THE EVOLUTION OF SIMILARITY-BIASED SOCIAL LEARNING

PAUL E. SMALDINO[1,2,*] AND ALEJANDRO PÉREZ VELILLA[1]

May 24, 2023

1. Introduction

Cultural evolutionary theory has focused a lot on social learning, and with good reason. Dual inheritance implies two channels for the transmission of traits—genes and cultural inheritance, the latter of which operates largely (though not entirely) through learning from others. Humans should rely more on social learning in cases where individual learning cant be trusted—e.g., in noisy, complex worlds with many options (Boyd and Richerson 1985; Turner et al. 2023). However, if the word is so noisy that what has worked for others is unlikely to work for us, social transmission may also fail to produce adaptive traits.

Here, we model the evolution of social learning strategies and allow for the coevolution of a bias for similarity. When individual learning is unreliable *and* there is diversity in the strategies that work for different people, a bias for learning only from similar others—ignoring social information from outgroup individuals—can be adaptive.

1.1. **Narrative Plan.** This paper will proceed with a number of modeling exercises to demonstrate the clarity and robust of the theory on offer here. Unlike many models, we do not treat individual learning and social learning as mutually exclusive activities. Rather, we follow both seminal (Boyd and Richerson 1985) and current (Turner et al. 2023) work in portraying social learning as an activity that can act in concert with individual learning.

(1) **Social learning evolves when individual learning is more uncertain.** We will replicate findings from a model by Boyd and Richerson (1985; chapter 4) and show that when individual learning becomes too noisy, unbiased social learning can evolve. We will focus on unbiased transmission (copying a random individual), and later explore other learning strategies.
(2) **Social learning relies on correlated environments.** When there are multiple traits that are not adaptive, there is less social learning. We will consider two groups, each of which has a different adaptive trait value, where individuals cannot distinguish between group members. The further apart the two adaptive trait values are, the weaker the selection for social learning. We will further

[1]University of California, Merced
[2]Santa Fe Institute
E-mail address: paul.smaldino@gmail.com.

explore differential group size, where there is a minority group for whom fewer individuals share the adaptive trait. We will show that social learning is still favored in large majority groups, because more individuals will share their adaptive trait values.

(3) **Similarity-bias recovers the evolution of social learning.** We will allow individuals to coevolve a second trait: parochialism (or similarity-bias). This is the disregard of social information from non-group members. We will give individuals group markers, which signify their association with a particular adaptive trait value with some reliability. We will see that the more reliably social markers indicate adaptive trait values, the more parochialism and therefore social learning are favored. We will also see that minority groups are influenced differentially. When targets for social learning are few in number, parochial social learning is not favored. When targets are plentiful, it is.

(4) **Our findings are robust to conformist and success-biased transmission.** We will repeat the above analyses where either conformist or success-biased transmission can also evolve. These strategies will tend to be favored over unbiased transmission, but the general logic of noisy individual learning, group variation, and similarity-bias all hold.

(5) **Parochialism is sticky.** Finally, we will inititalize populations with parochial social learning, in conditions where it does not initially evolve from scratch. How much does it persevere when it is no longer adaptive?

(6) **Further thoughts.** Social learning, including similarity-biased learning, should work differently when there are many traits, with different payoffs, which vary in their correlation with identity. This analysis is beyond the scope of this paper, but we should speculate. A future model might examine multiple traits, only some of which are correlated with group identity. What happens if one of the uncorrelated traits is very costly to get wrong, and there is a sort of game(s) theory bounded rationality so that individuals are stuck with only one sort of social learning strategy?

Note that this model is one of stabilizing selection, and we do not take into account environmental change in which the adaptive trait values change.

2. Model Description

2.1. **A model of the evolution of similarity-biased social learning.**

Consider a population of N individuals divided into 2 intermixed groups (representing norms, affordances, etc. in a cosmopolitan society). The frequency of group 0 is f, and the frequency of group 1 is $1 - f$. Each individual i is characterized by five heritable traits: their type (or group) $g_i \in \{0, 1\}$, their marker $m_i \in \{0, 1\}$, their parochialism $p_i \in \{0, 1\}$, their reliance on social learning $\rho_i \in [0, 1]$, and their social learning strategy, s_i (see below). Note that all traits take on one of a small number (often two) of discrete values except for the reliance on social learning, ρ_i, which varies continuously.[1]

During their lifespan, each individual i acquires a trait $\mathbf{x}_i$ using one of several learning strategies. In order to avoid path dependent effects where initial trait values were biased toward a particular adaptive trait, we represent traits in continuous two-dimensional space using polar coordinates, so that $\mathbf{x}_i = (r_i, \theta_i)$. This is also a minimal representation of the idea that traits have multiple components. For each group j, there is an optimally adaptive trait value, $\mathbf{H}_j$, which maximizes the fitness an individual of that group can receive. Without loss of generality, we always set $\mathbf{H}_0 = (1, 0)$ and $\mathbf{H}_1 = (1, \theta)$ (see Figure 1). We restrict $\theta \in [0, 180°]$, representing the spectrum from alignment to orthogonal to opposed trait values. For convenience, we will use the notation $\mathbf{H}_i$ to the group-specific adaptive trait of an agent i.

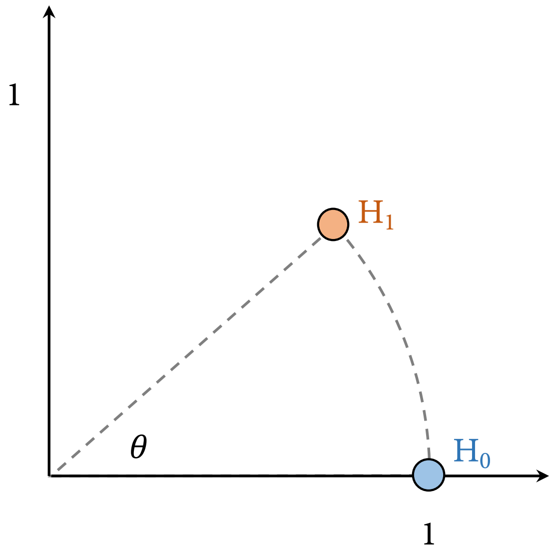

Figure 1. Depiction of the two adaptive traits, in polar coordinates. $H_0 = (1, 0)$, $H_1 = (1, \theta)$.

Individuals can acquire trait values through both individual and social learning. Social learning is presumed to be accurate, as individuals can readily see what others are doing. Individual learning, however, is less reliable, and for a member of group j produces a signal drawn randomly from a normal distribution with a mean of H_j and a standard deviation of σ_ℓ.

We will consider the heritability of social learning strategies as well as *parochialism*, which is a special category of social learning in which models are specifically target from among the learner's ingroup.

[1]Later, we may also want to consider simulations in which parochialism, p_i, also varies continuously (and therefore represents the probability of ignoring outgroup social information, but let's start with the discrete case.

2.2. **Initialization.** At initialization, agents are created and assigned to one of the two type-groups. Specifically, an agent is assigned to group 0 with probability f and group 1 otherwise. Agents are also assigned markers, which may or may not be reliable indicators of their group (that is, the adaptive trait that works for them). With probability R, agents are assigned the group marker that corresponds to their group, $m_i = g_i$. With probablility $1 - R$, agents are assigned a marker with probability equal to the frequency of the associated adaptive trait (i.e., they are assigned marker 0 with probability f and marker 1 with probability $1-f$), so as to keep the relative frequency of markers consistent across conditions. When R is close to 1, the markers are strongly correlated with the adaptive trait values of the agents who use them; when R is close to zero, they are not very informative. We assume that agents begin as non-parochial individual learners, so that $\forall_i, p_i = 0, \rho_i = 0, s_i = \text{UT}$.

Because the first generation relies entirely on individual learning, each agent i adopts a trait value equal to $\mathbf{H}_i + \mathcal{N}(0, \sigma_\ell)$, such that their learned trait values are points in Cartesian space drawn from a bivariate normal distribution centered on $\mathbf{H}_i$ and a normalized standard deviation of σ_ℓ. Each agent then acquires a payoff based on their trait value, described below.

2.3. **Dynamics.** After initialization, the model proceeds in discrete time steps, which are broken up into four stages: (1) reproduction, (2) model choice, (3) learning, and (4) payoff acquisition.

2.3.1. *Reproduction.* A new generation of N individuals is created. Each group produces a number of offspring equal to its current size (to keep both the population size and the relative group sizes constant). Each agent in the new generation has one parent, chosen from its own group with probability equal to the relative payoff of the parents in that group. Specifically, an agent i in group j in the parent generation is selected at random, and reproduces with a probability

$$\Pr(\text{reproduce}) = \frac{W_{ij}}{W_{\text{max}j}}, \tag{1}$$

where W_{ij} is the agent's payoff and $W_{\text{max}j}$ is the highest payoff in the group among the parent generation.

Each offspring inherits the social learning strategies and parochialism of its parent, though imperfectly. With probability $1 - \mu_R$, the offspring inherits the social learning reliance of its parent. Otherwise the agent i sets its social learning reliance to $\rho_i = \rho_k + \mathcal{N}(0, \sigma_R)$, where ρ_k is the social learning reliance of the parent, and the mutation amount is a randomly drawn value from a normal distribution with a mean of zero and a standard deviation of σ_R. The value of ρ_i will be truncated as necessary to remain in the range [0, 1].

For simulations with multiple social learning possible (e.g., conformist or success-biased learning), the social learning strategy can also evole. With probability $1 - \mu_L$, the offspring inherits the social learning strategy of its parent, and otherwise adopts a learning strategy at random from the set of allowable strategies.

Finally, parochialism can also evolve. With probability $1 - \mu_P$, the offspring inherits the parochialism of its parents, and otherwise adopts the parochial behavior opposite that of its parent.

2.3.2. *Model choice.* Each individual choose n other agents (targets) from the previous generation to learn from, $1 \leq n < N$. These targets will be chosen at random from the entire parent generation. The focal agent then excludes each target that does not share its marker with probability p_i. The agent uses its learning strategy (below) to learn from these targets. It is possible for a parochial agent to have zero targets left after excluding outgroup targets. In this case, the agent must rely exclusively on individual learning.

2.3.3. *Learning.* Each focal agent uses their learning strategy, s_i, to acquire a trait value from their set of targets. There are three possible options:

- Unbiased transmission (UT). Choose a target at random from those under consideration and register their trait value.
- Conformist transmission (CT). Adopt the median trait value among the targets.
- Payoff-biased transmission (PT). Adopt the trait from the target with the highest payoff. If two or more targets have the same high payoff, choose one at random.

In the first analyses, we will consider agents using exclusively UT. In later analyses, we will consider CT and PT separately, each in competition with UT.

Each agent i ascertains a socially learned trait value, $\mathbf{y}_i$ in this way, as well as an individually learned trait value, $\mathbf{z}_i$, such that

$$\mathbf{z}_i = \mathbf{H}_i + \mathcal{N}(0, \sigma_\ell), \tag{2}$$

such that their individually-learned trait values are points in Cartesian space drawn from a bivariate normal distribution centered on $\mathbf{H}_i$ and a normalized standard deviation of σ_ℓ. The trait the agent ultimately adopts is an average of the individually- and socially-learned values, weighted by their reliance on social learning.

$$\mathbf{x}_i = \rho_i \mathbf{y}_i + (1 - \rho_i)\mathbf{z}_i \tag{3}$$

Following learning, the agents in the previous generation die, ala Logan's Run.

2.3.4. *Payoffs.* Following Boyd and Richerson (1985), we use a Gaussian fitness function such that the maximum fitness for an agent in group j is $\mathbf{H}_j$, and their fitness declines with their trait's Euclidean distance from that value. Written out, the fitness $W(\mathbf{x}_i, \mathbf{H}_i)$ to an agent i is given by the following:

$$W(\mathbf{x}_i, \mathbf{H}_i) = \exp\left[-\frac{\|\mathbf{x}_i - \mathbf{H}_i\|^2}{2S}\right], \tag{4}$$

where the double-vertical lines represent the Euclidean distance between the two points, and the denominator S is a parameter controlling the intensity of selection (see Figure 2).

The first draft of this model was coded by Paul Smaldino in NetLogo 6.3.

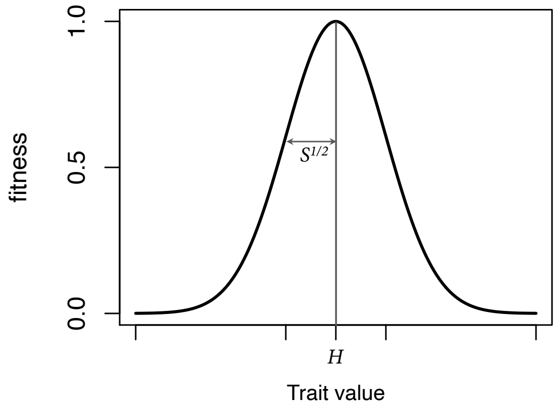

FIGURE 2. Gaussian fitness function. The horizontal axis represents the distance of the agent's trait value from their group-specific optimum, H, and the vertical axis is the corresponding payoff.

TABLE 1. Global parameters. Important to test them before doing all the batch runs.

Parameter	**Meaning**	**Value(s)**
N	Population size	{100, 1000}
f	Relative size of larger group	[0.5, 1]
H_0	Adaptive trait for group 0	1
θ	Distance between the two adaptive traits	$[0, 180°]$
n	Number of models for social learning	{1, 5, 15}
σ_ℓ	Individual learning uncertainty	$[0, 0.5]$
μ_P	Mutation rate for parochialism	{0, 0.01}
μ_L	Mutation rate for learning strategy	{0, 0.01}
μ_R	Mutation rate for social learning reliance	0.01
σ_R	Standard deviation of social learning reliance mutation	0.05
S	Intensity of selection	0.05

TABLE 2. Agent parameters.

Parameter	**Meaning**	**Values**
g_i	Agent type/group	$\{0, 1\}$
m_i	Agent marker	$\{0, 1\}$
p_i	Parochialism	$\{0, 1\}$
ρ_i	Reliance on social learning	$[0, 1]$
s_i	Learning strategy	{UT, CT, PT }

2.4. **Outcome measures.** We will keep track of the following outcome measures:

- Average reliance on social learning (ρ_i) in each group.
- Frequency of each learning strategy (s_i) in each group.
- Frequency of parochialism (p_i) (total and for each learning strategy) in each group.
- Average group and population payoff (total and for each learning strategy).

3. Analysis Plan

The results will be presented in the following order. Here are the simulation conditions for each of the subsection in the Results section.

3.1. **Social learning evolves when individual learning is more uncertain.**

- Fix $\theta = 0$, $f = 0$.
- No parochialism, and unbiased social learning only. Keep $\mu_P = \mu_L = 0$. Fix $n = 1$ since it doesn't matter.
- The main target of variation is the individual learning error, σ_ℓ. Vary this in intervals of 0.01 (which was 0.1 in the NetLogo sims).
- Run all conditions where social learning cannot evolve, and show how it facilitates higher payoffs.

3.2. **Social learning relies on correlated environments.**

- Vary $\theta \in [0, 180]$, in intervals of 5.
- Also vary the size of the minority group, $f \in [0.5, 1]$ in units of 0.1.
-

3.3. **Similarity-biased social learning mitigates the problems of uncorrelated environments.**

- Increase μ_P to something small but positive (0.01) to allow parochialism to evolve.
- Vary the number of models for social learning, $n \in \{1, 5, 15\}$. This will be important, especially the difference between $n = 1$ and $n > 1$ (might focus on just two cases).
- Initially focus on equal group size, but also examine differential group size ($f \in [0.5, 1]$).
- Vary θ as before.
- Vary the identity correlation, $R \in [0, 1]$.
- Consider how parochialism evolves when $theta$ is large, and R is sufficiently large.

3.4. **The integration of more complex learning strategies.** Re-run all the analyses above for cases in which agents can also evolve other social learning strategies: conformist and success-biased social learning. Let $\mu_L = 0.01$. In other words, there will be a batch identical to the runs above where CT (but not PT) is allowed to evolve, and another batch where PT (but not CT) is allowed to evolve. The case where both can evolve is not important to consider (I think).

3.5. **The stubbornness of parochial social learning.** If parochialism evolves, we will at some point examine the emergence of a new adaptive trait that is adaptive for both individuals. We can then examine how much parochialism persists. It may go away, but it may not. Initialize the model where all agents are already parochial and rely on social learning. Consider the case where everyone either has the same trait or group ids aren't well correlated with adaptive behaviors (R is small). Does parochialism stick, or go away?

4. Preliminary Results

Here are some preliminary results. All these are with $N = 100$, $n = 1$, $\sigma_\ell = 0.1$, $\mu_P \mu_R = 0.01$, $\sigma_R = S = 0.05$. The values shown in the axes may vary due to the scaling issue in the NetLogo model, in which the placement of adaptive trait values were at $r = 10$ instead of $r = 1$ for visualization purposes. The first two figures are from 100 runs of each condition, run out to 10,000 time steps. The later figures are from fewer runs, run out for less time and possibly not run to equilibrium.

4.1. **Social learning evolves when individual learning is more uncertain.** As individual learning uncertainty (σ_ℓ) increases, reliance on social learning also increases (Fig 3A). As this happens, the average payoff decreases in all cases since learning is more error prone. However, compared to the case where social learning cannot evolve, social learning does much better (Fig 3B).

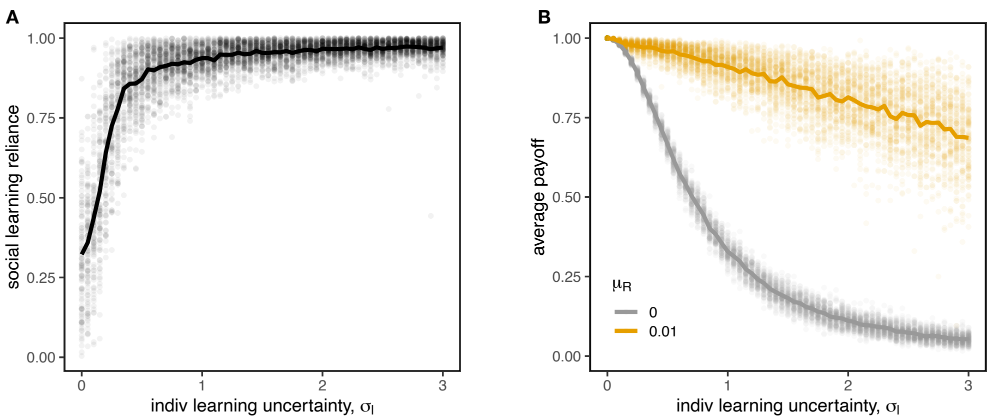

Figure 3

4.2. **Social learning relies on correlated environments.** Here we ran simulations for a condition where social learning tends to evolve where all agents have the same adaptive trait value. In this simulation, however, there are two different adaptive trait values, which differ by the amount θ. As θ increases, social learning is no longer favored. The only exceptions are for groups in the extreme minority (the larger group for $f > 0.5$), since most random individuals for them will tend to have the same adaptive trait (Fig 4).

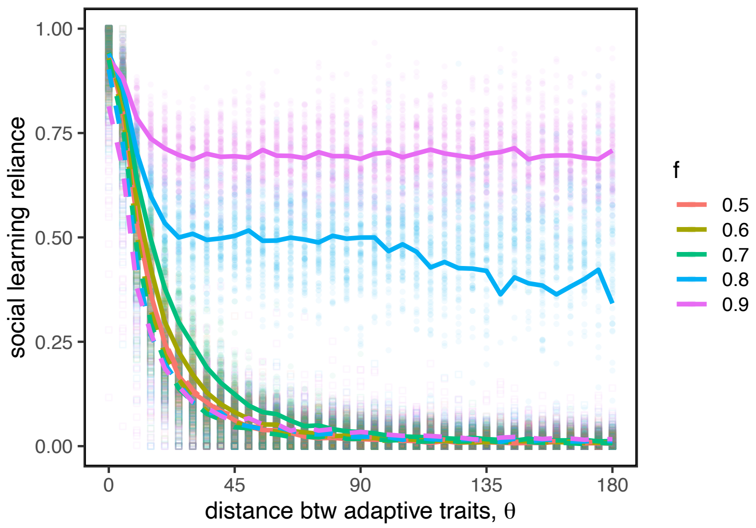

FIGURE 4

4.3. **Similarity-bias recovers the evolution of social learning.** Here we allow parochialism, a binary individual-level trait, to coevolve. We first find that as the distance between adaptive traits increases, so does parochialsm (Fig 5A), which in turn helps recover the evolution of reliance on social learning. We find that this is weakest in small minority groups. However, this effect may be driven by the fact that these simulations used $n = 1$, and so parochial social learners may simply be more likely to rely on individual learning only. I also found a bug in the code where individuals who threw away all social information used the trait value learned individually multiplied by $1 - s_i$, which probably selected against social learning reliance. I fixed this in the NetLogo code I sent, but anyway, this result may be an artifact (Fig 5B).

We also find that we get less selection for parochial social learning when R is small, because social markers are unreliable (Fig 5C,D). These results in the simulation were pretty noisy, so please may sure that it is coded correctly when translating to Julia (based on the logic of the model more than the NetLogo code).

4.4. **Our findings are robust to conformist and success-biased transmission.** It is important to test the robustness of all these results for conditions where success-biased and conformist social learning can evolve. This shows that this is not a phenomenon that relies on overly simple unbiased transmission, but also works for more sophisticated learning strategies. It appears to work (Fig 6).

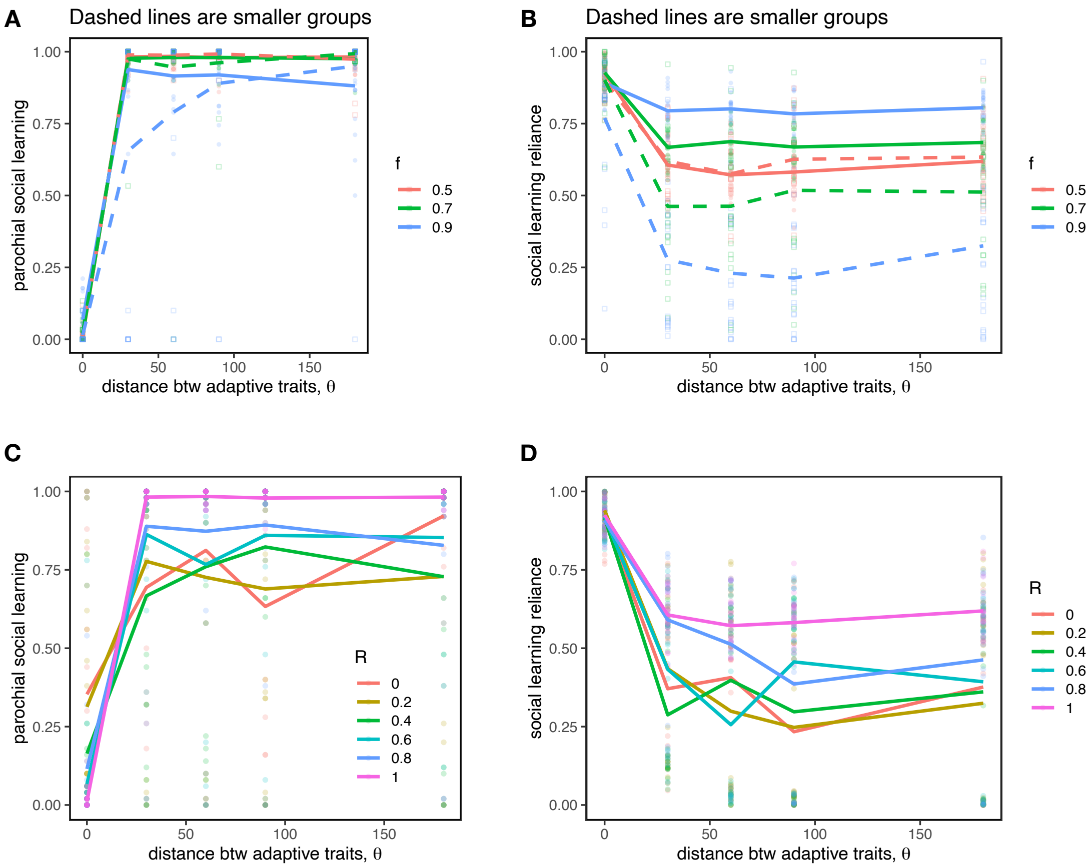

FIGURE 5. Caption

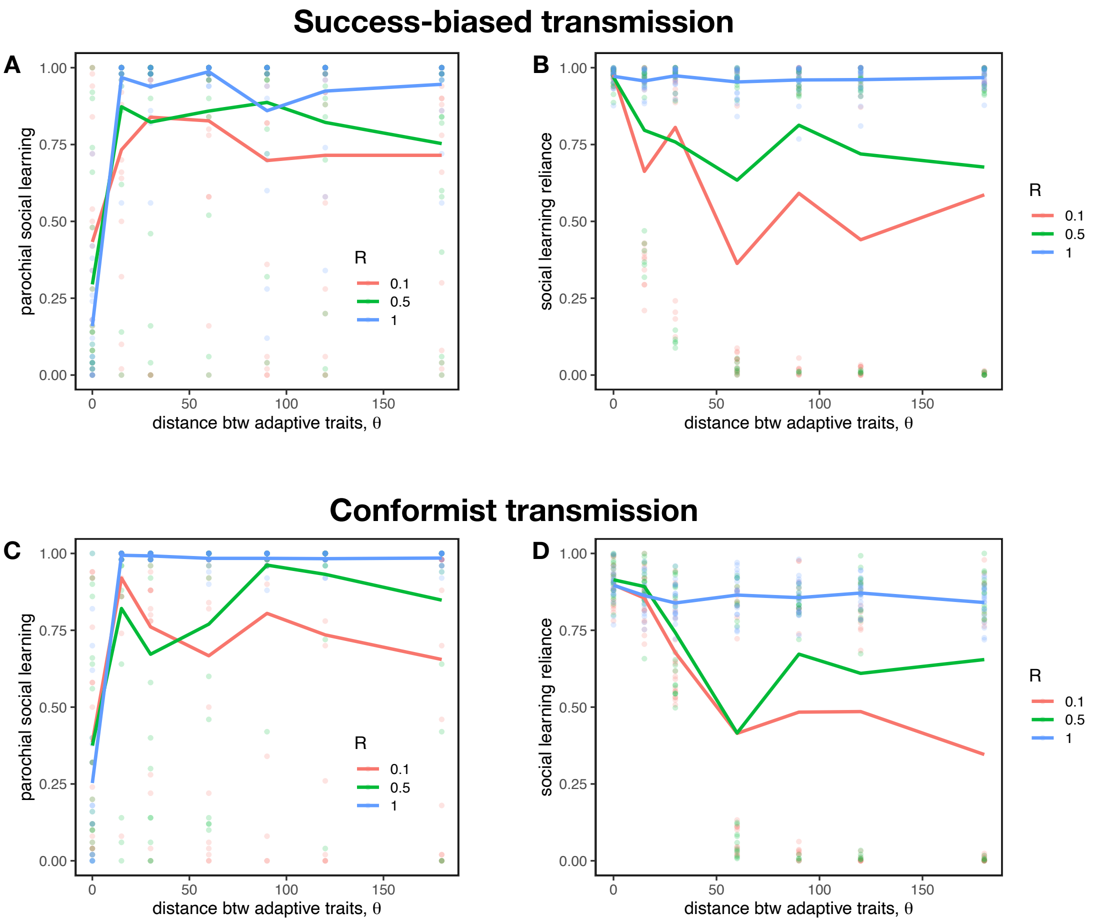

FIGURE 6. Caption