to determine the optimal reliance on social cues, it is necessary to model the co-evolution of the culturally transmitted pool of information, and the genes that determine how this information is transmitted. We accomplish this using numerical simulation.

3. Results

3.1. Payoff Bias

In this section, we extend the model analyzed in [18] to allow for payoff bias, where payoffs are assumed to be fitnesses. Consider a large population that lives in a environment which can exist in two states, labeled *state 1* and *state 2*. Each generation, the environment switches with probability γ and stays the same with probability $1 - \gamma$. This means that over the long run the environment is equally likely to be in each state. Individuals acquire one of two behaviors, *behavior 1* and *behavior 2* by either individual or social learning. Behavior 1 has fitness $W + d$ when the environment is in state 1 and fitness W when the environment is in state 2. Behavior 2 has fitness $W + d$ when the environment is in state 2 and fitness W when the environment is in state 1.

Each individual observes three cues that are predictive of the state of the environment. The adaptive problem is to determine the best way to use these cues.

> **Environmental cue.** Individuals observe an environmental cue, y, that can take on any real value. Let $\Pr(y|1)$ and $\Pr(y|2)$ be the probability that an individual observes cue value y in environments 1 and 2 respectively. The environmental cue is a normally

distributed random variable with mean ξ and variance ω when the environment is in state 1 and has mean $-\xi$ and ω when the environment is in state 2. This means that positive values of y indicate that it is more likely that the environment is in state 1 and negative values that it is in state 2. As the variance increases, a given cue value is a poorer predictor of environmental state.

Behavior. Each individual also observes the n models randomly sampled from the previous generation. For the jth model, social learners observe their behavior b_j and their payoff, x_j. Behaviors take on values 1 or 2, and payoffs are real numbers. The vector of behaviors is **b** and the vector of payoff values is **x**. Let p be the expected frequency of behavior 1 given that the population is experiencing environment 1. Due to the symmetry of the model, p is also the expected frequency of behavior 2 given that the population is experiencing environment 2. Then $\Pr(\mathbf{b}|1) = p^j(1-p)^{n-j}$ and $\Pr(\mathbf{b}|2) = p^{n-j}(1-p)^j$ where j is the number of individuals with the favored behavior.

Payoff. The payoff (aka fitness) of an individual with the favored behavior is a normally distributed random variable with mean $\mu + d$ and variance v and the payoff an individual with the disfavored behavior is a normally distributed variable with mean μ and variance v. The common mean, μ, is itself a normally distributed random variable with mean zero and a very large variance, V. This means that the absolute magnitudes of payoffs provide no information about the state of the environment, but that the difference between the payoffs is informative.

The information available to a given learner is the value of the non-social cue, y, and pair of vectors $\mathbf{b} = (b_1, b_2, \ldots b_n)$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Let $\Pr(b_j, x_j|k)$ be the joint probability that the jth individual has the favored behavior $b_j = k$ and payoff x_j in environment k. The probabilities of payoffs of different models conditioned on behavior and the state of the environment are independent so $\Pr(\mathbf{x}, \mathbf{b}|k) = \prod \Pr(x_j, b_j|k)$. The learner uses Bayesian methods to infer $\Pr(1|y, \mathbf{x}, \mathbf{b})$, the probability that the environment is in state 1 given that a learner observes an environmental cue y and models with behaviors **b** and payoffs **x**. Then the optimal decision rule is adopt behavior 1 if $\Pr(1|y, \mathbf{x}, \mathbf{b}) > 1/2$ otherwise adopt behavior 2.

It is shown in Appendix A.1 that this is equivalent to the following inequality.

$$j - \frac{n}{2} > G\left(\frac{j(n-j)}{n}\right)(\bar{x}_2 - \bar{x}_1) - gy$$

where

$$G = \frac{d}{v \ln\left(\frac{p}{1-p}\right)}$$

and

$$g = \frac{\xi}{\omega \ln\left(\frac{p}{1-p}\right)}$$

If the learner knew the density of y and frequency of the favored trait in the current environment, she could compute the values of these parameters. However, she does not know either of these things, and so cannot compute g and G. Instead we suppose that these are aspects of individual psychology. The value of g gives propensity to rely on non-social cues, and G the weight of payoff information both relative to the weight placed on the behavior of others. When $G = 0$, the learner ignores payoff information and the learning rule reverts to that given in [18]. When $G > 0$, the learner is more likely to adopt the trait the exhibits the higher mean payoff among its sample of cultural models.

This rule indicates that importance of payoff differences between traits should be scaled by the term $j(n-j)$. This means that payoff bias should be more important when cultural models exhibit a mixture of behaviors than when most cultural models behave in the same way. As a result, rare beneficial innovations will be less likely to be adopted than beneficial innovations that have become established. Notice that this is not the same as the

usual mass action effect which results from the probability that social learners observe an innovation. Instead, social learners are less likely to adopt even when they have observed the payoff advantage of the novel behavior. As far we know, no previous formulation of payoff bias incorporates this phenomenon.

Finally, this rule encourages a different view of payoff bias than given in other work (e.g., [2]) where payoff bias is often conceptualized as a mechanism that determines who are the most attractive models. Here, individual payoffs are just data about the effects of alternative behaviors, and the rule simply weights the observed mean payoff of each behavior.

We assume that g and G are heritable attributes of the organism's psychology that are shaped by natural selection. To model their evolution, we assume that the values of g and G are affected by a large number of alleles at a two haploid loci. Individuals first acquire their genotype through genetic transmission. Then, they observe members of the previous generation and an environmental cue, and determine whether they should adopt behavior 1 or behavior 2. Finally, viability selection adjusts the genotypic frequencies.

We used an agent-based simulation to investigate how natural selection shapes individual's learning psychology, G and g, under different sets of environmental conditions. We are particularly interested in the benefits of payoff information when individual learning is hard, that is, when the environmental cues are noisy. This assumption captures the edge case that we think makes culture so adaptive: when the solution to an ecological challenge is difficult to discover through individual learning alone, but once discovered, has benefits that are easy to observe.

We also want to test the hypothesis that payoff biased transmission is helpful when the environment changes rapidly. Rapid environmental change makes social learning less adaptive because the social information acquired from previous generations become outdated. We suspect that payoff-biased transmission may help shield social learners from the deleterious impact of environmental change by providing them with another source of information that is less frequency dependent. This is important because it would mean that payoff-biased transmission can increase the range of environmental conditions under which social learning can evolve.

The agent-based simulation keeps track of the evolution of alleles that affect the values of G and g in a population of organisms. We assume haploid, asexual genetic transmission. An individual with the ith G allele and the kth g allele has a learning rule characterized by the parameters G_i and g_k that can take any real value. Individuals observe the behavior of the members of the previous generation, an environmental cue and a payoff cue and determine whether they should adopt behavior 1 or 2 using the optimal decision rule presented above. Then, viability selection adjusts the genotypic frequencies. Every time steps corresponds to a generation. Each generation, the following happens:

1. The state of the environment switches from state 1 and 2 with probability γ.
2. The individuals get a social cue, an environmental cue, and a payoff cue:
   - The social cue is the number j of individuals with behavior 1 among n social models drawn randomly from the previous generation. As there are two behavioral variants, the social cue is binomial with parameters p and n, where p is the frequency of behavior 1 in the population at previous generation.
   - The environmental cue y is drawn from a normal distribution with mean ξ in environment 1 and $-\xi$ in environment 2, and standard deviation e_{sd}.
   - The payoff cue, $(\bar{x}_2 - \bar{x}_1)$, is drawn from a normal distributed with mean d in environment 1 and $-d$ in environment 2, and standard deviation $\sqrt{v}$.
3. Individuals combine those three cues using the optimal decision rule analytically derived above to chose a behavior.
4. Viability selection occurs. The baseline fitness is W. Individuals with the favored behavior, given the current state of the environment, get a fitness benefit d. Reproduction is based on relative fitness, which is calculated by dividing individual fitness by the maximum fitness in the population. The fitness of individuals relative to the maximum fitness in the population is used as a vector of weights in the *sample* function in

the R language *base* package in order to sample the individuals that reproduce and transmit their alleles to the next generation.

5. Mutations in *G* and *g* alleles in the next generation occur with probability *M*. The values of the mutant alleles are drawn from a normal distribution with mean equal the allele of the parent and standard deviation m_{sd}. *G* and *g* are unlinked and mutate independently.

Several parameters of the simulation were kept constant throughout all the runs examined in this paper: number of agents = 10,000, $e_{sd} = 1$, $W = 0.5$, $d = 1$, $M = 0.05$, and $m_{sd} = 0.5$.

At the start of every run, the state of the environment, the behavior of each agent, as well as the values of the *G* and *g* alleles, are all set to 1. Each simulation ran for at least 5000 environmental shifts. After 5000 environmental shifts, the simulation continued until the distribution of *G* and *g* alleles became stationary. Stationary conditions were met when, for both *G* and *g*, the slope of a linear regression models fitted to the median allele value in the population, over the last 2000 generations, was smaller than 0.001.

In order to measure the fitness benefits associated with payoff information, we also ran simulations with the decision rule that includes only social and environmental cue [18]:

$$j - \frac{n}{2} > -gy$$

The simulation results suggest that payoff bias is most adaptive when learning from the environment is hard and when the environment is unstable. Figure 1 shows the average fitness in the population for high- and low-quality environmental information as a function of the quality of the payoff information. The fitness values plotted are relative to the average fitness in population that evolved in the same conditions but without payoff information. Thus, relative fitness greater than one mean payoff bias increases fitness, above and beyond the fitness provided by social and environmental cues alone. In the "high" regime of quality of environmental information, $\xi = 0.1$ (circles). In this regime, the chance of adopting the favored behavior via individual learning alone is 0.54. In the "low" regime, $\xi = 0.1$ (squares), i.e., the probability of adopting the favored behavior is 0.504 — a hair better than flipping a coin to make a decision.
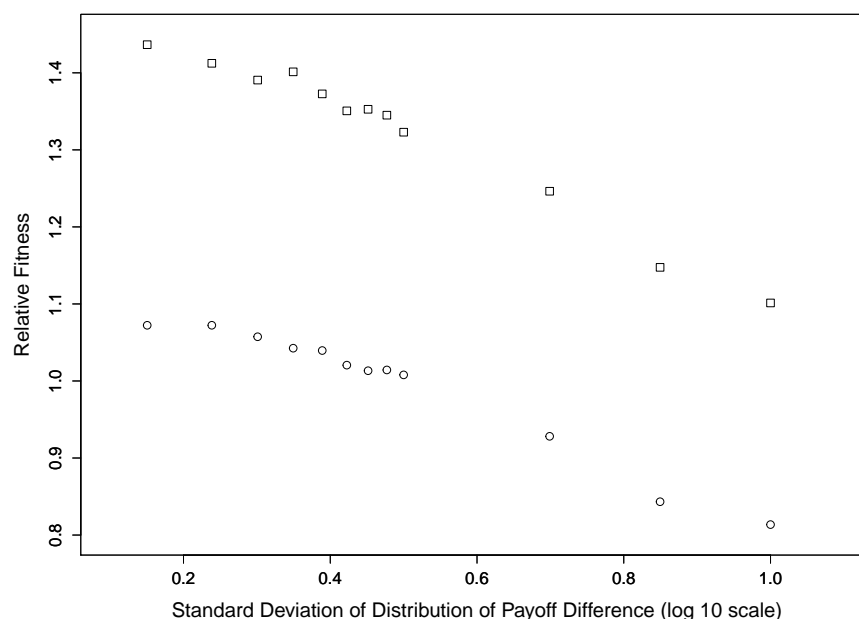

Figure 1. The effect of quality of the environmental cue on relative fitness gain associated with payoff bias. The squares represent cases where the quality of the environmental cue is low ($\xi = 0.01$) and the circles the cases where it is high ($\xi = 0.1$). The x-axis represents the quality of the payoff information, i.e., $\sqrt{v}$ on a log10 scale. Other parameters are as follow: $n = 3$, $\gamma = 0.01$.

Within each regime of environmental information quality, the relative fitness benefits of payoff biased transmission decrease with the standard deviation of the payoff cue distribution, $\sqrt{v}$ (*x*-axis). These results plotted in Figure 1 suggest that, unless it is very noisy, payoff information generally leads to higher fitness. More interestingly, the relative fitness benefits are highest in the low-quality of environmental information regime. As the quality of the environmental cue decreases from $\xi = 0.1$ to $\xi = 0.01$, relative fitness increase by approximately 0.33. Notice that the evolutionary stable cultural transmission rule may reduce average fitness because selection does not maximize average fitness. Selection favors more social learning that is optimal for the population.

We also found that payoff bias leads to higher relative fitness in fast changing environments. Figure 2 shows relative fitness in fast changing environments ($\gamma = 0.1$, circles) and a slow changing environments ($\gamma = 0.01$, squares). Given the same quality of payoff cue (*x*-axis), organisms in the fast changing environment condition enjoyed relative fitness benefits that were, on average, greater by 0.09 units. Payoff biases are most beneficial in rapidly changing environments because unstable environments lead to higher behavioral variation in the population. For instance, a population in a rapidly changing environment will spend more time in intermediate values of *p* than a population in stable environment, in which the frequency of the favored behavior can remain high for several generations in a row. Since the value of payoffs depends in part on the variance in behavior among the social models observed, payoffs are more useful in rapidly changing environments.
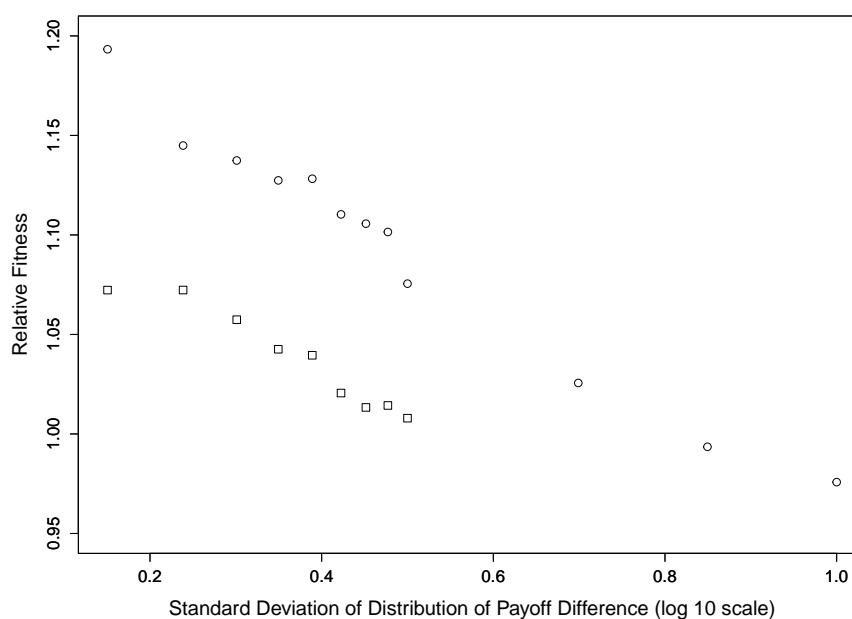

Figure 2. The effect of rates of change in the environment on relative fitness gain associated with payoff biases. The squares represent cases where rate of change is low ($\gamma = 0.01$, squares) and the circles the cases where it is high ($\gamma = 0.1$), circles The x-axis represents the quality of the payoff information, i.e., $\sqrt{v}$ on a log10 scale. Other parameters are as follow: $n = 3$, $\xi = 0.1$.

Overall, our results suggests that selection will favor payoff-biased social transmission under a wide range of conditions. In particular, ecological problems that are hard to solve individually, as well as fast changing environments, will strongly favor using payoff information.

3.2. Content Bias

In both [18] and the payoff bias model analyzed above, we assumed that the variable environment was symmetric. Each environment was equally likely, and the two behaviors had the same advantage in the environment in which they were favored. Here, we show when these assumptions are relaxed, selection favors a decision rule which is biased in favor of the behavior favored in the more common environment and the behavior which has a larger fitness advantage.

Consider a large population that lives in a environment which can exist in two states, imaginatively labeled *state 1* and *state 2*. Each generation, the environment switches from state 1 to state 2 with probability γ_2 and switches from state 2 to state 1 with probability γ_1. The means that over the long run the environment will be in state 1 with probability

$$\pi = \frac{\gamma_1}{\gamma_1 + \gamma_2}$$

Individuals acquire one of two behaviors, *behavior 1* and *behavior 2* by either individual or social learning. Behavior 1 has fitness $1 + d_1$ in environment 1 and 1 in environment 2. Behavior 2 has fitness $1 + d_2$ in environment 2 and 1 in environment 1. Each individual observes an environmental cue, x, that can takes on a range of values. Let $\Pr(y|k)$ be the probability that an individual observes cue value y in environment k. We assume that this probability is normal with mean μ and variance v in environment 1 and mean $-\mu$ and variance v in environment 2. Each individual also observes n models sampled at random from the previous generation. We ignore differences in prestige, age, etc., and assume that all models are identical, so the only thing that matters is the number of models exhibits trait one or two. Let j be the number of models who exhibit the favored behavior the current environment. Let the probability of j conditioned on the environment being in state k is $\Pr(j|k)$.

Let $\Pr(j|k)$ be the probability that the j individuals exhibit the favored behavior in environment $k = (1, 2)$. The learner uses Bayesian methods to infer $\Pr(k|y, j)$, the probability that the environment is in state k given that a learner observes an environmental cue y and j models with behavior k. Then the optimal decision rule is adopt behavior 1 if it has higher expected fitness.

$$\Pr(1|y, j)(1 + d_1) + \Pr(2|y, j) > \Pr(1|y, j) + \Pr(2|y, j)(1 + d_2)$$

or

$$\Pr(1|y, j) > \frac{d_1}{d_1 + d_2}$$

It is shown in the Appendix A.2 that this is equivalent to the following inequality.

$$j - a\frac{n}{2} > -gx + b$$

where

$$g = \frac{\frac{\mu}{v}}{\ln\left(\frac{p_1 p_2}{(1-p_1)(1-p_2)}\right)}$$

This is the same parameter as in the symmetric, unbiased case, except now modified to account for the asymmetric transition probabilities.

$$b = \frac{\ln\left(\frac{1-\pi}{\pi}\right) + \ln\left(\frac{d_1}{d_2}\right)}{\ln\left(\frac{p_1 p_2}{(1-p_1)(1-p_2)}\right)}$$

and

$$a = \frac{\ln\left(\frac{p_2}{1-p_1}\right)}{\ln\left(\frac{p_2}{1-p_1}\right) + \ln\left(\frac{p_1}{1-p_2}\right)}$$

This learning rule indicates that nature of content bias in favor of a behavior depends on whether that behavior has relatively (1) higher fitness in environments in which it is favored, or (2) has higher frequency environments in which it is favored. In the first case, learners are more likely to adopt that behavior independent of the number of cultural models who display that behavior, while in the second case, the number of models displaying a behavior necessary to motivate the learner to adopt that behavior is reduced. To see

this, set $d_1 = d_2$ so that the advantage of behavior 1 in environment one is the same as the advantage of behavior 2 in environment 2 and $\pi = 0.5$ so that both environments are equally likely. Then $b = 0$, and the right-hand side of the decision rule is then same as in the unbiased case studied by [18]. If $\pi > 0.5$ and $d_1 > d_2$, environment 1 is more likely and behavior one has a bigger relative payoff, then both terms in the numerator are negative and $b < 0$. This means that other things being equal, learners are more likely to adopt behavior 1. Similarly, if $\pi < 0.5$ and $d_1 < d_2$, learners are more likely to adopt behavior 2. If the terms have opposite sign then the effect on the decision depends on their relative magnitude. Finally, if $p_1 = p_2$, then $a = 1/2$, and the left-hand side reduces to the same expression as in the unbiased case. Since $p_1, p_2 > 0.5$.

$$\ln\left(\frac{p_1}{1 - p_2}\right) > \ln\left(\frac{p_2}{1 - p_1}\right) > 0$$

and therefore $1 > a > 0$. This means when behavior 1 is more common in environment 1 than behavior 2 is in environment 2, the rule favors behavior 1 when $j > \frac{an}{2}$. Even if fewer than half of the models exhibit behavior one, the social cue may favor the choice of behavior 1.

4. Discussion

The human species is unusual because people make much more use of social learning than other species, and this fact has important implications for understanding our evolutionary history and our current behavior [20,21]. Thus, it is important to understand the evolutionary forces that shape social learning. Previous work on this topic has been criticized because it assumes that social and individual learning are distinct processes. Perreault et al. [18] analyzed a model in which social and individual learning result from a single learning mechanism. However, this model did not allow for biased social learning.

Here, we extend this framework to allow for payoff bias and content bias. Our results are qualitatively similar to previous work suggesting that the assumption that social and individual learning are distinct processes is not crucial. We also derived more detailed predictions about the nature of both forms of biased social learning that can be tested with laboratory data on social learning.

The model present here has several limitations. First, it assumes that there are only two discrete behaviors. Many real-world behaviors have many possible variants. Second, the environment varies only in time. Human populations also face spatial variation. Third, sets of cultural models are assembled at random, and there is no age structure. While previous work (e.g., [2]) suggests that adding these complexities does not lead to qualitatively different predictions, one cannot be sure until the work is done.

These results suggest that natural selection will favor the use of payoff cues under a wide range of environmental conditions. People in small-scale societies solve many difficult adaptive problems. They depend on complex technologies and opaque ecological knowledge that arises in the solution of very hard problems. Such problems, particularly in fast changing environments strongly favor the use of payoff information. We believe that combining payoff information with social and environmental cues expanded the range of ecological habitats ancestral human populations could occupy and made these populations better able to survive the high frequency, high amplitude climatic fluctuations that prevailed during the Upper Pleistocene, and may have facilitated the rapid range of expansion of the human species at approximately 60 kya BP.

The power of general purpose learning mechanisms like payoff bias does not mean that human minds are predicted to be a blank slate. If some environments are reliably more frequent, or if some traits reliably have bigger fitness advantages, our results suggest that selection will favor learning rules that, all other things being equal, make it more likely that individuals will adopt traits that are more adaptive in expectation.

Appendix A

Here, we provide more details about the derivations sketched in the body of this paper. In the first section, we deal with payoff bias; in the second, we provide derivations for content bias. The code for the simulation is available at https://github.com/PerreaultC/Boyd-Perreault-Payoff-and-Content-Bias.git.

Appendix A.1. Optimal Social Learning with Payoff Information

Payoff bias occurs when social learners observe both behavior and payoff and are more likely to adopt behaviors that are statistically associated with higher payoffs. In this paper, we derive the optimal payoff bias rule, and study the conditions under which it is adaptive. We assume that there is also individual learning based on an environmental cue as in Perreault et al. (2012).

Consider a large population that lives in a environment which can exist in two states, imaginatively labeled *state 1* and *State 2*. Each generation, the environment switches with probability γ and stays the same with probability $1 - \gamma$. The means that over the long run, the environment is equally likely to be in each state. Individuals acquire one of two behaviors, *behavior 1* and *Behavior 2* by either individual or social learning. Behavior 1 is favored by selection when the environment is in state 1 and behavior 2 is favored when the environment is in state 2. The relative fitness advantage of the favored behavior in each environment is the same.

Each individual observes an environmental cue that can takes on a range of values. Let $\Pr(y|1)$ and $\Pr(y|2)$ be the probability that an individual observes cue value y in environments 1 and 2 respectively.

Each individual also observes n models sampled from the previous generation. For the jth model, social learners observe the behavior b_j and the individuals payoff, x_j. Otherwise models are identical. Thus, the entire sample is given by the pair of vectors $\mathbf{b} = (b_1, b_2, \ldots b_n)$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Let $\Pr(b_j, x_j|k)$ be the joint probability that the jth individual has behavior b_j and payoff x_j in environment k. The probabilities of payoffs of different models conditioned on behavior and the state of the environment are independent so $\Pr(\mathbf{x}, \mathbf{b}|k) = \prod \Pr(x_j, b_j|k)$.

Let $\Pr(1|y, \mathbf{x}, \mathbf{b})$ be the probability that the environment is in state 1 given that a learner observes an environmental cue y and models with behaviors $\mathbf{b}$ and payoffs $\mathbf{x}$. Then the optimal decision rule is adopt behavior 1 if $\Pr(1|y, \mathbf{x}, \mathbf{b}) > 1/2$ otherwise adopt behavior 2.

The first step in calculating $\Pr(1|y, \mathbf{x}, \mathbf{b})$ is to calculate the joint probability that the environment is in state 1, and that the individual observes y, $\mathbf{x}$ and $\mathbf{b}$, $Pr(1, y, \mathbf{x}, \mathbf{b})$.

$$\Pr(1, y, \mathbf{x}, \mathbf{b}) = \Pr(y|1)\Pr(\mathbf{x}, \mathbf{b}|1)\Pr(1) = \Pr(y|1)\Pr(\mathbf{x}|\mathbf{b}, 1)\Pr(\mathbf{b}|1)\Pr(1)$$

Now, we use Bayes' law to calculate $\Pr(1|y, \mathbf{x}, \mathbf{b})$

$$\Pr(1|y, \mathbf{x}, \mathbf{b}) = \frac{\Pr(1, y, \mathbf{x}, \mathbf{b})}{\Pr(y, \mathbf{x}, \mathbf{b})}$$

However, $\Pr(y, \mathbf{x}, \mathbf{b}) = \Pr(y, \mathbf{x}, \mathbf{b}|1)\Pr(1) + \Pr(y, \mathbf{x}, \mathbf{b}|2)\Pr(2)$. Using the expression for these conditional joint probabilities and the fact that the environment is equally likely to be in either state so $\Pr(1) = \Pr(2)$, this becomes

$$\Pr(1|y, \mathbf{x}, \mathbf{b}) = \frac{\Pr(\mathbf{x}|\mathbf{b}, 1)\Pr(\mathbf{b}|1)\Pr(y|1)}{\Pr(y|1)\Pr(\mathbf{x}|\mathbf{b}, 1)\Pr(\mathbf{b}|1) + \Pr(y|2)\Pr(\mathbf{x}|\mathbf{b}, 2)\Pr(\mathbf{b}|2)}$$

Thus, the probability that the environment is in state 1 given the cues is greater than $1/2$ if

$$\begin{aligned}
\Pr(y|1)\Pr(\mathbf{x}|\mathbf{b}, 1)\Pr(\mathbf{b}|1) \quad > \quad & \frac{1}{2}\Pr(y|1)\Pr(\mathbf{x}|\mathbf{b}, 1)\Pr(\mathbf{b}|1) \\
& + \frac{1}{2}\Pr(y|2)\Pr(\mathbf{x}|\mathbf{b}, 2)\Pr(\mathbf{b}|2) \\
\Pr(y|1)\Pr(\mathbf{x}|\mathbf{b}, 1)\Pr(\mathbf{b}|1) \quad > \quad & \Pr(y|2)\Pr(\mathbf{x}|\mathbf{b}, 2)\Pr(\mathbf{b}|2) \\
\frac{\Pr(\mathbf{b}|1)}{\Pr(\mathbf{b}|2)} \quad > \quad & \frac{\Pr(\mathbf{x}|\mathbf{b}, 2)}{\Pr(\mathbf{x}|\mathbf{b}, 1)}\frac{\Pr(y|2)}{\Pr(y|1)}
\end{aligned}$$

So far, the derivation is completely general. Models can be sampled in any way, and the payoffs can take any distribution. Now, let us first assume that models are sampled at random from the previous generation. To do this, let p be the expected frequency of behavior 1 given that the population is experiencing environment 1. Due to the symmetry of the model, p is also the expected frequency of behavior 2 given that the population is experiencing environment 2. Then $\Pr(\mathbf{b}|1) = p^j(1-p)^{n-j}$ and $\Pr(\mathbf{b}|2) = p^{n-j}(1-p)^j$, where j is the number of individuals with the favored behavior. Second, assume that the payoff of an individual with the favored behavior is a normally distributed random variable with mean $\mu + d$ and variance v and the payoff an individual with the disfavored behavior is a normally distributed variable with mean μ and variance v. The common mean, μ, is itself a normally distributed random variable with mean zero and a very large variance, V. This means that the absolute magnitudes of payoffs provide no information about the state of the environment, but that the difference between the payoffs is informative. Finally, assume the environmental cue in is a normally distributed random variable with mean ξ and variance ω when the environment is in state 1 and has mean $-\xi$ and ω when the environment is in state 2.

With these assumptions, the optimal decision rule is to adopt behavior 1 if

$$\frac{p^j(1-p)^{n-j}}{p^{n-j}(1-p)^j} > \frac{\int N(\mathbf{x}, \mu)d\mu}{\int D(\mathbf{x}, \mu)d\mu}\frac{\exp\left(\frac{-(y+\mu)^2}{2\omega}\right)}{\exp\left(\frac{-(y-\mu)^2}{2\omega}\right)}$$

where

$$N(\mathbf{x}, \mu) = \exp\left(\frac{-\mu^2}{2V}\right)\exp\left(-\frac{\sum_{b_i=2}(x_i - \mu - d)^2 + \sum_{b_i=1}(x_i - \mu)^2}{2v}\right)$$

and

$$D(\mathbf{x}, \mu) = \exp\left(\frac{-\mu^2}{2V}\right)\exp\left(-\frac{\sum_{b_i=1}(x_i - \mu - d)^2 + \sum_{b_i=2}(x_i - \mu)^2}{2v}\right)$$

Now, we manipulate the integrands, first in the numerator. Let

$$
\begin{aligned}
N(\mathbf{x}, \mu) &= \exp\left(\frac{-\mu^2}{2V} - \frac{\sum_{b_i=2}(x_i - \mu - d)^2 + \sum_{b_i=1}(x_i - \mu)^2}{2v}\right) \\
&= \exp\left(\frac{-\sum(x_i^2 - 2\mu x_i + \mu^2) - \sum_{b_i=2}(-2dx_i + 2\mu d + d^2)}{2v} - \frac{\mu^2}{2V}\right) \\
&= \exp\left(\frac{-2(n-j)\mu d + 2\mu n\bar{x} - n\mu^2}{2v} - \frac{\mu^2}{2V}\right) \\
&\quad \exp\left(-\frac{\sum x_i^2}{2v}\right)\exp\left(-\frac{(n-j)d^2 - 2d(n-j)\bar{x}_2}{2v}\right)
\end{aligned}
$$

where $\bar{x}_1$ is the mean of the n observed payoff values and $\bar{x}_2$ is the mean of the $(n-j)$ observed payoff values for individuals with behavior 2. Since the right two terms do not involve μ, they can be taken outside of the integral. Then rewriting the term containing μ

$$
\exp\left(\frac{-2(n-j)\mu d + 2\mu = n\bar{x} - n\mu^2}{2v} - \frac{\mu^2}{2V}\right)\exp\left(-\frac{k_1}{2}\left(\mu^2 - 2k_2\mu\right)\right)
$$

where $k_1 = \frac{1}{V} + \frac{n}{v}$ and $k_2 = \frac{\left(\frac{n\bar{x} - (n-j)d}{v}\right)}{\left(\frac{1}{V} + \frac{n}{v}\right)}$. Then

$$
\exp\left(\frac{-2(n-j)\mu d + 2\mu n\bar{x} - n\mu^2}{2v} - \frac{\mu^2}{2V}\right) = \exp\left(-\frac{k_1}{2}\left((\mu - k_2)^2\right)\right)\exp\left(\frac{1}{2}k_1 k_2^2\right)
$$

This means that

$$
\begin{aligned}
\exp\left(\frac{-2(n-j)\mu d + 2\mu n\bar{x} - n\mu^2}{2v} - \frac{\mu^2}{2V}\right) &= \exp\left(-\frac{k_1}{2}\left((\mu - k_2)^2\right)\right) \\
&\quad \exp\left(\frac{\left(\frac{n\bar{x} - (n-j)d}{v}\right)^2}{2\left(\frac{1}{V} + \frac{n}{v}\right)}\right)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\lim_{V\to\infty}\int N(x, \mu)d\mu &= \exp\left(-\frac{\sum x_i^2}{2v}\right)\exp\left(\frac{-(n-j)d^2 + 2(n-j)d\bar{x}_2}{2v}\right) \\
&\quad \lim_{V\to\infty}\left(\exp\left(\frac{\left(\frac{n\bar{x} - (n-j)d}{v}\right)^2}{2\left(\frac{1}{V} + \frac{n}{v}\right)}\right)\int \exp\left(-\frac{k_1}{2}\left((\mu - k_2)^2\right)\right)d\mu\right) \\
&= \exp\left(-\frac{\sum x_i^2}{2v}\right)\exp\left(-(n-j)d\left(\frac{d - 2\bar{x}_2}{2v}\right)\right) \\
&\quad \exp\left(\frac{(n\bar{x} - (n-j)d)^2}{2nv}\right)C
\end{aligned}
$$

where $C = \int \exp\left(-\frac{k_1}{2}(\mu - k_2)^2\right)d\mu$ is a constant that depends on k_1 but not k_2. A similar derivation yields an expression for the denominator of the fraction on the right-hand side of the decision rule:

$$
\lim_{V\to\infty}\int D(x, \mu)d\mu = \exp\left(\frac{(n\bar{x} - jd)^2}{2nv}\right)\exp\left(-\frac{\sum x_i^2}{2v}\right)\exp\left(-jd\left(\frac{d - 2\bar{x}_1}{2v}\right)\right)C
$$

where $\bar{x}_2$ is the mean payoff of individuals who use behavior 2. Then the decision rule becomes

$$\left(\frac{p}{1-p}\right)^{(2j-n)} > \exp\left(\frac{(n\bar{x})^2 - 2n\bar{x}(n-j)d + ((n-j)d)^2 - (n\bar{x})^2 + 2n\bar{x}jd - (jd)^2}{2nv}\right)$$

$$\exp\left(\frac{jd(d-2\bar{x}_1) - (n-j)d(d-2\bar{x}_2)}{2v}\right)$$

$$\exp\left(\frac{-(y^2 + 2\xi y + \xi^2) + (y^2 - 2\xi y + \xi^2)}{2\omega}\right)$$

$$> \exp\left(\frac{-2n^2\bar{x}d + 4n\bar{x}jd + n^2d^2 - 2njd^2}{2nv}\right)$$

$$\exp\left(d\frac{j(d-2\bar{x}_1) - (n-j)(d-2\bar{x}_2)}{2v}\right)\exp\left(\frac{-2\xi y}{\omega}\right)$$

$$> \exp\left(d\frac{-2\bar{x}\left(\frac{n}{2}-j\right) + d\left(j-\frac{n}{2}\right)}{v}\right)$$

$$\exp\left(d\frac{d\left(j-\frac{n}{2}\right) - j\bar{x}_1 + (n-j)\bar{x}_2}{v}\right)\exp\left(\frac{-2\xi y}{\omega}\right)$$

$$> \exp\left(\frac{-\bar{x}(n-2j) - j\bar{x}_1 + (n-j)\bar{x}_2}{\frac{v}{d}}\right)\exp\left(\frac{-2\xi y}{\omega}\right)$$

$$> \exp\left(\frac{-\left(\frac{j}{n}\bar{x}_1 + \frac{n-j}{n}\bar{x}_2\right)(n-2j) - j\bar{x}_1 + (n-j)\bar{x}_2}{\frac{v}{d}}\right)\exp\left(\frac{-2\xi y}{\omega}\right)$$

$$> \exp\left(\frac{-j\left(1 + \frac{(n-2j)}{n}\right)\bar{x}_1 + (n-j)\left(1 - \frac{(n-2j)}{n}\right)\bar{x}_2}{\frac{v}{d}}\right)\exp\left(\frac{-2\xi y}{\omega}\right)$$

$$> \exp\left(\frac{-j\left(\frac{n+(n-2j)}{n}\right)\bar{x}_1 + (n-j)\left(\frac{n-(n-2j)}{n}\right)\bar{x}_2}{\frac{v}{d}}\right)\exp\left(\frac{-2\xi y}{\omega}\right)$$

$$> \exp\left(d\left(\frac{\frac{-2j(n-j)}{n}\bar{x}_1 + \frac{2j(n-j)}{n}\bar{x}_2}{v}\right)\right)\exp\left(\frac{-2\xi y}{\omega}\right)$$

$$> \exp\left(2d\frac{j(n-j)}{n}\left(\frac{\bar{x}_2 - \bar{x}_1}{v}\right)\right)\exp\left(\frac{-2\xi y}{\omega}\right)$$

Taking the log of both sides of the inequality yields

$$\left(j - \frac{n}{2}\right)\ln\left(\frac{p}{1-p}\right) > \frac{d}{v}\left(\frac{j(n-j)}{n}\right)(\bar{x}_2 - \bar{x}_1) - \frac{\xi}{\omega}y$$

or

$$j - \frac{n}{2} > G\left(\frac{j(n-j)}{n}\right)(\bar{x}_2 - \bar{x}_1) - gy$$

where

$$G = \frac{d}{v\ln\left(\frac{p}{1-p}\right)}$$

and

$$g = \frac{\xi}{\omega\ln\left(\frac{p}{1-p}\right)}$$

If the environment is in state 1, the expected value of $\bar{x}_1$ is $\mu + d$ and the expected value of $\bar{x}_2$ is μ. The variance of the sum of two normally distributed random variables is the sum of the variances so $\bar{x}_2 - \bar{x}_1$ is normally distributed with mean $-d$ and variance

$v\left(\frac{1}{j} + \frac{1}{n-j}\right) = \frac{nv}{j(n-j)}$. In environment 2, the variance is the same but the mean is d. As before, y is normally distributed, but now with mean ξ and variance ω. So for an individual with genotype g_i and G_i, you can calculate the right-hand side of the inequality from the sum of two cumulative distributions, and then proceed as before.

Appendix A.2. Optimal Content Bias

The first step in calculating $\Pr(1|x_i, j)$ is to calculate the joint probability that the environment is in state 1, and that the individual observes x_1 and j, $Pr(1, x_i, j)$. Since $\Pr(x_i|1)$ and $\Pr(j|1)$ are independent it follows that

$$\Pr(x_i, j|1) = \Pr(x_i|1)\Pr(j|1)$$

and thus

$$\Pr(1, x_i, j) = \Pr(x_i, j|1)\Pr(1) = \Pr(x_i|1)\Pr(j|1)\Pr(1)$$

Now, we use Bayes' law to calculate $\Pr(1|x_i, j)$

$$\Pr(1|x_i, j) = \frac{\Pr(1, x_i, j)}{\Pr(x_i, j)}$$

However, $\Pr(x_i, j) = \Pr(x_i, j|1)\Pr(1) + \Pr(x_i, j|2)\Pr(2)$. Using the expression for these conditional joint probabilities derived above, this becomes

$$\Pr(1|x_i, j) = \frac{\Pr(x_i|1)\Pr(j|1)\Pr(1)}{\Pr(x_i|1)\Pr(j|1)\Pr(1) + \Pr(x_i|2)\Pr(j|2)\Pr(2)}$$

The probability that the environment is in state 1 is $\Pr(1) = \pi$ and the probability it is in state 2 is $\Pr(2) = 1 - \pi$ thus

$$\Pr(1|x_i, j) = \frac{\Pr(x_i|1)\Pr(j|1)\pi}{\Pr(x_i|1)\Pr(j|1)\pi + \Pr(x_i|2)\Pr(j|2)(1 - \pi)}$$

It is useful to rewrite this expression as

$$\Pr(1|x_i, j) = \frac{\frac{\Pr(j|1)}{\Pr(j|2)}\pi}{\frac{\Pr(j|1)}{\Pr(j|2)}\pi + \frac{\Pr(x_i|2)}{\Pr(x_i|1)}(1 - \pi)}$$

Substituting the expression for $\Pr(1|x_i, j)$ into the optimal decision rule yields

$$\frac{\frac{\Pr(j|1)}{\Pr(j|2)}\pi}{\frac{\Pr(j|1)}{\Pr(j|2)}\pi + \frac{\Pr(x_i|2)}{\Pr(x_i|1)}(1 - \pi)} > \frac{d_1}{d_1 + d_2}$$

$$\frac{\Pr(j|1)}{\Pr(j|2)}\pi > \left(\frac{d_1}{d_1 + d_2}\right)\left(\frac{\Pr(j|1)}{\Pr(j|2)}\pi + \frac{\Pr(x_i|2)}{\Pr(x_i|1)}(1 - \pi)\right)$$

$$\frac{\Pr(j|1)}{\Pr(j|2)}\pi\left(1 - \frac{d_1}{d_1 + d_2}\right) > \frac{\Pr(x_i|2)}{\Pr(x_i|1)}(1 - \pi)\left(\frac{d_1}{d_1 + d_2}\right)$$

$$\frac{\Pr(j|1)}{\Pr(j|2)} > \frac{\Pr(x_i|2)}{\Pr(x_i|1)}\left(\frac{(1 - \pi)\left(\frac{d_1}{d_1 + d_2}\right)}{\pi\left(1 - \frac{d_1}{d_1 + d_2}\right)}\right)$$

$$\frac{\Pr(j|1)}{\Pr(j|2)} > \frac{\Pr(x_i|2)}{\Pr(x_i|1)}\left(\frac{1 - \pi}{\pi}\right)\left(\frac{d_1}{d_2}\right)$$

Notice that when environments are equally likely ($\pi = 1/2$) and payoff advantages are symmetrical ($d_1 = d_2$), this condition is the same as that given in [18].

Now, assume that (1) the expected frequency of individuals with behavior 1 in environment 1 is p_1 and the frequency of individuals with behavior 2 in environment 2 is p_2 and that models are sampled at random from the previous generation so that $\Pr(j|1)$ and $\Pr(j|2)$ binomial with parameters p_1 and n, and p_2 and n, and (2) the environmental cue is a normally distributed random variable with mean μ and variance v when the environment is in state 1 and has mean $-\mu$ and v when the environment is in state 2. This means that individuals adopt behavior 1 if

$$
\frac{\frac{n!}{j!(n-j)!}p_1^j(1-p_1)^{n-j}}{\frac{n!}{j!(n-j)!}p_2^{n-j}(1-p_2)^j} > \frac{e^{-\frac{(x+\mu)^2}{2v}}}{e^{-\frac{(x-\mu)^2}{2v}}}\left(\frac{1-\pi}{\pi}\right)\left(\frac{d_1}{d_2}\right)
$$

This expression can be simplified to become

$$
\left(\frac{p_2}{1-p_1}\right)^{j-n}\left(\frac{p_1}{1-p_2}\right)^{j} > e^{-\frac{x\mu}{v}}\left(\frac{1-\pi}{\pi}\right)\left(\frac{d_1}{d_2}\right)
$$

Since the logarithm function is monotonic, we can take the logarithm of both sides of the inequality which yields a simple linear form for the decision rule

$$
j\ln\left(\frac{p_1}{1-p_2}\right) + (j-n)\ln\left(\frac{p_2}{1-p_1}\right) > -\frac{x\mu}{v} + \ln\left(\left(\frac{1-\pi}{\pi}\right)\left(\frac{d_1}{d_2}\right)\right)
$$

$$
j\ln\left(\frac{p_1 p_2}{(1-p_1)(1-p_2)}\right) - n\ln\left(\frac{p_2}{1-p_1}\right) > -\frac{x\mu}{v} + \ln\left(\left(\frac{1-\pi}{\pi}\right)\left(\frac{d_1}{d_2}\right)\right)
$$

$$
j - n\frac{\ln\left(\frac{p_2}{1-p_1}\right)}{\ln\left(\frac{p_1 p_2}{(1-p_1)(1-p_2)}\right)} > \frac{\frac{\mu}{v}x + \ln\left(\left(\frac{1-\pi}{\pi}\right)\left(\frac{d_1}{d_2}\right)\right)}{\ln\left(\frac{p_1 p_2}{(1-p_1)(1-p_2)}\right)}
$$

This leads to the three parameter decision rule

$$
j - a\frac{n}{2} > -gx + b
$$

where

$$
g = \frac{\frac{\mu}{v}}{\ln\left(\frac{p_1 p_2}{(1-p_1)(1-p_2)}\right)}
$$

This is the same parameter as in the symmetric, unbiased case.

$$
b = \frac{\ln\left(\frac{1-\pi}{\pi}\right) + \ln\left(\frac{d_1}{d_2}\right)}{\ln\left(\frac{p_1 p_2}{(1-p_1)(1-p_2)}\right)}
$$

First, suppose that $d_1 = d_2$ so that the advantage of behavior 1 in environment one is the same as the advantage of behavior 2 in environment 2 and $\pi = 0.5$ so that both environments are equally likely, $b = 0$, and the right-hand side of the decision rule is the same as in the unbiased case studied by [18]. If $\pi > 0.5$ and $d_1 > d_2$, so environment 1 is more likely and behavior one has a bigger relative payoff, then both terms in the numerator are negative and $b < 0$. This means that other things being equal, learners are more likely to adopt behavior 1. Similiarly, if $\pi > 0.5$ and $d_1 > d_2$, learners are more likely to adopt behavior 2. If the terms have opposite sign then the effect on the decision depends on their relative magnitude. Finally,

$$
a = \frac{\ln\left(\frac{p_2}{1-p_1}\right)}{\ln\left(\frac{p_2}{1-p_1}\right) + \ln\left(\frac{p_1}{1-p_2}\right)}
$$

So, if $p_1 = p_2$ then $a = 1/2$, and the left-hand side reduces to the same expression as in the unbiased case. Since $p_1, p_2 > 0.5$.

$$\ln\left(\frac{p_1}{1 - p_2}\right) > \ln\left(\frac{p_2}{1 - p_1}\right) > 0$$

and therefore $1 > a > 0$. This means when environment 1 is more common, the rule favors behavior 1 when $j > \frac{an}{2}$. Even if fewer than half of the models exhibit behavior one, the social cue may say choose behavior 1.

References

1. Cavalli-Sforza, L.L.; Feldman, M.W. *Cultural Transmission and Evolution: A Quantitative Approach*; Princeton University Press: Princeton, NJ, USA, 1981.
2. Boyd, R.; Richerson, P.J. *Culture and the Evolutionary Process*; University of Chicago Press: Chicago, IL, USA, 1985.
3. Rogers, A.R. Does biology constrain culture? *Am. Anthropol.* **1988**, *90*, 819–831. [CrossRef]
4. Kameda, T.; Nakanishi, D. Does social/cultural learning increase human adaptability?: Rogers's question revisited. *Evol. Hum. Behav.* **2003**, *24*, 242–260. [CrossRef]
5. McElreath, R.; Strimling, P. When natural selection favors imitation of parents. *Curr. Anthropol.* **2008**, *49*, 307–316. [CrossRef]
6. Rendell, L.; Boyd, R.; Cownden, D.; Enquist, M.; Eriksson, K.; Feldman, M.W.; Fogarty, L.; Ghirlanda, S.; Lillicrap, T.; Laland, K.N. Why copy others? Insights from the social learning strategies tournament. *Science* **2010**, *328*, 208–213. [CrossRef] [PubMed]
7. Kendal, R.L.; Boogert, N.J.; Rendell, L.; Laland, K.N.; Webster, M.; Jones, P.L. Social learning strategies: Bridge-building between fields. *Trends Cogn. Sci.* **2018**, *22*, 651–665. [CrossRef] [PubMed]
8. Morgan, T.J.; Acerbi, A.; Van Leeuwen, E.J. Copy-the-majority of instances or individuals? Two approaches to the majority and their consequences for conformist decision-making. *PLoS ONE* **2019**, *14*, e0210748. [CrossRef] [PubMed]
9. Schlag, K.H. Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits. *J. Econ. Theory* **1998**, *78*, 130–156. [CrossRef]
10. Schlag, K.H. Which one should I imitate? *J. Math. Econ.* **1999**, *31*, 493–522. [CrossRef]
11. Hirshleifer, D.; Plotkin, J.B. Moonshots, investment booms, and selection bias in the transmission of cultural traits. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, 1–9. [CrossRef] [PubMed]
12. Kendal, J.; Giraldeau, L.A.; Laland, K. The evolution of social learning rules: Payoff-biased and frequency-dependent biased transmission. *J. Theor. Biol.* **2009**, *260*, 210–219. [CrossRef] [PubMed]
13. Miu, E.; Morgan, T.J. Cultural adaptation is maximised when intelligent individuals rarely think for themselves. *Evol. Hum. Sci.* **2020**, *2*, 1–18. [CrossRef]
14. Fogarty, L.; Kandler, A. The fundamentals of cultural adaptation: Implications for human adaptation. *Sci. Rep.* **2020**, *10*, 1–11. [CrossRef]
15. Mesoudi, A. Cultural evolution: A review of theory, findings and controversies. *Evol. Biol.* **2016**, *43*, 481–497. [CrossRef]
16. Heyes, C.M. Social learning in animals: Categories and mechanisms. *Biol. Rev.* **1994**, *69*, 207–231. [CrossRef]
17. Laland, K.N. Social learning strategies. *Anim. Learn. Behav.* **2004**, *32*, 4–14. [CrossRef] [PubMed]
18. Perreault, C.; Moya, C.; Boyd, R. A Bayesian approach to the evolution of social learning. *Evol. Hum. Behav.* **2012**, *33*, 449–459. [CrossRef]
19. Perfors, A.; Tenenbaum, J.B.; Griffiths, T.L.; Xu, F. A tutorial introduction to Bayesian models of cognitive development. *Cognition* **2011**, *120*, 302–321. [CrossRef]
20. Boyd, R. *A Different Kind of Animal*; Princeton University Press: Princeton, NJ, USA, 2017.
21. Henrich, J. *The Secret of our Success*; Princeton University Press: Princeton, NJ, USA, 2015.