Human biases limit algorithmic boosts of cultural evolution.

Levin Brinkmann*[1], Deniz Gezerli[1], Kira v. Kleist[1], Thomas F. Müller[1], Iyad Rahwan[1], and Niccolò Pescetelli[1]

[1] *Center for Humans and Machines, Max Planck Institute for Human Development*

14th April, 2021

1 Abstract

Humans are impressive social learners. Researchers of cultural evolution have studied the many biases that enable solutions and behaviours to spread socially from one human to the next, selecting from whom we copy and what we copy. In a digital society, algorithmic and human agents both contribute to transmission of knowledge. One hypothesis is that machines may influence the patterns of social transmission not only by providing a means for spreading human behavior but also by providing novel behaviors themselves. We propose that certain algorithms might show (either by learning or by design) different behaviors, biases and problem-solving abilities than their human counterparts. This may in turn foster better decisions in environments where diversity in problem-solving strategies is beneficial. In this study, we ask whether machines with complementary biases to humans could boost cultural evolution in a lab-based planning task, where humans show suboptimal biases. We conducted a large behavioral study and an agent-based simulation to test the performance of transmission chains with human and algorithmic players. In half of the chains, an algorithmic bot replaced a human participant. We show that the algorithm boosts the performance of immediately following participants in the chain, but this gain is lost for participants further down the transmission chain. Our findings suggest that algorithm can potentially improve performance, but human bias can hinder algorithmic solutions from being preserved. Our results suggest that the conditions for hybrid social learning and cultural evolution may be limited by task environment and human biases.

2 Introduction

When the first superhuman computer program in the game of Go—AlphaGo— beat the world champion Lee Sedol in 2016, its gameplay was considered surprising and unconventional, apparently violating longstanding Go traditions. In particular, for move 37, AlphaGo calculated the chance of a human player making the same move as 1 in 10000 [1]. Its unconventional play might originate from the fact that AlphaGo [2], and more so its successor AlphaZero [3], learned through self-play and, in the case of AlphaZero, do not rely on historic records of human gameplay. In the following years different teams created open-source reimplementations of AlphaZero, namely Leela Zero [4] and OpenGO [5]. These algorithms have in common, that they self-learned their gameplay independently from humans, and thus allow for an exciting new way to evaluate human cultural evolution of GO play. Replaying historic human games of the last 300 centuries, OpenGo, over the years, increasingly often chooses the same move then humans, suggesting a convergence towards an similar gameplay [6]. Remarkably, there has been a steep increase in this alignment since 2017 when Leela Zero became available to the public [6, 7]. Similar patterns of increased alignment between human and algorithmic play have been suggested in the game of chess [8].

The use of tools, such as books or software, for human training in games such as Go and chess is not a novel phenomenon and represent one way how knowledge is socially transmitted among people through technology. Yet, current development in AI made it possible not only for algorithms to play chess but to play creatively and without the need to rely on human games. This opened up the fascinating possibility of social learning, namely learning by observation [9], between algorithms and humans [10]. We hypothesized that social learning between human and algorithms may be especially beneficial when human and algorithms

*corresponding author: brinkmann@mpib-berlin.mpg.de

show diversity in problem-solving. Diversity of information and problem-solving strategies tend to reduce herding and error cascades [11–15]. By self-learning or by design, algorithms showing complementary biases to humans could foster the discovery of new solutions in domain-specific problems and improve outcomes compared to human-only problem solvers[16].

Researchers have already pointed out that the widespread use of digital technology can influence the processes of social learning in humans by providing new and faster means of communication and social learning [17, 18]. Going one step further, we argue that rather than a mere means for cultural evolution (such as books or Twitter), digital technologies and AI may also play an active role in shaping cultural evolution processes. Domain-specific algorithms like AlphaZero, either by learning or by design, may improve on cultural adaptation by increasing problem-solving diversity in domains where human bias is damaging performance. Algorithms learning from interaction with their environment, rather than from observing human data, may be in a better position to foster innovation, as in our AlphaZero example. Similarly, when human biases are known, algorithms can be designed to exhibit complementary biases to their human counterpart.

For example, when humans face decisions under cognitive constraints, they apply heuristics [19–22], which then under different circumstances can manifest themselves as a misadapted bias [23–25]. One instance of a human bias is a tendency for myopic behavior when facing a sequential decision [26–29]. Many problem-solving tasks are composed of sequential decisions, and a solution strategy is the exploration of the corresponding decision tree. With an increasing number of decisions, humans and algorithms do not explore the decision tree's total depth [2, 29–31]. For instance, Huys et al. [29] showed that people tend to selectively prune the decision tree after the anticipation of a large loss. Such heuristic can lead to sub-optimal solutions if a great effort or loss has to be accepted first for a large reward to be acquired.

In their experimental study Huys et al. [29] introduced a simple sequential, goal-directed decision-making task where participants plan and make sequential moves on a directed network (Figure 1). Each move is associated with gains or losses of different magnitude. Participants have full knowledge of the network and the rewards associated with each transition between two nodes. For this task, which we will call the reward network task, a policy based on a state-action value function that selectively discounts rewards following large losses, is best in describing the human decision-making [29, 32]. We call the corresponding bias the aversive pruning bias. As it can lead to sub-optimal behavior in the reward network task, Lieder et al. [33] investigated if an algorithm can augment human decisions by providing pseudo-rewards, which render a myopic strategy optimal. We address the same question from a different angle, namely whether humans can overcome their inherent bias via social learning from an algorithm and whether humans further transmit this new behavior. Social learning is especially important in complex problems and decisions taken under uncertainty [9, 34, 35]. Hybrid social learning may thus be beneficial in contexts where humans show biases that limit their potential for exploration and problem-solving.

This paper explores this possibility in the lab, using a transmission chain paradigm [36–39] where human and algorithmic players play the reward network task sequentially. Participants who are playing on the same network repeatedly tend to reuse similar actions between different games [32]. To exclude such non-social learning, we developed a novel randomized version of the reward network task, in which each participant plays the same network only once. Additionally, we developed a classification strategy, which allowed us to compare environments for which the human aversive pruning bias is either adapted or misadapted. We ask whether humans can learn from the observation of an algorithm whose play shows complementary problem-solving biases. We designed the algorithmic player to show a bias opposite to humans, namely to explore decision branches associated with initial losses. In the treatment condition, an algorithm replaced a human in the second generation. We predicted better performance in the treatment condition than in a control condition of humans-only due to the increased strategic diversity of the hybrid chain.

Transmission chain experiments have been used to investigate how biases shape cultural evolution [40–43]. Previous research has shown that humans have different biases of what (content bias) and who (context bias) is copied [9, 44, 45]. A bias based on characteristics of who is copied has important implications on hybrid social learning because humans tend to quickly lose trust in algorithms [46, 47]. In this study, however, we are interested in how algorithmic and human biases interact and therefore we control for a context bias by hiding whether an artefact was created by a human or the algorithm. Prior to the experiment, we expected the performance of a solution (fitness) to be the decisive feature driving social learning, yet we learned that the bias we intended to compensate with the algorithm, the aversion of large losses, plays an equally important role.

In line with our preregistered hypothesis, we found an increase in performance over generations and a short term performance improvement due to algorithmic solutions. However, in contrast to our hypotheses, the positive impact introduced by the algorithm was not sustained over time. Solutions that conflict with human aversive pruning bias, had higher copy error rates and therefore quickly disappeared in the long run. We discuss our results in terms of content bias, which defines the limits under which behaviours are socially transmitted and stable across generations.

We contribute to the literature on social learning and cultural evolution by hypothesising hybrid human-algorithm social learning, while previous studies focused on human behaviour [36, 48]. We also show that although theoretically possible, under conditions of time-pressure, working memory load or uncertainty, the benefits introduced by diverse algorithms may be limited. Ultimately, the conditions for successful hybrid cultural evolution may be limited by task characteristics and human biases.

3 Methods

3.1 Participants

The study was approved by the ethics committee of the Max Planck Institute for Human Development. All 177 Participants were recruited through Prolific (www.prolific.co), where they were forwarded to an external website to complete the experiment. Before starting, they received a consent form and instructions. The experiment, including two practice rounds, took around 60 minutes in total. Participants were paid £7 for the completion of the experiment. Furthermore there was a reward of one penny given for every 100 points gained during the experiment. Participants received on average £3.20 bonus payments, depending on their performance. In some cases, where participants had to drop out because of technical issues on their sides (failed network connection, etc.), they were paid £3.50 for their participation. The experiment was run in multiple sessions. Two sessions failed for technical reasons. The data from those sessions was disregarded entirely and the experiment restarted with a database snapshot created before those sessions. The only entry requirement was speaking English. We did not exclude any participants from the analyses.

3.2 Task

Huys et al. [29] developed the reward network task in which participants were ask to find an optimal sequence of moves on a carefully designed directed network of 6 nodes. We generalized the task, by randomly sampling the networks, instead of using a single fixed network. From each node there were exactly two moves to other nodes possible, each being associated with one of four possible payoffs (-100, -20, 20, or 140) (see Figure 1.a). Possible moves were visualized by directed arrows with colors matching their respective payoffs (red, orange, blue, and green). The objective was to find a path of 8 moves beginning at a fixed starting node, which maximizes the accumulated payoff. Unlike in [29], networks were randomly generated (see below). We called a network together with a specific starting position an environment. The experiment was implemented using a customized version of the Empirica framework [49] and consisted of 3 consecutive stages (see Figure 1.a). In the first stage, participants were asked to watch the solution of a previous player. They saw the score of the previous player and a 15 seconds animation of the 8 moves. Each move between two nodes was animated individually for 2 seconds with both nodes being highlighted by a darker color and the corresponding directed arrow and reward being thickened (see Supplementary Figure 1 for a screenshot). In the second stage, participants were shown the same environment and were asked to select a path of 8 moves. The path could be entered by clicking on the nodes involved in sequence. The currently occupied node was displayed in darker color. If a node was selected which could not directly be reached from the current node, the erroneously selected node was colored in bright red. The participant was then able to select a different node instead. Participants were able to see their current accumulated score of this round, the number of steps remaining and a score of the last moves entered. This information got updated immediately, whenever a participant clicked a possible target node (see Supplementary Figure 2 for a screenshot). The answer of the participant was considered to be valid if all 8 moves were played in the given time. Participants received a punishment of -500 for the round if they did not provided a valid answer within time, to strongly incentives participants to responds even if being uncertain about the solution. In the third and final stage, the final score of the current round was displayed in large letters (see Supplementary Figure 3 for a screenshot). Participants were also informed if they received a punishment for failing to provide a valid solution. The first and the third stage had a fixed duration of 15 and 5 seconds, respectively. The second stage had a maximum duration of 15 seconds, which ended once a valid solution was entered. Participants were shown their remaining time of each stage at all times together with their total score across all previous rounds.

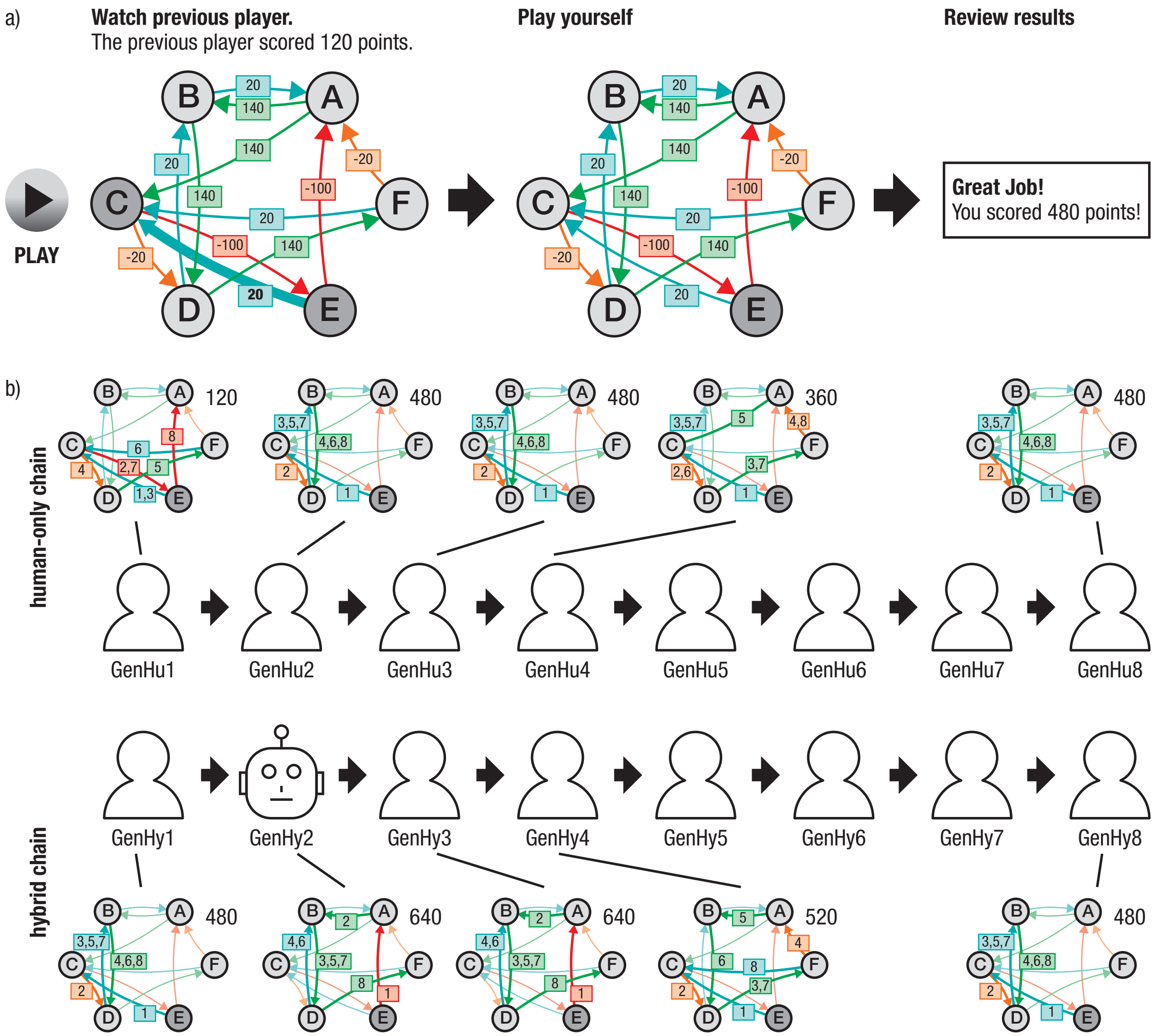

Figure 1: (a) In the first stage of the network task participants were seeing a animation of the solution of the previous player. Here we are depicting a snapshot showing the transition from node E to node C. In the second stage the participants could enter a path by clicking on the respective nodes in sequence. The node with gray background color indicates the current node, in this case the starting node. The network presented here is classified as human regretful. In the last stage the total score of the sequence is presented. (b) For each environment we constructed two chains of 8 player. In hybrid chains, the second player was replaced by an algorithm. The networks depict the solutions of the first 4 generations as well as the last generation for a handpicked environment (corresponding to (a)). The value on the arrows denotes the step at which a player was choosing the move. The cumulative reward is shown in the upper right corner of each graphics. In this example for the human-only chain the cumulative reward increases at first, but stagnates quickly. For the hybrid chain, the algorithm shows a performance beyond that observed in the human-only chain, but this increase gets lost over the subsequent human generations.

3.2.1 Experimental Design

We created chains of 8 different players. The positions within the chain were called generations. Players could be human participants or an algorithm. Within each chain, each player was exposed to the solution of the previous player. Players in the first generation were exposed to a random solution.

In a 2x2 design, chains of two different types were created (see Figure 1.b). In human-only chains, the control condition, all 8 generations comprised human participants. In hybrid chains, the treatment condition, an algorithm (described below) replaced a human and provided the solution of the 2nd generation. The rest of the chain comprised human participants. 800 environments of two different types were investigated. As described below in detail, the two types, 'human rewarding' and 'human regretful', differ in whether aversive pruning is increasing or reducing the expected reward, respectively. For each of the 800 environments two chains where constructed, one for each of the two conditions. This lead to a total of 1600 chains and 12800 games, of which 800 are played by the algorithm.

Participants were assigned to new environments on the fly at random based on availability, with the constraints that the previous generation in the chain has been successfully finished and that the participant plays each environment at most once. If a participants did not entered a path of 8 moves in time, the solutions was considered as invalid and the corresponding position in the chain was reopened for a new participant. Each participant had to play a maximum of 80 rounds. Towards the end of the experiment, participants left the experiment earlier, when no further matching game was available. Due to the random assignment procedure, participants were likely playing in each of the chain types as well as the environment types throughout the experiment. However, participants entering the experiment at the beginning were more likely to be placed in earlier generations, compared to participants who entered the game at a later stage.

3.2.2 Aversive Pruning Model

Huys et al. [29] described a pruned tree search algorithm for this type of task. The model calculates the state-action value $\mathcal{Q}(a, s)$ of each action (move) a in state s. The value of a particular action is given by the sum of the immediate reward $\mathcal{R}(a, s)$ and the maximum value of the next action a' from the next state s' $= \mathcal{T}(a,s)$ where $\mathcal{T}$ is the deterministic transition function. At each level of depth of the search tree, future rewards are discounted by a factor of $(1 - \gamma_{a,s})$. Together this leads to the well known Bellman equation

$$\mathcal{Q}_d(a, s) = \mathcal{R}(a, s) + (1 - \gamma_{a,s}) \max_{a'} \mathcal{Q}_d(a', \mathcal{T}(a, s)). \tag{1}$$

The parameter $\gamma_{a,s}$ is interpreted as the rate of pruning of the search tree in a mean field approximation. Correspondingly, rewards k steps ahead are discounted by a factor of $(1-\gamma_{a,s})^{(1-k)}$. Scaling the state-action value $\mathcal{Q}_d$ by the inverse temperature β and applying a softmax leads to the policy

$$\pi(a_t|s_t) = \frac{e^{\beta \mathcal{Q}_d(a_t,s_t)}}{\sum_{a'} e^{\beta \mathcal{Q}_d(a',s_t)}}. \tag{2}$$

Central to their work, Huys et al. [29] defined a 'Pruning' version of this model to account for stronger pruning when the participants encounter a large negative reinforcement. A large negative reinforcement is defined as a reward of -100 in this experiment. In this model, which we will call the aversive pruning model, the $\gamma_{a,s}$ parameter takes two different values, a specific pruning rate γ_s in the case of large negative losses and general pruning rate γ_g in all other cases [3].

$$\gamma_{a,s} = \begin{cases} \gamma_s, & \text{if } \mathcal{R}(a, s) = -100 \\ \gamma_g, & \text{else} \end{cases} \tag{3}$$

3.3 Network Generation, Selection and Classification

800 environments were created before the experiment, each one characterised by a directed network of nodes and a starting node, with each edge of the network defining a possible move between two nodes. We started with creating a pool of 60000 strongly connected directed networks and sampled uniformly for each link between two nodes one of four possible reward (-100, -20, 20, 140). Considering 6 possible starting nodes for each network, this yielded 360000 environments. We then calculated for each environment a path maximising the reward. To gain a more compact reward distribution, environments with a maximum reward in the upper and lower quartile were removed from the pool. To avoid trivial solutions, environments were rejected if the maximum path did not cover at least 4 distinct nodes. Finally, to exclude environments with myopic optimal

solutions, we compared for each node on the optimal path, the reward of the optimal move with the reward of the alternative, sub-optimal move. We required environments to have at least four moves, in which the optimal one has the same or a lower direct reward then the sub-optimal one.

The final selection of environments was based on the sensitivity of aversive pruning on the expected total reward. The aversive pruning sensitivity for each environments was examined by choosing a reference model ([2], $\gamma_g = \gamma_s = 0.35$ and $\beta = 0.03$) and calculating the derivative of the expected reward in respect of the aversive pruning parameter γ_s. We then randomly selected 400 environments each with their aversive pruning sensitivity in the lowest and the highest decentile. Environments for which the aversive pruning bias led to lower rewards were named 'human regretful'. The other type of environments were correspondingly named 'human rewarding'.

3.4 Matching the algorithmic performance

We aimed for the algorithm to have a performance comparable with a human participant, to similarly discount future rewards and yet to have a different bias. On a pilot study we estimated with a Bayesian model fit (see supplementary material) the model parameter of human participants as $\gamma_g = 0.20$ (CI_{90}: $(0.15, 0.25)$), $\gamma_s = 0.45$ (CI_{90}: $(0.32, 0.58)$) and $\beta = 0.012$ (CI_{90}: $(0.011, 0.014)$). Note that γ_g and γ_s are of comparable magnitude of what was found by Huys et al. [29], however in our pilot we observed a lower inverse temperature β our policy. In their work participants could learn the environment as they were trained repeatedly on one specific network, while in our work participants where playing different environments and each only once.

To obtain a risk seeking algorithm with a bias inverse to humans but with comparable performance we fixed $\gamma_g = 0.5$ and $\gamma_s = 0.05$ and then fitted $\beta = 0.0264$ to match the performance of human participants on the pilot study. To mimic social learning the algorithm is using an additional heuristic at run time. First, a solution, i.e. a sequence of 8 actions, is sampled using the aversive pruning model with the parameter describe above, then the total reward of this solution is compared with the one of the previous player. If the solutions of the algorithm matches or surpasses the previous reward, it is enter into the chain. Otherwise an exact copy of the solution of the previous player is entered.

3.5 Statistical Analysis

We run a hypothesis driven linear regression with the reward of a solution as the response variable. The reward of a solution is the sum of the rewards of each of the 8 moves of a single round. Additional, we run explorative Poisson regressions with a logarithmic link on the number of actions copied between solutions as the response. Transformations of factors are described alongside the results. Different models were compared by performing a pairwise ANOVA and comparing the AIC. We used a single model for both types of environments and consequently added interactions between each fixed effect of interest and the environment type. As our main focus within this work is on environments challenging for humans, we used the human regretful environments as the baseline to ease interpretation. 95% confidence intervals are reported throughout. The code of the statistical analysis and the corresponding data is published with this work.

Prior to the experiment, we expected as preregistered that (H1) in human-only chains, individual solutions will improve across generations, within each environment, via social learning. We expected that placing the machine in the chain at generation two ($GenHy2$), will (H2) locally increase performance so that a score boost is observed in generation three ($GenHy3$) and following compared to the first generation ($GenHy1$). We expected our manipulation to (H3) globally increase performance as measured by normalized score accrued in the game, to (H4) accelerate solution discovery as measured by the slope of score improvement and reduction of error compared to the global optimal solution, and to (H5) increase the likelihood of chains discovering the best solution. We furthermore expected (H6) to not have an effect of the machine intervention on human rewarding networks as people will be able to judge that their own solution is better than the machine's.

4 Results

4.1 Algorithm impacts following generations, but effect quickly decays

To investigate and compare the evolution of the performance of solutions in the different chains, we run a linear mixed-effects model predicting the reward of a individual solution, by considering (a) the numeric position in the chain (generation), (b) individual generations following the algorithm and (c) the number of rounds participants had previously played (max 80) as additive effects. For the first two effects (a, b) we added an interaction with the environment type. Additionally we added random effects for the (d) individual participants and (e) individual environments. The round of a participant (c) was added to account for non-social learning of participants. We considered the first generation of the human regretful environments as the baseline. Algorithmic solutions were not considered in this analysis because they were part of our treatment.

We encoded (b) the influence of the algorithm on the performance of following generations by adding two independent effects for the two generations directly following the algorithm ($GenHy3$ and $GenHy4$). All further generations were assigned a single effect ($GenHy5+$) and we considered solutions not following an algorithm in the chain as the baseline. This includes all solutions in human-only chains as well as the first generation ($GenHy1$) in the hybrid chains, where the algorithm has not yet been introduced. This most parsimonious model (Supplementary Table 2) had a smaller AIC ($/chi^2 = 1.99$, $df = 2$, $p = 0.37$ and $/chi^2 = 5.13$, $df = 6$, $p = 0.52$) compared to one that included independent effects for all generations following the algorithm (Supplementary Table 1) and one that included independent effects on the three generations ($GenHy3$, $GenHy4$ and $GenHy5$) following the algorithm (Supplementary Table 3).

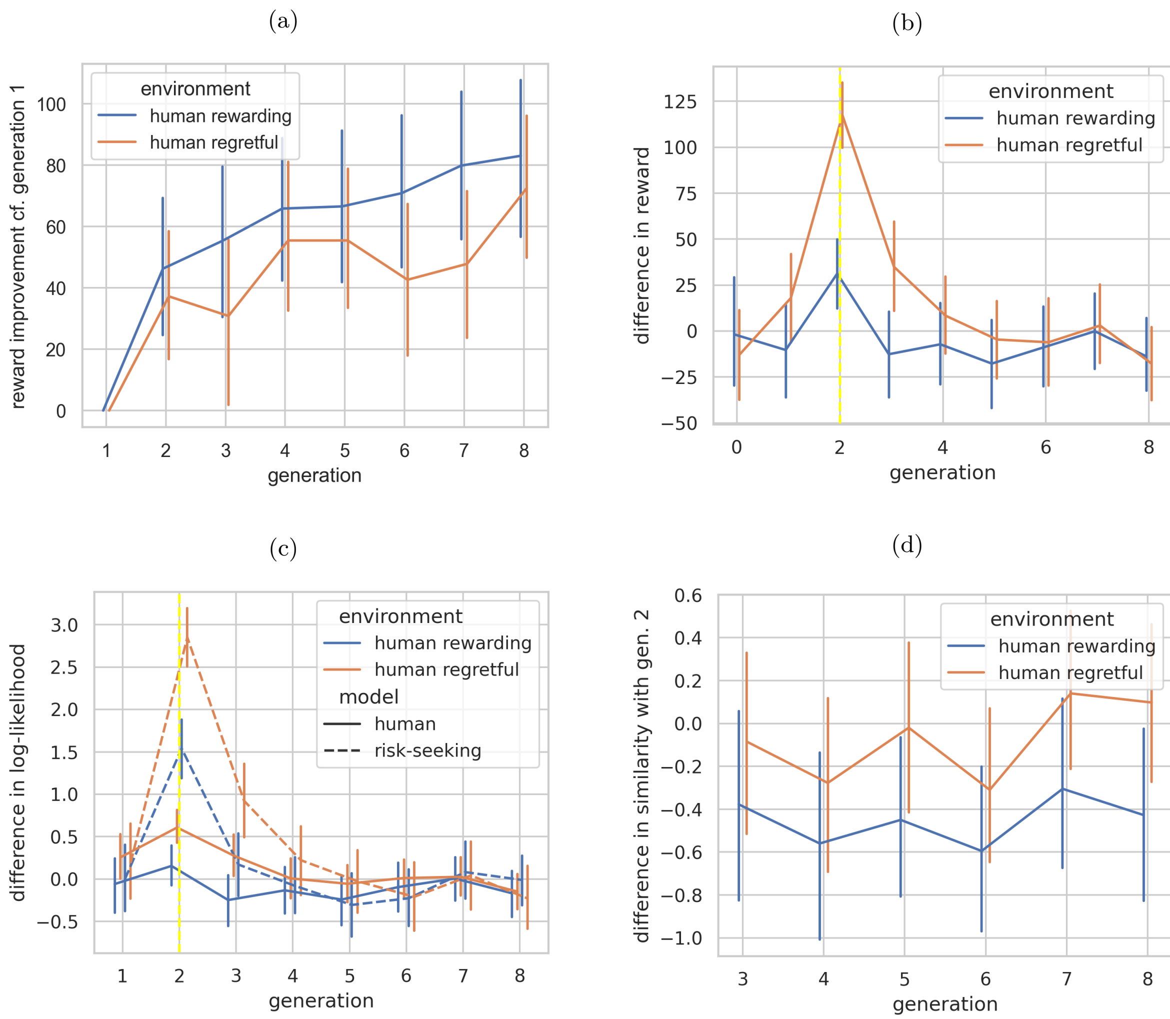

Figure 2: (a) Average within chain improvement of performance over generations; (b) average performance difference between human-only and hybrid chains; (c) average difference (hybrid − human-only) in the log-likelihood of two models (solid: model fit to human data, dashed: algorithm used in hybrid chains) to choose the same move then the player; (d) Average number of matching actions with generation 2 over generations in hybrid chains. Vertical bars are indicating bootstrapped 95 % confidence intervals of the mean. (d) Difference of average count of matching actions with generation 2 over generations between treatment and control groups. Vertical bars are indicating bootstrapped 95 % confidence intervals of the mean. Dashed vertical line at generation 2 shows the algorithm's position.

As a first validation of our experimental setup, we quantified the effect of social learning by investigating the impact of generation (a) on reward. We found for human regretful environments an improvement of

3.867 ($\sigma = 1.244$, $Z = 3.109$, p $= 0.002$, $CI = (1.429, 6.305)$) points from generation to generation, and for the human rewarding environment an additional improvement of 4.859 ($\sigma = 1.390$, $Z = 3.495$, p < 0.001, $CI = (2.134, 7.583)$) points per generation. Figure 2a depicts the average reward of solutions in human-only chains in relation to the reward of the first player in the chain. The increase in performance over the 8 generation suggests the presence of social learning as predicted (H1) and the accumulation of higher performing solutions in later generations. Social learning appears to lead to larger increases in performance for 'human rewarding' environments where the human bias is beneficial.

Having found, that social learning does occur, we investigated the impact of the algorithm on following generations (b) in mixed chains. Figure 2b depicts the average reward difference comparing hybrid chains with human-only chains. We found for human regretful environments an effect of 30.786 ($\sigma = 7.974$, $Z = 3.861$, p < 0.001, $CI = (15.157, 46.415)$) points for participants directly following the algorithm and a weak effect of 13.225 ($\sigma = 7.922$, $Z = 1.669$, p $= 0.095$, $CI = (-2.302, 28.753)$) points for the second generation following. No effect was found the remaining generations -2.473 ($\sigma = 5.094$, $Z = -0.485$, p $= 0.627$, $CI = (-12.457, 7.511)$). We did not find evidence for interactions of these effects with the environment type (see Supplementary Table 2). Human participants in the generation following the algorithm (generation 3) have higher rewards than their counterparts in human-only chains. However, this effect appears to quickly wear off. These findings support a boost in performance of participants by social learning from the algorithm (H2). We could not find evidence for global increase in performance for hybrid chains (H3) when considering in the second half of the chain (generation 5-8).

Given that the algorithm had an significant, but short lived, effect on following humans performance, we then investigated if participants behavior was likewise effected. As a measure of behavior we utilized two models, one resembling the algorithm as used in the experiment and a second one fitted on human data from a pilot study (see 3.4), and calculated the likelihood of the models to select the same move then the player in the experiment. We obtained a single score for each solution by summing the log-likelihoods of all 8 moves. We then calculated the average log-likelihood for each condition and generation as a measure of behavioral similarity between the player and the model. Figure 2c depicts the difference in the likelihood between hybrid and human-only chains for the different generations (see also Supplementary Figure 5 showing both conditions separately).

For the human regretful environments we found in the third generation, directly following the algorithm, a significant difference in the log-likelihood for the risk-seeking model (...) comparing hybrid and human-only chains, but no difference for the human model (...). In the last generation, we did not found an evidence for a difference in the log-likelihood for the risk seeking model between both conditions (...). For the human regretful model we did not find evidence for a relative increase in the log-likelihood of the risk-seeking model in the third generation in hybrid chains (...). These findings suggests that solutions transmitted from the algorithm to humans contain characteristics of the algorithm.

Finally, we investigated the rate at which participants in human-only and mixed chains do follow optimal strategies (see Supplementary Figure 4). We found for hybrid chains an increased rate at which optimal solutions are discovered in generation 3 (...) compared to human-only chain suggesting social learning from the algorithm, however the difference decays quickly and we do not found a significant difference in the solution discovery in the final generation (...). Correspondingly these findings do not support the hypotheses of an accelerated solution discovery (H4) and sustained increase in discovery rate (H5) induced by the algorithm.

4.2 Of equally scoring solutions, the algorithmic ones are copied less.

Figure 2b depicts the average number of matching moves between solutions of the second generation (either from a human or the algorithm) and the following (human) solutions. Despite the higher performance of algorithmic solutions, those do not appear to be copied at a higher rate. We suspected that there are two effects at play with opposite sign. On the one hand the higher reward of algorithm solutions, could lead to a higher rate of imitation. On the other hand the mismatch with the inherent bias of participants, might reduce copying.

To explore further the potential drivers of imitation we created a set of exploratory Poisson regressions predicting the number of actions being copied. As fixed effects we considered (a) the creator of the previous solution (algorithm or human), (b) the reward of the previous solution and (c) the number of large losses in the previous solution. We scaled the reward (b) by its variance. We added random effects for the participants and the environment. A full model (Supplementary Table 4) showed very weak interactions in general, with the exception of an interaction with the environment type. For this reason we added interactions with the environment type to each factor and did not considered any additional interaction in the smaller models.

In a first analysis we focused on the solutions of participants in generation 3 as there the effect the algorithm should be the strongest. A model (Supplementary Table 4) only including whether the previous solution was from a human or the algorithm (a) as dependent variable did not showed a significant effect ($\beta(\sigma) = -0.011(0.038)$, $Z = -0.280$, $p = 0.780$, $CI = (-0.085, 0.064)$). However, when including the reward of the previous solution (Supplementary Table 5) we found an increased rate of copying of solutions with higher scores (($\beta(\sigma) = 0.394(0.034)$, $Z = 11.423$, $p < 0.001$, $CI = (0.326, 0.461)$)) and a lower rate of copying of algorithmic solutions ($\beta(\sigma) = -0.189(0.041)$, $Z = -4.625$, $p < 0.001$, $CI = (-0.269, -0.109)$). This findings suggests that once controlling for the higher reward of algorithmic solutions, that those solutions are copied at a lower rate compared to human solutions. A model (Supplementary Table 6) including an interaction between the reward (b) and whether the previous solution was from a human or the algorithm (a), did not found this interaction to be significant (($\beta(\sigma) = -0.027(0.059)$, $Z = -0.468$, $p = 0.640$, $CI = (-0.143, 0.088)$)), suggesting that the two effects were additive. Also, in none of the presented models interactions suggest a significant difference between the two types of environments.

Participants did not know whether the previous solution they see is from an algorithm or from another human. We thus investigated which features mediated the apparent difference in social transmissions. The large number of paths collected in the control condition gave us the opportunity to independently test the hypotheses that both higher rewards and fewer number of large losses lead to a higher chance of a solution being copied. In this second analysis we used the data from generation 2 to 8 in human-only chains and run another Poission regression (Supplementary Table 7) on the number of actions being copied, in human-only chains. We excluded the first generation as participants there are exposed to a random solutions of in general very poor performance. We found a positive effect on the reward of the previous solution (b) ($\beta(\sigma) = 0.393(0.014)$, $Z = 28.162$, $p < 0.001$, $CI = (0.366, 0.420)$) and a negative effect on the number of large losses within the previous solution (c) ($\beta(\sigma) = -0.040(0.011)$, $Z = -3.544$, $p < 0.001$, $CI = (-0.061, -0.018)$).

These findings indeed suggest a content bias in the social transmission in which solutions with higher reward and less large loss are transmitted with increased fidelity. Consequently solutions alien to humans, such as those from our algorithm, are less well preserved.

4.3 Agent-based model shows that only repeated algorithmic participation has a sustained effect.

We developed a simple agent based model mimicking our experimental setup to theoretically explore the impact of a specific biases on social learning in hybrid cultural evolution (see supplementary methods for details). We stripped out most of the details of the experiment and modeled solution as a point in a two dimensional space with two independent qualities. The dimension s^g represents the general quality of a solution and the second dimension s^s the specialization of a solution, i.e. the adaptation to an environment. This specialization can be either adaptive or misadaptive. Correspondingly we model a 'human-rewarding' and a 'human-regretful' environment in which higher values of s^s lead to higher and lower rewards, respectively. Higher values of s^g lead to higher rewards in both environments. We model human agents as being adapted to a fix positive value of $s^s = 0.5$ and algorithmic agents to be adapted to fixed negative value of $s^s = -0.5$.

As in the experiment we construct chains of 8 agents. Agents first accessed a perceived quality of the previous players solution, then depending on this perceived quality decided whether to copy it, and finally either perfectly copied or sampled an entirely new solution. The perceived quality determines the content bias of the agent. We compare agents with adapted content-biases who consider both the score of the previous solutions and the match with their own specialization (dashed in Fig. 3) to agents with a utilitarian content-bias who consider only the score of the previous solution (solid in Fig. 3).

Both humans and algorithm agents sample new solutions from a 2-dimensional Gauss distribution which is shifted to higher values of s^s for humans and to lower values of s^s for the algorithm agents, and thereby mimicking the inverse biases of those two agent types. Chains are initialised with a sampled solution with neutral bias $s^s = s^g = 0$.

Figure 3 depicts the averaged reward over 8 generations. We depict chains of human agents, algorithmic agents and two hybrid chains. The 'single-algorithm' condition corresponds to the hybrid chains in the experiments and only has a single algorithm in the second generation of a chain of otherwise human agents. In 'random-hybrid' chains we randomly mix human and algorithmic agents.

In the 'single-algorithm' condition the algorithm shows strong super-human performance in the the second generation in 'human-regretful' environments. Agents with an adapted content-bias (left, green, solid) following the algorithm show a higher performance then their peers in the human-only condition,

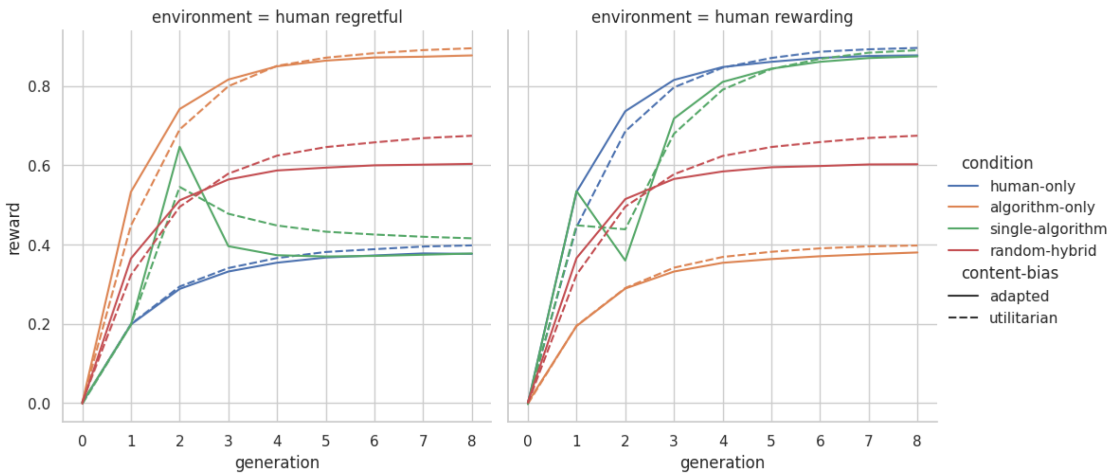

Figure 3: Average reward of the solutions of 100000 modeled agents. Human-only chains are depicted in blue, algorithm-only in orange, hybrid chains with a single algorithm (as in the experiment) in green and randomly mixed hybrid chains in red. On the left panel the environment favors the algorithmic bias, on the right panel it favors the human bias. We compare two type of content bias, one with a bias for higher performing solutions (solid) and a second with an additional bias to match the specific bias of the agent.

however, as in the experiment, the performance boost is quickly lost and chains with (left, green, solid) and without (left, blue, solid) the algorithm converge to the same performance level. The convergence is much quicker for agents with an adapted content-bias (left, green/blue, solid) compared to agents with an utilitarian content-bias (left, green/blue, dashed). On ‘human-rewarding‘ environments the introduction of the algorithm (right, green) lead to a performance drop in comparison to ‘human-only‘ chains (right, blue), however this performance drop is then made up for in the subsequent generations.

Having shown that within our model, a single algorithm has no sustained effect on the performance, we investigate the effect of randomly mixing humans and the algorithm. Humans and the algorithm are modeled symmetrical, and hence ‘human-only‘ chains on ‘human rewarding‘ environments correspond in their performance with a ‘algorithmic-only‘ on ‘human regretful‘ environments and visa versa. Randomly mixed hybrid chains (red) do show a performance in-between the performance of solely-adapted, e.g. algorithmic agents in human regretful environments, and solely-misadapted, e.g. human agents in human regretful environments. For those ‘random-hybrid‘ chains agents with an ‘utilitarian‘ content-bias (red, dashed) do eventually show a higher average performance compared to agents with an ‘adapted‘ content-bias (red, solid), however in the first two generation agents with an ‘adapted‘ content-bias have a slight edge over their peers. The same can be observed for chains with solely-adapted agents.

In the simplified setting explored by the agent-based model the performance of chains in the same condition converge to a fix value irrespective of the algorithm participation at the beginning of the chain. For the algorithms better adapted to the environment then humans, this lead to a decrease of performance in the generations following the algorithm. This loss in performance is sped up when human agents preferentially imitate solutions which matches their own biases and hence mismatch the bias of the algorithmic solutions.

5 Discussion

In this work, we investigated the long-term impact of algorithmic strategies on social learning using a transmission chain experiment. We adapted the reward network task introduced by Huys et al. [29] to a transmission chain paradigm to test for improvement in performance via social learning over generations of the chains. In this task, people are known to show an aversive pruning bias in exploring the decision tree. As expected, we found evidence of an improvement of performance over generations due to social learning. Contrary to our expectations, adding an algorithm with a problem-solving bias complementary to humans into the chain did not improve chain performance. While humans did copy solutions from the algorithm, they appeared to do so at a lower rate than they copied other humans’ solutions with the same performance.

Our first contribution is expanding previous proposals in cultural evolution by suggesting a relatively unexplored area of investigation, namely hybrid social learning. Scholars of cultural evolution have long investigated how social learning could lead to the unmatched explosion of human cultural complexity in comparison to non-human animals [35, 50, 51]. Similarly, we might ask if the advent of self-learning algorithms can lead to another acceleration in cultural evolution via hybrid human-algorithm social learning. Going one step beyond prior work looking into cultural evolution with digital technology [17, 52], we suggest that in a hybrid society, algorithm may not be just a medium for cultural transmission and evolution, but they may play an active role in the production of cultural artifacts. In particular, we suggest that successful hybrid social learning may occur when algorithms, either by design or by self-learning, show different biases than their human counterparts. Algorithms increasingly learn from interactions with their environments, thereby showing behaviors and biases that are independent of humans. Greater variance in problem-solving and copying skills has been associated with greater cultural variance [53, 54] and—as long as there are selection biases in who to learn from—greater innovation. We looked at situations where human biases are known to constrain human performance [29], and therefore humans could benefit most from observing an algorithmic strategy.

In our experiment, we tested these hypotheses by introducing algorithmic players adopting different decision-making strategies than human players. Investigating hybrid groups of human and algorithmic players provides the experimenter with the advantage to closely control the behavior of algorithmic agents while observing the effect on the rest of the population [55–57], yet, to the best of our knowledge, bots have not prominently featured in transmission chain experiments.

Our second contribution lies in our empirical findings. We showed that humans did not preserve algorithmic solutions if they were incongruent with human exploration. Although human and algorithmic biases have been thoroughly investigated in their respective fields (psychology/economics and computer science), how the two interact together is still poorly understood. We show that learning from algorithms might be limited by task and cognitive constraints. This is in contrast with evidence showing that algorithms take in human biases [58–60].

In our experiment, higher-performing solutions that were incongruent with human biases were not copied, and consequently lost over generations. Such preference for copying congruent solutions may limit high-fidelity copying and thus the accumulation of algorithmic solutions into human repertoire [50, 54]. These findings are in agreement with Griffiths, Kalish, and Lewandowsky [61] analytical works, which suggests that when the bias of participants and fitness of solutions mismatch, high fidelity in the transmission is needed for the superior solution to be adopted. In follow up work, Thompson and Griffiths [62] modeled cultural evolution in transmission chain experiments as being influenced by attraction towards preexisting biases, and local innovations. If both conflict, they showed experimentally that participants' solutions converge to a middle ground. In their model, inductive biases transform artifacts; in our agent-based model, content biases control imitation, yet biases hinder the discovery of optimal solutions in both cases. Our work goes beyond their findings and suggests that even if an algorithm aids humans in archiving optimal solutions, humans' bias in whom and what to copy can lead to those solutions being quickly lost in successive human-human transmissions [9, 44, 47, 61].

A lot of work in the field of cultural evolution tried to scale findings in the lab to the real-world [36, 63], including critical discussions about the limitations of such extrapolations [64]. We suggest that hybrid social interaction among human and algorithmic players may play an increasingly critical role in today's digital society. Investigating human-algorithmic social learning in the lab is the first step to study how these phenomena might unfold in the real-world, and how interactions in hybrid social systems may foster or hinder innovations and collective performance.

Designing algorithms to nudge collective behavior may add to an already long list of ethical concerns in AI [58, 60, 65, 66]. However, from an empirical point of view, our results seem to suggest that even if algorithms were designed to improve human performance, the features of what behaviors people copy and who are they willing to copy from may limit social learning, especially in situations of uncertainty, high cognitive demand, or high time pressure. Under these conditions, humans are likely to follow well-known and adaptive biases[20, 22]. We acknowledge the limitations of our study, both in terms of generalisability and sample size. Future studies will need to address whether AI-human collaboration may be more successful in other domains or simpler environments.

In our experiment, we were interested to isolate cultural transmission by exposing participants to one solution only. This may limit the generalisability of our study. Outside the lab, people can copy from multiple models, which may give them the option to compare human and algorithmic solutions. Also, while in our experiment we tested the effect of a single algorithmic player, the frequency of encountering algorithmic generated solutions in the real-world may be higher. Our agent-based model (Figure 3) suggests that sustained improvement in performance might be observed with greater chances to copy from algorithms, but more work is needed. Finally, in our experiment people visited each environment only once. This likely reduced the effect of individual learning as well as given participants inadequate feedback on their performance. Repeated unsuccessful feedback with the same environment before being exposed to a superior algorithmic solution might give participants additional opportunities to copy the algorithm.

In this work we focused on the transmission of behaviour, rather then then the transmission of strategy. Social learning seem to be more effective when copying exact behaviors rather than strategies and reasoning [35, 67, 68]. Copying reasoning strategies is not a prerequisite for cultural evolution [69]. Yet, an more interactive communication between generations, i.e. teaching, could allow for the transmission of strategy. It has been shown that communication of intention improves human-algorithm cooperation [70]. Correspondingly, we hypothesise that an algorithm that communicates the reasoning behind a solution might be copied at a higher rate and allow following humans to contest their preexisting beliefs.

To conclude, in this work, we found limited influence of bots on long-term human cultural evolution. This finding by no means suggests that there are no algorithmic influences on human culture. The relationship between human biased strategies and the algorithmic strategies derived by self-play might look different outside the lab where complex AI algorithms are at play. However, studying these phenomena in controlled environment is an important first step to understand hybrid social learning. In this study we suggested that differences between human and AI behavior might be relevant for emerging properties in cultural evolution.

References

[1] Wired. *In Two Moves, AlphaGo and Lee Sedol Redefined the Future | WIRED*. 2016. URL: https://www.wired.com/2016/03/two-moves-alphago-lee-sedol-redefined-future/.

[2] David Silver et al. "Mastering the game of Go with deep neural networks and tree search". In: *Nature* 529.7587 (Jan. 2016), pp. 484–489. ISSN: 14764687. DOI: 10.1038/nature16961. URL: https://www.nature.com/articles/nature16961.

[3] David Silver et al. "Mastering the game of Go without human knowledge". In: *Nature* 550.7676 (Oct. 2017), pp. 354–359. ISSN: 14764687. DOI: 10.1038/nature24270. URL: https://www.nature.com/articles/nature24270.

[4] Gary Linscott, Alexander Lyashuk, and Gian-Carlo Pascutto. *Leela Chess Zero*. 2018. URL: https://lczero.org/.

[5] Yuandong Tian et al. "Elf OpenGo: An analysis and open reimplementation of Alphazero". In: *36th International Conference on Machine Learning, ICML 2019* 2019-June (Feb. 2019), pp. 10885–10894. URL: http://arxiv.org/abs/1902.04522.

[6] Yuandong Tian et al. *Open-sourcing a new ELF OpenGo bot and related Go research*. URL: https://ai.facebook.com/blog/open-sourcing-new-elf-opengo-bot-and-go-research/.

[7] Minkyu Shin, Jin Kim, and Minkyung Kim. "Measuring Human Adaptation to AI in Decision Making: Application to Evaluate Changes after AlphaGo". In: (2020). URL: http://arxiv.org/abs/2012.15035.

[8] Anthony Strittmatter, Uwe Sunde, and Dainis Zegners. "Life cycle patterns of cognitive performance over the long run". In: *Proceedings of the National Academy of Sciences of the United States of America* 117.44 (2020), pp. 27255–27261. ISSN: 10916490. DOI: 10.1073/pnas.2006653117.

[9] Rachel L. Kendal et al. *Social Learning Strategies: Bridge-Building between Fields*. July 2018. DOI: 10.1016/j.tics.2018.04.003. URL: http://www.cell.com/article/S1364661318300949/fulltext%20http://www.cell.com/article/S1364661318300949/abstract%20https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(18)30094-9.

[10] Iyad Rahwan et al. *Machine behaviour*. Apr. 2019. DOI: 10.1038/s41586-019-1138-y. URL: https://doi.org/10.1038/s41586-019-1138-y.

[11] Jan Lorenz et al. "How social influence can undermine the wisdom of crowd effect". In: *Proceedings of the National Academy of Sciences of the United States of America* 108.22 (May 2011), pp. 9020–9025. ISSN: 00278424. DOI: 10.1073/pnas.1008636108. URL: www.pnas.org/cgi/doi/10.1073/pnas.1008636108.

[12] Asher Koriat. "When two heads are better than one and when they can be worse: The amplification hypothesis". In: *Journal of Experimental Psychology: General* 144.5 (2015), pp. 934–950. ISSN: 00963445. DOI: 10.1037/xge0000092. URL: http://dx.doi.org/10.1037/xge0000092.

[13] Asher Koriat. “When are two heads better than one and why?” In: *Science* 336.6079 (Apr. 2012), pp. 360–362. ISSN: 10959203. DOI: 10.1126/science.1216549. URL: http://science.sciencemag.org/.

[14] Ramsey M. Raafat, Nick Chater, and Chris Frith. *Herding in humans.* Oct. 2009. DOI: 10.1016/j.tics.2009.08.002.

[15] Lu Hong and Scott E. Page. “Groups of diverse problem solvers can outperform groups of high-ability problem solvers”. In: *Proceedings of the National Academy of Sciences of the United States of America* 101.46 (Nov. 2004), pp. 16385–16389. ISSN: 00278424. DOI: 10.1073/pnas.0403723101. URL: www.pnas.orgcgidoi10.1073pnas.0403723101.

[16] Philip E. Tetlock and Philip E. Tetlock. *Expert Political Judgment.* STU - Stud. Princeton University Press, Mar. 2017. ISBN: 9780691128719. DOI: 10.1515/9781400888818. URL: http://www.jstor.org/stable/j.ctt7spbt.

[17] Alberto Acerbi. *Cultural Evolution in the Digital Age.* Oxford University Press, 2019. DOI: 10.1093/oso/9780198835943.001.0001.

[18] Cecilia Heyes. *Cognitive Gadgets.* Harvard University Press, Aug. 2018. DOI: 10.4159/9780674985155. URL: /record/2017-47901-000.

[19] Amos Tversky and Daniel Kahneman. “Judgment under uncertainty: Heuristics and biases”. In: *Science* 185.4157 (1974), pp. 1124–1131. ISSN: 00368075. DOI: 10.1126/science.185.4157.1124. URL: http://science.sciencemag.org/.

[20] Mike Oaksford and Nick Chater. “A Rational Analysis of the Selection Task as Optimal Data Selection”. In: *Psychological Review* 101.4 (1994), pp. 608–631. ISSN: 0033295X. DOI: 10.1037/0033-295X.101.4.608.

[21] John M.C. Hutchinson and Gerd Gigerenzer. “Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet”. In: *Behavioural Processes.* Vol. 69. 2. 2005, pp. 97–124. DOI: 10.1016/j.beproc.2005.02.019.

[22] Peter M. Todd and Gerd Gigerenzer. “Environments that make us smart: Ecological rationality”. In: *Current Directions in Psychological Science* 16.3 (June 2007), pp. 167–171. ISSN: 09637214. DOI: 10.1111/j.1467-8721.2007.00497.x. URL: http://journals.sagepub.com/doi/10.1111/j.1467-8721.2007.00497.x.

[23] Noel G. Machado and Jerrin Samuel. *Predictably Irrational Hiring.* Harper Audio New York, NY, 2021, p. 263145412098841. DOI: 10.1177/2631454120988415.

[24] W. P. Bottom. *Heuristics and Biases: The Psychology of Intuitive Judgment.* Vol. 29. 4. Cambridge university press, 2004, pp. 695–698. DOI: 10.5465/amr.2004.14497675.

[25] Niccolò Pescetelli and Nicholas Yeung. “The Role of Decision Confidence in Advice-Taking and Trust Formation”. In: *Journal of Experimental Psychology: General* 150.3 (2020). ISSN: 00963445. DOI: 10.1037/xge0000960. URL: https://pubmed.ncbi.nlm.nih.gov/33001684/.

[26] George Loewenstein and Jon Elster. *Choice over time.* Russell Sage Foundation, 1992.

[27] Joel Myerson and Leonard Green. “Discounting of Delayed Rewards: Models of Individual Choice”. In: *Journal of the Experimental Analysis of Behavior* 64.3 (Nov. 1995), pp. 263–276. ISSN: 0022-5002. DOI: 10.1901/jeab.1995.64-263. URL: http://doi.wiley.com/10.1901/jeab.1995.64-263.

[28] Shane Frederick, George Loewenstein, and Ted O’Donoghue. “Time discounting and time preference: A critical review”. In: *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice* XL.June (2003), pp. 13–86. ISSN: 0022-0515. DOI: 10.1257/002205102320161311.

[29] Quentin J.M. Huys et al. “Bonsai trees in your head: How the pavlovian system sculpts goal-directed choices by pruning decision trees”. In: *PLoS Computational Biology* 8.3 (Mar. 2012), p. 1002410. ISSN: 1553734X. DOI: 10.1371/journal.pcbi.1002410. URL: www.ploscompbiol.org.

[30] Arthur Guez, David Silver, and Peter Dayan. “Scalable and efficient bayes-adaptive reinforcement learning based on Monte-Carlo tree search”. In: *Journal of Artificial Intelligence Research* 48 (Nov. 2013), pp. 841–883. ISSN: 10769757. DOI: 10.1613/jair.4117. URL: https://www.jair.org/index.php/jair/article/view/10853.

[31] Donald E Knuth and Ronald W Moore. *An Analysis of Alpha-Beta Priming’.* Tech. rep. 1975, pp. 293–326.

[32] Quentin J.M. Huys et al. "Interplay of approximate planning strategies". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.10 (2015), pp. 3098–3103. ISSN: 10916490. DOI: 10.1073/pnas.1414219112. URL: www.pnas.org/lookup/suppl/doi:10.1073/pnas.1414219112/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1414219112.

[33] Falk Lieder et al. "Cognitive prostheses for goal achievement". In: *Nature Human Behaviour* 3.10 (Oct. 2019), pp. 1096–1106. ISSN: 23973374. DOI: 10.1038/s41562-019-0672-9. URL: https://doi.org/10.1038/s41562-019-0672-9.

[34] L. Rendell et al. "Why copy others? insights from the social learning strategies tournament". In: *Science* 328.5975 (Apr. 2010), pp. 208–213. ISSN: 00368075. DOI: 10.1126/science.1184719. URL: https://science.sciencemag.org/content/328/5975/208%20https://science.sciencemag.org/content/328/5975/208.abstract.

[35] L. G. Dean et al. "Identification of the social and cognitive processes underlying human cumulative culture". In: *Science* 335.6072 (Mar. 2012), pp. 1114–1118. ISSN: 10959203. DOI: 10.1126/science.1213969. URL: http://science.sciencemag.org/.

[36] Alex Mesoudi and Andrew Whiten. "The multiple roles of cultural transmission experiments in understanding human cultural evolution". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 363.1509 (2008), pp. 3489–3501. ISSN: 09628436. DOI: 10.1098/rstb.2008.0129. URL: http://dx.doi.org/10..

[37] Christine A. Caldwell and Ailsa E. Millen. *Studying cumulative cultural evolution in the laboratory.* Nov. 2008. DOI: 10.1098/rstb.2008.0133. URL: /pmc/articles/PMC2607341/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2607341/.

[38] Christine A. Caldwell and Ailsa E. Millen. "Experimental models for testing hypotheses about cumulative cultural evolution". In: *Evolution and Human Behavior* 29.3 (May 2008), pp. 165–171. ISSN: 10905138. DOI: 10.1016/j.evolhumbehav.2007.12.001.

[39] Amanda J. Lucas et al. "The value of teaching increases with tool complexity in cumulative cultural evolution: Value of teaching in cumulative culture". In: *Proceedings of the Royal Society B: Biological Sciences* 287.1939 (2020), p. 20201885. ISSN: 14712954. DOI: 10.1098/rspb.2020.1885rspb20201885.

[40] F. C. Bartlett and Cyril Burt. "REMEMBERING: A STUDY IN EXPERIMENTAL AND SOCIAL PSYCHOLOGY". In: *British Journal of Educational Psychology* 3.2 (June 1933), pp. 187–192. ISSN: 00070998. DOI: 10.1111/j.2044-8279.1933.tb02913.x. URL: http://doi.wiley.com/10.1111/j.2044-8279.1933.tb02913.x.

[41] Yoshihisa Kashima. "Maintaining cultural stereotypes in the serial reproduction of narratives". In: *Personality and Social Psychology Bulletin* 26.5 (May 2000), pp. 594–604. ISSN: 01461672. DOI: 10.1177/0146167200267007. URL: http://journals.sagepub.com/doi/10.1177/0146167200267007.

[42] Alex Mesoudi, Andrew Whiten, and Robin Dunbar. "A bias for social information in human cultural transmission". In: *British Journal of Psychology* 97.3 (2006), pp. 405–423. ISSN: 20448295. DOI: 10.1348/000712605X85871.

[43] Michael L. Kalish, Thomas L. Griffiths, and Stephan Lewandowsky. "Iterated learning: Intergenerational knowledge transmission reveals inductive biases". In: *Psychonomic Bulletin and Review.* Vol. 14. 2. Springer New York LLC, 2007, pp. 288–294. DOI: 10.3758/BF03194066. URL: https://link.springer.com/article/10.3758/BF03194066.

[44] Joseph Henrich and Richard McElreath. *The Evolution of Cultural Evolution.* Jan. 2003. DOI: 10.1002/evan.10110. URL: https://onlinelibrary.wiley.com/doi/full/10.1002/evan.10110%20https://onlinelibrary.wiley.com/doi/abs/10.1002/evan.10110%20https://onlinelibrary.wiley.com/doi/10.1002/evan.10110.

[45] Alex Mesoudi. "Pursuing Darwin's curious parallel: Prospects for a science of cultural evolution". In: *Proceedings of the National Academy of Sciences of the United States of America* 114.30 (July 2017), pp. 7853–7860. ISSN: 10916490. DOI: 10.1073/pnas.1620741114. URL: www.pnas.org/cgi/doi/10.1073/pnas.1620741114.

[46] Berkeley J. Dietvorst, Joseph P Simmons, and Cade Massey. "Algorithm aversion: People erroneously avoid algorithms after seeing them err". In: *Journal of Experimental Psychology: General* 144.1 (2015), pp. 114–126. ISSN: 00963445. DOI: 10.1037/xge0000033. URL: http://repository.upenn.edu/fnce_papershttp://dx.doi.org/10.1037/xge0000033.

[47] César A Hidalgo et al. *How Humans Judge Machines.* MIT Press, 2021.

[48] Alberto Acerbi, Claudio Tennie, and Alex Mesoudi. "Social learning solves the problem of narrow-peaked search landscapes: Experimental evidence in humans". In: *Royal Society Open Science* 3.9 (Sept. 2016), p. 160215. ISSN: 20545703. DOI: 10.1098/rsos.160215. URL: https://royalsocietypublishing.org/doi/10.1098/rsos.160215.

[49] Abdullah Almaatouq et al. "Empirica: a virtual lab for high-throughput macro-level experiments". In: *Behavior Research Methods* (Mar. 2021), pp. 1–14. ISSN: 15543528. DOI: 10.3758/s13428-020-01535-9. URL: https://doi.org/10.3758/s13428-020-01535-9.

[50] Claudio Tennie, Josep Call, and Michael Tomasello. "Ratcheting up the ratchet: On the evolution of cumulative culture". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364.1528 (2009), pp. 2405–2415. ISSN: 14712970. DOI: 10.1098/rstb.2009.0052.

[51] Alex Mesoudi and Alex Thornton. "What is cumulative cultural evolution?" In: *Proceedings of the Royal Society B: Biological Sciences* 285.1880 (2018). ISSN: 14712954. DOI: 10.1098/rspb.2018.0712.

[52] Simon Carrignon, R. Alexander Bentley, and Damian Ruck. "Modelling rapid online cultural transmission: evaluating neutral models on Twitter data with approximate Bayesian computation". In: *Palgrave Communications* 5.1 (Dec. 2019), pp. 1–9. ISSN: 20551045. DOI: 10.1057/s41599-019-0295-9. URL: https://doi.org/10.1057/s41599-019-0295-9.

[53] Michael Muthukrishna and Joseph Henrich. *Innovation in the collective brain*. Mar. 2016. DOI: 10.1098/rstb.2015.0192. URL: http://dx.http//dx.doi.org/10.1098/rstb.2015.0192orviahttp://rstb.royalsocietypublishing.org..

[54] Joseph Henrich. "Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses—The Tasmanian Case". In: *American Antiquity* 69.2 (Apr. 2004), pp. 197–214. ISSN: 0002-7316. DOI: 10.2307/4128416.

[55] Hirokazu Shirado and Nicholas A. Christakis. "Locally noisy autonomous agents improve global human coordination in network experiments". In: *Nature* 545.7654 (May 2017), pp. 370–374. ISSN: 14764687. DOI: 10.1038/nature22332. URL: https://www.nature.com/articles/nature22332.

[56] Malte Jung and Pamela Hinds. "Robots in the Wild". In: *ACM Transactions on Human-Robot Interaction* 7.1 (2018). ISSN: 25739522. DOI: 10.1145/3208975. URL: https://doi.org/10.1145/3208975.

[57] Margaret L. Traeger et al. "Vulnerable robots positively shape human conversational dynamics in a human-robot team". In: *Proceedings of the National Academy of Sciences of the United States of America* 117.12 (Mar. 2020), pp. 6370–6375. ISSN: 10916490. DOI: 10.1073/pnas.1910402117. URL: https://www.pnas.org/content/117/12/6370%20https://www.pnas.org/content/117/12/6370.abstract.

[58] Francis Schrag. *Review of Weapons of math destruction: How big data increases inequality and threatens democracy*. Vol. 24. USA: Crown Publishing Group, 2017. ISBN: 0553418815. DOI: 10.14507/er.v24.2197.

[59] Aylin Caliskan, Joanna J. Bryson, and Arvind Narayanan. "Semantics derived automatically from language corpora contain human-like biases". In: *Science* 356.6334 (Apr. 2017), pp. 183–186. ISSN: 10959203. DOI: 10.1126/science.aal4230. URL: http://science.sciencemag.org/.

[60] Joy Buolamwini. *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification **. Tech. rep. Jan. 2018, pp. 1–15. URL: http://proceedings.mlr.press/v81/buolamwini18a.html.

[61] Thomas L. Griffiths, Michael L. Kalish, and Stephan Lewandowsky. "Theoretical and empirical evidence for the impact of inductive biases on cultural evolution". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 363.1509 (Nov. 2008), pp. 3503–3514. ISSN: 0962-8436. DOI: 10.1098/rstb.2008.0146. URL: https://royalsocietypublishing.org/doi/10.1098/rstb.2008.0146.

[62] Bill Thompson and Thomas L Griffiths. *Inductive Biases Constrain Cumulative Cultural Evolution*. Tech. rep. 2019.

[63] Andrew Whiten and Alex Mesoudi. *Establishing an experimental science of culture: Animal social diffusion experiments*. Nov. 2008. DOI: 10.1098/rstb.2008.0134. URL: https://royalsocietypublishing.org/doi/10.1098/rstb.2008.0134.

[64] Helena Miton and Mathieu Charbonneau. "Cumulative culture in the laboratory: Methodological and theoretical challenges". In: *Proceedings of the Royal Society B: Biological Sciences* 285.1879 (May 2018), p. 20180677. ISSN: 14712954. DOI: 10.1098/rspb.2018.0677. URL: https://royalsocietypublishing.org/doi/10.1098/rspb.2018.0677.

[65] Shoshana Zuboff. "The age of surveillance capitalism: The fight for a human future at the new frontier of power". In: *Yale Law Journal* 129.5 (2020), pp. 1460–1515. ISSN: 00440094. DOI: 10.26522/brocked.v29i2.849.

[66] Sankalp Bhatnagar et al. *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation Authors are listed in order of contribution Design Direction*. Tech. rep. 2018.

[67] Iyad Rahwan et al. "Analytical reasoning task reveals limits of social learning in networks". In: *Journal of the Royal Society Interface* 11.93 (Apr. 2014). ISSN: 17425662. DOI: 10.1098/rsif.2013.1211. URL: http://dx.doi.org/10.1098/rsif.2013.1211.

[68] Edwin J.C. Van Leeuwen, Josep Call, and Daniel B.M. Haun. "Human children rely more on social information than chimpanzees do". In: *Biology Letters* 10.11 (Nov. 2014). ISSN: 1744957X. DOI: 10.1098/rsbl.2014.0487. URL: http://dx.doi.org/10.1098/rsbl.2014.0487orviahttp://rsbl.royalsocietypublishing.org..

[69] Maxime Derex et al. "Causal understanding is not necessary for the improvement of culturally evolving technology". In: *Nature Human Behaviour* 3.5 (2019), pp. 446–452. ISSN: 23973374. DOI: 10.1038/s41562-019-0567-9. URL: http://dx.doi.org/10.1038/s41562-019-0567-9.

[70] Jacob W. Crandall et al. "Cooperating with machines". In: *Nature Communications* 9.1 (Dec. 2018), pp. 1–12. ISSN: 20411723. DOI: 10.1038/s41467-017-02597-8. URL: www.nature.com/naturecommunications.