

Social Learning of Avoidance: Reinforcement learning and experimental results

Björn R. Lindström^{1,2}, Olof Hjorth¹, Andreas Olsson^{1,2}

¹ Department of Clinical Neuroscience, Karolinska Institutet, Sweden ² Stockholm Brain Institute

www.emotionlab.se

Introduction

- Learning by observing the behavior of others is wide-spread in nature (social learning) [1]
- Little is known about how humans learn to avoid potentially dangerous behaviors by observing others.
- We asked how avoidance learning is affected by observing the actions of, but not the direct consequences for, someone else:
 - I. How do we learn by observing avoidance without visible reinforcement?
 - II. How does individual trial and error learning alter or sustain avoidance learned by observation?

- We used Reinforcement Learning [2] to formalize our hypotheses about the learning mechanisms involved in social avoidance learning.
- The model-derived predictions were confirmed in two experiments.
- We extrapolated from the individual to the group-level using Agent-Based Modeling (ABM) [3].

Reinforcement Learning (RL) Model

We hypothesized that humans learn about the expected value (Q) of actions (A,B) through observation, even in the absence of any directly observable consequences of these actions:

$$Q_A(t+1) = Q_A(t) + \alpha_{OBS} * (R(t) - Q_A(t)) \quad [\text{Eq. 1}]$$

The observationally learned value will directly impact the probability of copying the observed action when confronted with the same alternatives (via the softmax function):

$$P_A(t) = \exp(Q_A/\beta) / (\exp(Q_A/\beta) + \exp(Q_B/\beta)) \quad [\text{Eq. 2}]$$

The expected value of the actions are modified by individual trial and error learning. We hypothesized that the absence of punishment function as (internal) reward (R = 1):

$$Q_A(t+1) = Q_A(t) + \alpha_{OWN} * (R(t) - Q_A(t)) \quad [\text{Eq. 3}]$$

The combination of social learning and internal reward predicts very stable preferences for the observed action.

Testing the model: Experimental Methods

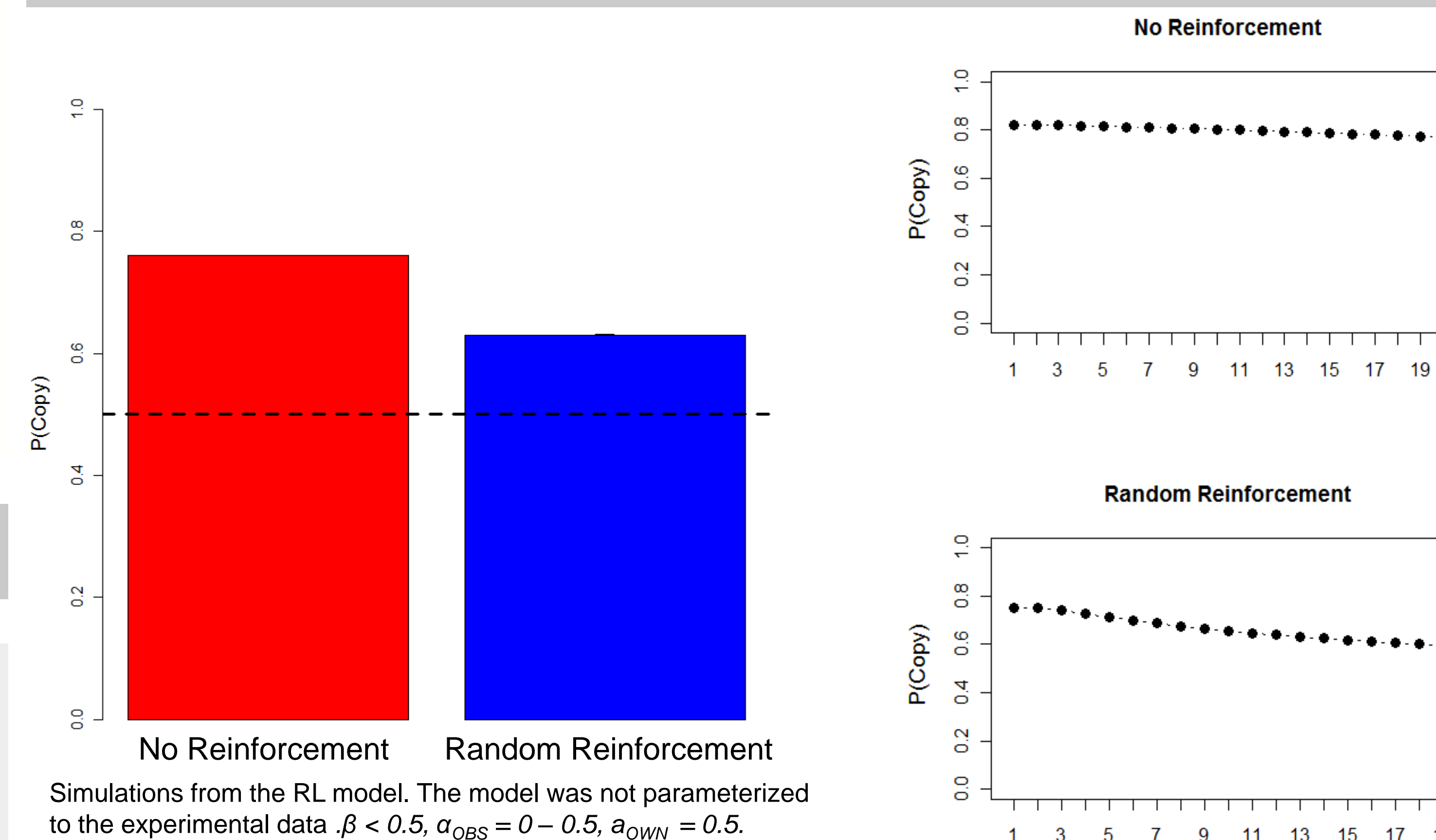
- Two-alternative (A,B) "shock avoidance" task
- Two phases: Observation (20 trials) and Own (20 trials).
- The choices of the computer-controlled "Demonstrator" during the Observation phase were always fixed to P(A) = 1 using a cover story.
- The subject observed the actions of, but not the consequences for, the Demonstrator.
- In the Own phase, the subject chose among the same choices (A,B)

Experiment 1 (n=27): No reinforcement

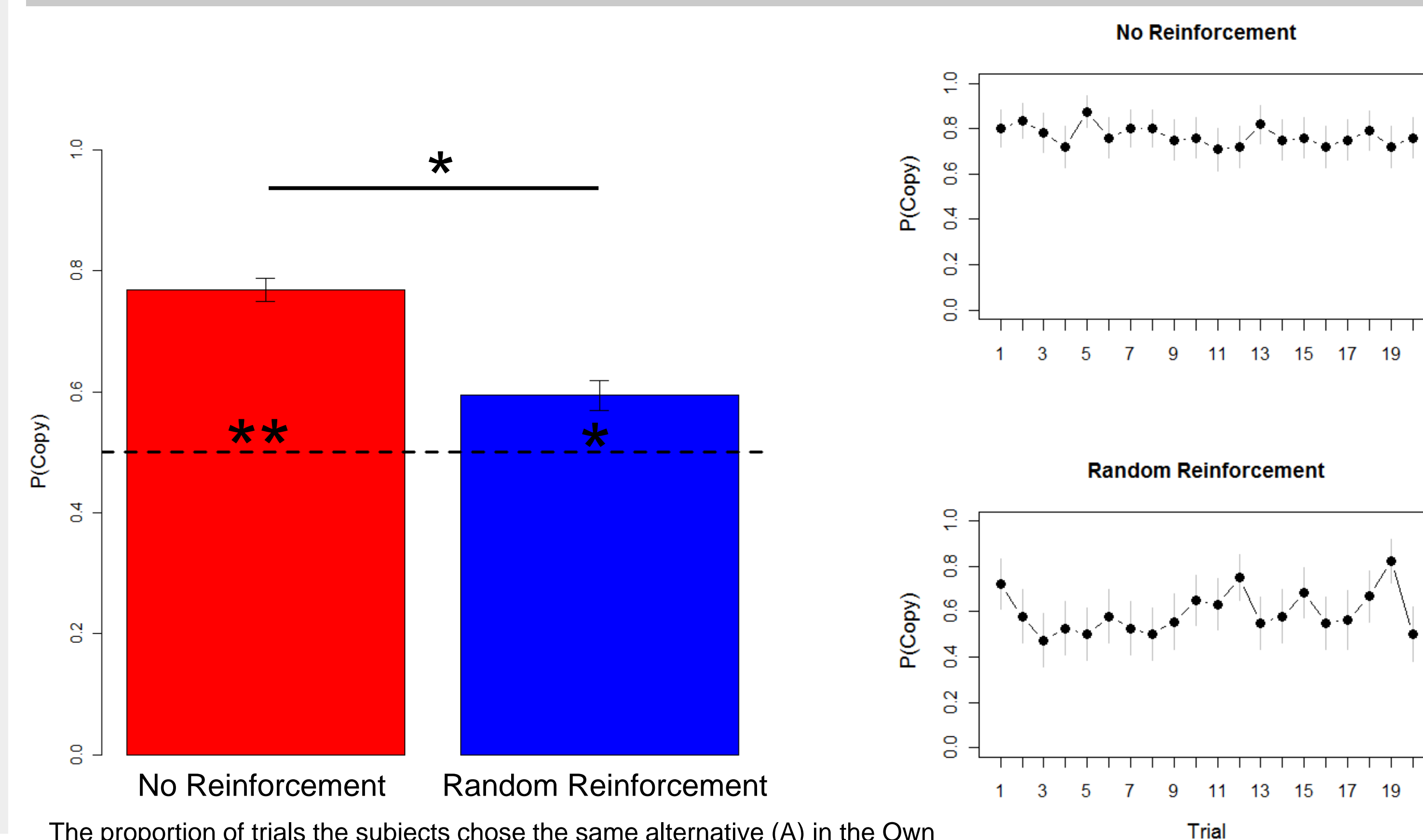
Experiment 2 (n=20): Random reinforcement (shock)

If the empirical probability of choosing (A) > 0.5 = social learning

Testing the model: Model predictions



Testing the model: Experimental results



Agent-Based Modeling (ABM)

Could the same learning mechanisms generate stable avoidance traditions on the group-level?

We simulated the predicted group-level behavior using ABM [3]. The behavior of each individual agent was based on the RL model.

Simulation procedure:

Initialize agents (e.g., n = 50) with random age

FOR n generations

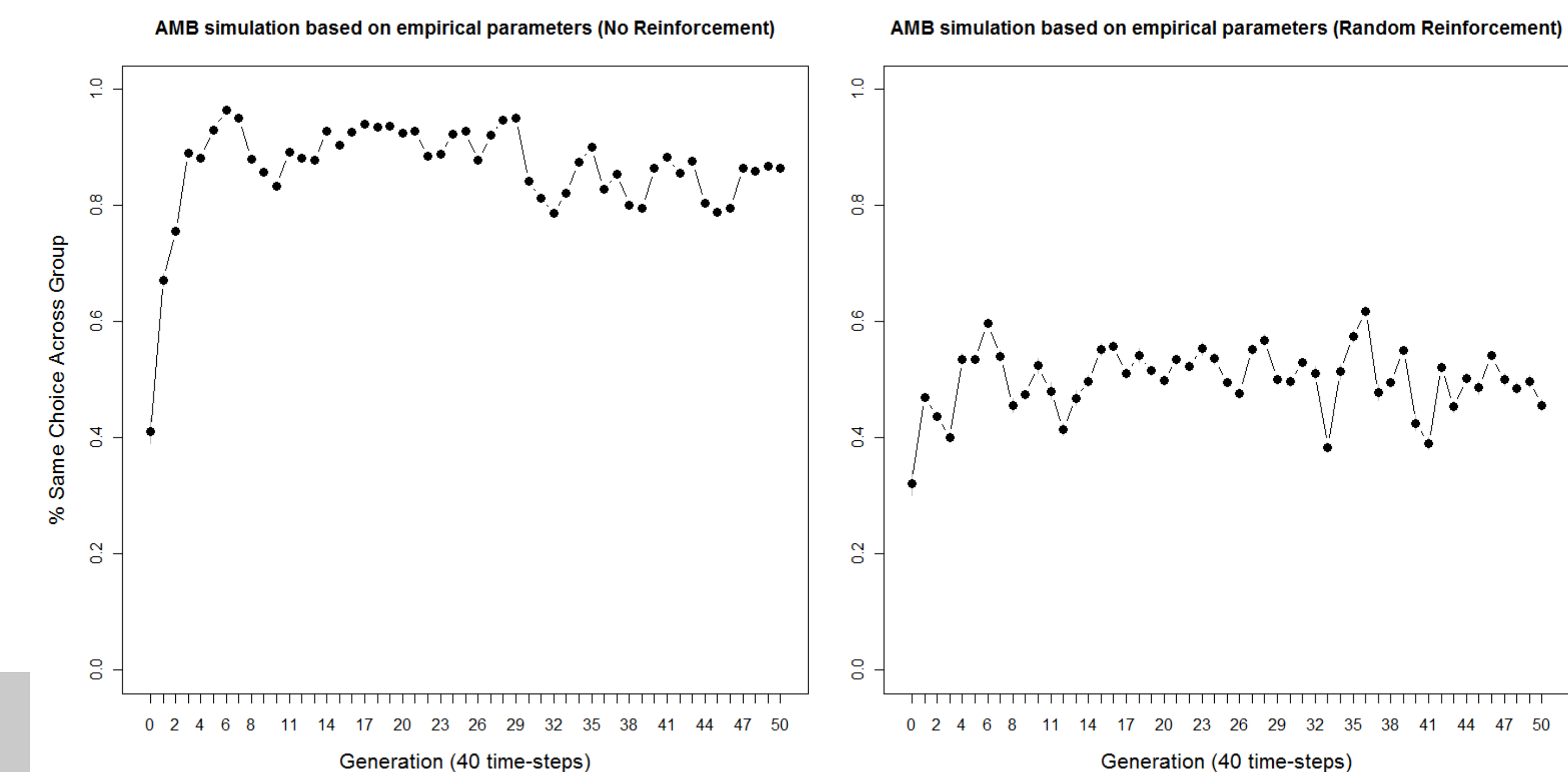
IF age <= 20 observe the actions of a random agent (Eq. 1)

ELSE individual decision making and learning (Eq. 2 - 3)

KILL oldest agent and replace with naive agent with age = 0

LOOP

Simulation of avoidance traditions



Conclusions

- Social learning has a powerful impact on avoidance behavior in the absence of overt reinforcement
- Social avoidance learning can be modeled by a simple RL model, which combine observational learning and internal reinforcement when expected punishment is absent.
- The model predicts stable avoidance traditions at the group-level.

References

1. Morgan, T.J.H., Rendell, L.E., Ehn, M., Hoppitt, W. & Laland, K.N. The evolutionary basis of human social learning. *Proceedings. Biological sciences / The Royal Society* **279**, 653-62 (2012).

2. Dayan, P. & Balleine, B.W. Reward, motivation, and reinforcement learning. *Neuron* **36**, 285-98 (2002).

3. Smith, E.R. & Conroy, F.R. Agent-based modeling: a new approach for theory building in social psychology. *Personality and social psychology review* **11**, 87-104 (2007).