





ARTICLE



<https://doi.org/10.1057/s41599-022-01305-2>

OPEN

Quantifying effects of tasks on group performance in social learning

Gengjun Yao^{1,3}, Jingwei Wang^{2,3}, Baoguo Cui¹ & Yunlong Ma²  

Social learning is a learning process in which new behaviors can be acquired by observing and imitating others. It is the key to cultural evolution because individuals can exchange profitable information culturally within the group. Recent studies have over-focused on social learning strategies but paid rare attention to the learning tasks. In particular, in these studies, individuals rely on perfect imitation, directly copying the solutions of others, to improve their performance. However, imperfect imitation, a prevalent form of social learning in cultural evolution, has received little discussion. In this paper, the effects of three task features (task types, task complexity, and task granularity) on group performance are simulated with an agent-based model and quantified with decision trees. In the proposed model, individuals in a network learn from others via imperfect imitation, which means individuals make a trade-off between their solutions and socially acquired solutions. Here, status quo bias is introduced to represent the degree to which individuals adhere to their solutions. Results show that the performance of a group is not affected by task complexity in hard-to-easy tasks but declines with the task complexity rising in easy-to-hard tasks. Besides, groups usually perform better in fine-grained tasks than in coarse-grained ones. The main reason is that in coarse-grained tasks, conservative individuals encounter learning bottlenecks that prevent them from exploring superior solutions further. Interestingly, increasing task granularity can mitigate this disadvantage for conservative individuals. Most strikingly, the importance scores given by decision trees suggest that tasks play a decisive role in social learning. These findings provide new insights into social learning and have broad implications for cultural evolution.

¹School of Journalism and Communication, Tsinghua University, Beijing, China. ²College of Electronic and Information Engineering, Tongji University, Shanghai, China. ³These authors contributed equally: Gengjun Yao, Jingwei Wang. ✉email: evanma@tongji.edu.cn

Introduction

One of the forces driving human cultural evolution is the transmission of information capable of affecting individuals' behavior (Boyd et al., 2011; Flinn, 1997). Many studies have suggested that social learning, i.e., learning facilitated by observing and imitating other members' behavior (Cantor et al., 2015; Derex and Boyd, 2016; Rogers, 1988; Van Leeuwen et al., 2018), plays a central role in human and other animal species' cultural evolution (Mesoudi and Thornton, 2018). At a micro level, individuals relying on social learning can seek superior solutions by imitating the behavior of others. At a macro level, such individual-level interactions via social learning allow for the diffusion of excellent solutions among the population, resulting in the improvement of group-level performance (Barkoczi and Galesic, 2016; Csaszar and Science, 2010; Mason and Watts, 2012; Rendell et al., 2010).

Recent work has made significant progress in explaining the mechanisms of how social learning contributes to group performance. These studies have extensively discussed the effects of social learning strategies (how individuals learn from others) (Barkoczi and Galesic, 2016; Fogarty et al., 2012; Molleman et al., 2014) and the structure of communication networks (in which information diffusion occurs) (Shi et al., 2017; Lazer and Friedman, 2007; Wisdom et al., 2013; Laland, 2004) among individuals on group performance (Kendal et al., 2018; Shore et al., 2015; Lamberson, 2010; Fang et al., 2010). However, they still have some limitations in helping us to fully understand social learning. First, many studies conducted experiments only on a single type of task, and only a few considered the diversity of task types (Acerbi et al., 2016; Morgan et al., 2012) or assumed that different tasks differ only in complexity (Barkoczi et al., 2016; Almaatouq et al., 2021). However, the difference between tasks is not only in complexity but also in granularity, a seriously overlooked feature. Such neglect of task features may lead to a significant underestimation of the role of tasks in social learning.

Second, much of the work regards social learning as perfect imitation, i.e., a simple replication mechanism. For example, in the study of Barkoczi and Galesic (Barkoczi and Galesic, 2016), individuals relying on social learning are only responsible for exploitation (i.e., copying existing solutions), and new solutions are explored by individuals alone. In contrast, a few recent studies (Derex et al., 2015; Morin et al., 2021) found that social learning leads to the generation of new solutions combining information from multiple sources rather than a simple dichotomous choice of whether to copy other solutions. This finding suggests that social learning should be viewed as imperfect imitation, allowing individuals to diffuse superior solutions and explore new solutions (Haviland et al., 2013).

Third, many recent studies (Barkoczi and Galesic, 2016; Mason and Watts, 2012) have revealed complex interactions between variables, such as social learning strategies and network structure. These findings make it extremely difficult to quantify the effect of each variable on social learning. The above problems indicate that our current understanding of social learning may be incomplete. Therefore, this paper aims to address the following three questions on social learning: First, how do tasks affect group performance? Second, what performance can be achieved by groups relying on social learning? Third, how do we quantify the effects of different variables on group performance?

To address these problems, we develop a multi-agent model of social learning, where a group consisting of many individuals repeatedly searches for solutions that can improve group performance. Following previous studies (Kauffman and Levin, 1987; Levinthal, 1997), we model social learning as the search on rugged landscapes with peaks and valleys. Individuals are poorly informed about the rugged landscapes and can only interact with

their neighbors to explore better solutions. The great challenge individuals face is not to get stuck in local optima (peaks), which is common in such environments.

We create different rugged landscapes to represent two typical learning tasks: the hard-to-easy and easy-to-hard tasks (Kurt and Ehret, 2010; Levinthal, 1997). The learning curves for these two tasks are the exact opposite (Lieberman, 1987; Kauffman and Levin, 1987). In the hard-to-easy task, individuals feel difficult at the beginning of learning and then find it easier as they go along. In contrast, individuals feel easy at first but find it increasingly difficult over time in the easy-to-hard task. The two types of tasks are common in real life. For example, beginner violinists need to spend substantial time finding pitches, so they have difficulty playing a complete piece in the early stage of learning. Beginning pianists, on the other hand, can do this easily. However, they have to devote more effort to practice as the complexity of the piece increases. Thus, learning the violin can be viewed as a hard-to-easy task, while learning the piano can be considered an easy-to-hard task (Jordà, 2004).

Learning is not always easy, and learners will experience major or minor setbacks in the process. For example, piano learners can easily learn to play the beginning repertoire but struggle to master the more complex techniques (such as chords) required for the advanced repertoire. They have to face repeated failures before they can play these advanced pieces. In many studies, the factor characterizing the degree of this fluctuation is called task complexity. In our model, the complexity of a task is denoted by the number of peaks (Barkoczi et al., 2016; Barkoczi and Galesic, 2016), which are tunable in the two above tasks. Peaks can be viewed as locally optimal solutions, and complex tasks contain more locally optimal solutions than simple ones.

Especially, we introduce task granularity, an important task feature ignored in previous studies, to quantify the degree to which a task can be decomposed (Ethiraj et al., 2008; Pil and Cohen, 2006) or the roughness of individual perception of a task. Many studies in cognitive science have shown that human cognitive processes require discrete representations (Boyd and Henrich, 2002; Dietrich and Markman, 2003; Sperber, 1996). One reason is that there is usually an upper limit to the human perceptual system, and they are not sensitive to subtle changes. For example, the smallest frequency change that normal-hearing adults can detect is of the order of 0.2–0.3% (about 8–12 Hz) for frequencies between 250 and 4000 Hz (Moore, 1974). This means that an adult cannot accomplish the task of distinguishing 1000 Hz tones and 1001 Hz tones. Similar limitations exist for manufactured systems, such as the porcelain and precision industries. In these systems, the measurement and control of physical quantities such as position and temperature are extremely important. If the sensing or control precision of a system is not up to the required level, it will not be able to produce qualified products. This partly explains why developing countries need to import advanced equipment from developed countries.

In our model, all individuals in a group are embedded in a communication network with a specific structure, and one can only interact with its neighbors. We consider four typical network structures: fully connected network, locally connected lattice, Watts-Strogatz network, and Barabasi-Albert network (Barabási and Albert, 1999; Watts and Strogatz, 1998). These networks cover a broad range of possible social connection patterns between individuals. Individuals in a network interact with others through social learning. Here, social learning is regarded as an imperfect imitation that individuals partially imitate the behavior of others (Derex et al., 2015; Posen et al., 2013; Rogers, 2010). One reason is that individuals usually have status quo bias (SQB) (Crawford, 1995; Friedkin and Johnsen, 1990; Morin et al., 2021;

Posen et al., 2013), which refers to the preference of individuals for their current state; so that they do not copy others' solutions exactly. In contrast, they will combine information from themselves and others to produce new solutions. An individual's willingness to imitate the behavior of others depends on its SQB, and each individual is assigned a value representing its degree of SQB in our model. Conservative individuals who prefer their solution strongly have large SQB values, whereas open individuals are assigned small SQB values. Then, the group conservativeness can be denoted by the SQB distribution of all individuals in a group; that is, conservative-biased groups contain more conservative individuals than open-biased groups.

Before each simulation starts, we specify the task and the network structure of a group and its SQB distribution. Initially, each individual is randomly assigned a solution with a payoff sampled from the task environment. Then, individuals continue to find new solutions by learning from their neighbors. On each step, an individual can use three different strategies to select a neighbor for learning, including the best strategy (learn from the best neighbor), conformity strategy (learn from the majority of neighbors), and the random strategy (learn from anyone neighbor) (Barkoczi and Galesic, 2016; Kendal et al., 2018; Laland, 2004; Zhang and Gläscher, 2020). After that, one individual observes whether the payoff of its neighbor's solution is higher than itself. If yes, the individual employs an imperfect imitation policy and obtains a new solution fused by its solution and the socially acquired solution; otherwise, the individual holds its original solution. We iterate the procedure for 500 steps and record the average payoff in the group on each step separately for each combination of the task features, group features, and strategy-related variables. Results reported are averaged across multiple repetitions.

We first identify five variables greatly affecting group performance through extensive simulation experiments: three task features (task types, task complexity, and task granularity), group conservativeness, and social learning strategies. Specifically, task granularity and group conservativeness interact to impact group performance. Our results show that other variables, which have been suggested to affect group performance in previous studies, such as network structure, network density, and group size, do not have significant effects (Supplementary Note 4). Moreover, we use a popular machine learning method, the decision tree, to quantify the importance of the above five variables. As we show next, three task features account for 78.6% of the importance in predicting group performance, suggesting that tasks play a decisive role in social learning.

Methods

Task environments. A few studies used the NK model to generate "tunable rugged" landscapes (Barkoczi and Galesic, 2016; Csaszar and Science, 2010). The ruggedness of a landscape generated by the NK model is determined by N , the number of components that make up each solution, and K , the number of interdependencies between the N components. In particular, each solution is represented by an N -length vector composed of binary digits, leading to a total of 2^N possible solutions in a task environment. The payoff of each solution is calculated as the average of the payoff contribution of each element, which is determined by the other $K - 1$ elements (see for more details). We found that the NK model has inherent defects in generating task environments for social learning. First, the payoff contribution of each element is a random number drawn from a uniform distribution between 0 and 1. Thus, the payoff of a solution, calculated as the average of payoff contributions of the N elements, is also an uncertain number each time the landscape is regenerated. This result means

the same solution may have different payoffs even in landscapes created by the same NK model. Second, interdependencies between N elements in the NK model are randomly assigned after the landscape is created. As a result, an element is likely to be interdependent with different $K - 1$ other elements across repetitions under the same NK model configuration. This further increases the chance that the payoff of the same solution is different in repeated landscapes. Taken together, the great randomness in the NK model makes the performance of social learning strategies unstable and unreproducible.

Given the severe uncertainty of the NK model, we use the popular optimization test functions (Barkoczi et al., 2016; Mason and Watts, 2012; Mesoudi, 2008) to create multi-peaked task environments. Optimization test functions are regularly used in operations research and the field of global optimization to study how different optimization algorithms perform. These functions have been designed and studied carefully because they pose challenges to adaptive optimization algorithms, covering a wide range of possible environmental structures with regard to the variability of high-quality solutions, the ruggedness of the landscape, and the average payoff.

An optimization test function generally consists of two components, the trend term, and the fluctuation term. The trend term represents a certain trend of the function, such as an upward or downward trend, composed of polynomial, exponential or logarithmic functions. The fluctuation term composed of trigonometric functions drives the value of the function to change periodically, leading to many local minima. The trend term reflects the evolution of learning efficiency, while the fluctuation term reflects the learning risk.

We chose two popular optimization test functions (Ackley and Rastrigin) to represent two different task environments. For simplicity, we only consider the one-dimensional form $y = f(x)$ of the two functions, where the independent variable x represents the solution of the task and the dependent variable y represents the payoff of x . For each function, we transformed it so that both the solution and the payoff take values in $[0, 1]$, and there is only one globally optimal solution with the highest payoff (see Supplementary Note 1 for more details).

In particular, we use the transformed Ackley and Rastrigin functions to represent the hard-to-easy and the easy-to-hard tasks, respectively. The hard-to-easy tasks refer to those difficult when the individuals first learn them and then get easier as they go along. This characteristic fits well with the shape of the transformed Ackley function (Fig. 1a), where the function is flat far from the global optimum while changing rapidly closer to it. As the trend term of the Ackley function relies on the exponential function. Thus, in the hard-to-easy tasks, if an individual's initial solution is far from the optimal solution, then it can only make small progress with each step; conversely, if its initial solution is close to the optimal solution, then it can achieve a large increase in the payoff each time it moves. The easy-to-hard tasks are the opposite of the hard-to-easy tasks, which can be depicted by the transformed Rastrigin function (Fig. 1b). Since the trend term of the Rastrigin function depends on the quadratic function, this function is flat near the global optimum while falls rapidly far from it. Therefore, in the easy-to-hard tasks, individuals with initial solutions far from the global optimum learn fast at first while they get slow when their solutions approach the global optimum. We raise the payoff of the transformed Rastrigin function to the power of 8, following Lazer and Friedman (Lazer and Friedman, 2007). This operation ensures there are only a few solutions with high payoffs and most with low payoffs in the easy-to-hard task.

Task complexity is represented by the number of peaks in each function. Here, we test both functions with a wide range of complexity, varying the number of peaks from 1 to 16383. A

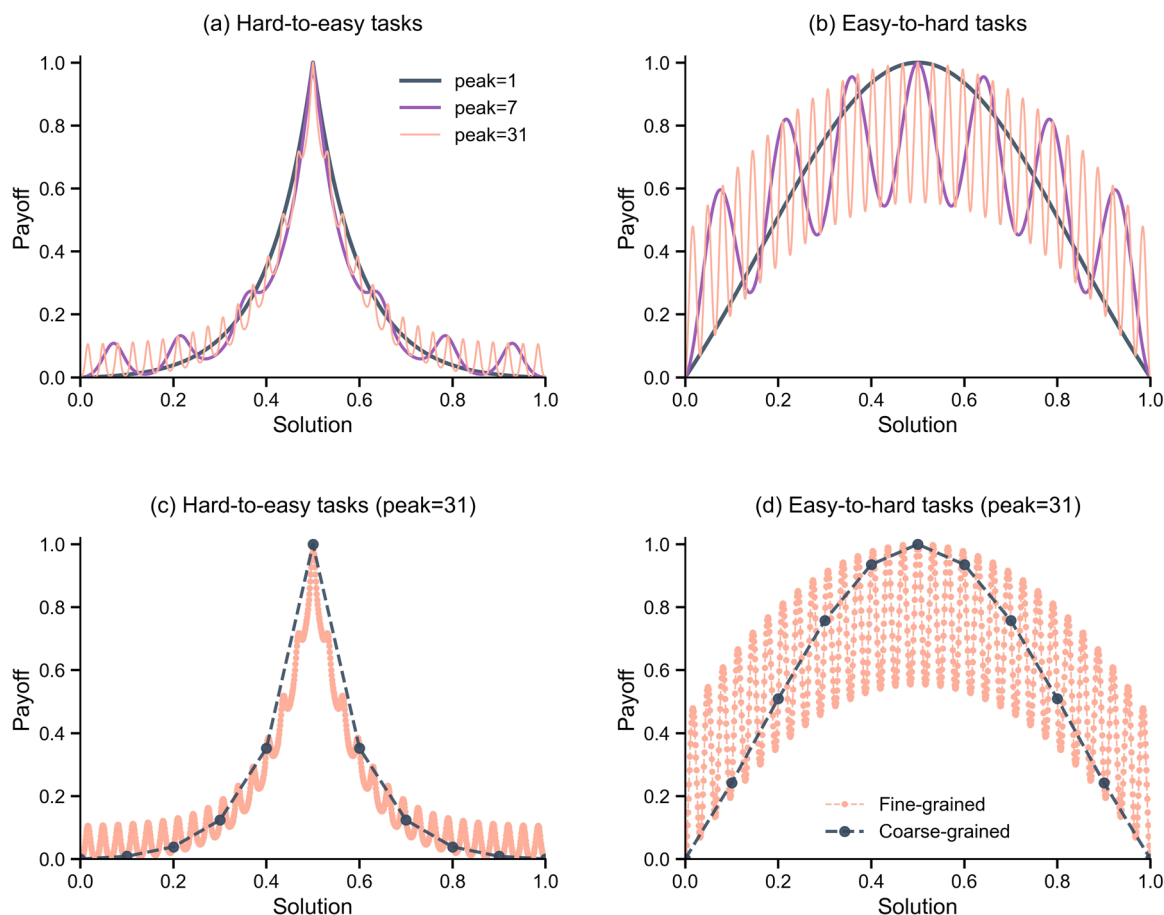


Fig. 1 The hard-to-easy and easy-to-hard tasks with different complexity and granularity. **a** The hard-to-easy tasks with 1, 7, and 31 peaks. **b** The easy-to-hard tasks with 1, 7, and 31 peaks. **c** Fine-grained and coarse-grained hard-to-easy tasks with 31 peaks. **d** Fine-grained and coarse-grained easy-to-hard tasks with 31 peaks. Fine-grained: $\omega = 3$. Coarse-grained: $\omega = 1$.

function with 16383 peaks is sufficient to represent a task with very high complexity, and its complexity is greater than that of the NK models in previous studies (Barkoczi et al., 2016; Barkoczi and Galesic, 2016). The mapping between task complexity C and the number of peaks p is defined as $C = \log_2(p + 1)$, that is, $p = 2^C - 1$ ($C = 1, 2, \dots, 14$). This mapping allows us to control the complexity of a task by adjusting the number of peaks, and the more complex the task, the more rugged the landscape becomes (see Fig. 1a, b). Depending on the value of C , we categorize the task complexity into three levels, namely low complexity ($1 \leq C \leq 5$), medium complexity ($6 \leq C \leq 10$), and high complexity ($11 \leq C \leq 14$).

Task granularity is used to represent the number of solutions in a task. Coarse-grained tasks have a few solutions, while fine-grained tasks have a great number of solutions (see Fig. 1c, d). As mentioned above, the optimization test function is defined in the interval $[0, 1]$. By dividing this interval into 10^ω parts equally, we can obtain $10^\omega + 1$ solutions for a task. Further, ω can be used to denote the granularity level, and a larger ω indicates finer granularity and more solutions. Figure 1c, d show the hard-to-easy and easy-to-hard tasks under coarse granularity ($\omega = 1$) and fine granularity ($\omega = 3$). In this paper, we vary the value of ω from 1 to 10. The task granularity was designed with reference to the studies in cognitive science (Boyd and Henrich, 2002; Dietrich and Markman, 2003). (See Supplementary Notes 1 and 2 for more details).

Generating the networks. We consider four kinds of networks, namely the fully connected (FC) network, locally connected

lattice (Lattice), Watts-Strogatz (WS) network, and Barabasi-Albert (BA) network (more details about all networks are provided in Supplementary Note 4). In the experiments presented in the main text, we use the FC network with 100 nodes, where each node represents an individual engaged in social learning. Then, we investigate the effect of the network size varying the number of nodes in the FC network from 50 to 5000 (Supplementary Fig. 12). To study the effect of network density on social learning, we generate various Lattice networks with 100 nodes varying the average degree of nodes from 10 to 98 (Supplementary Fig. 14). Note that the degree of a node is defined as the number of its neighbors. The network density is defined as $K/(N - 1)$, where K is the average degree of nodes and N is the number of nodes. Next, we use the WS network to study the effect of connection randomness, varying the rewiring probability from 0 to 1 (Supplementary Fig. 15). The rewiring probability that a node randomly reconnects to other nodes determines the randomness of the WS network. Finally, we perform simulation experiments on the four networks to investigate the effect of network structure on social learning (Supplementary Fig. 13). All networks are assumed to have 100 nodes and a fixed average degree of $K = 10$, except for the FC network where $K = N - 1$. The rewiring probability is 0.1 in the WS network.

Setting of group conservativeness. We use the status quo bias (SQB) to indicate the conservativeness of an individual, that is, the degree of preference for its current solution. We use $\phi \in [0, 1]$ to represent the SQB, and an individual with a large ϕ prefers to

believe in itself rather than accept others' solutions. Then, we define individuals with $\phi > 0.5$ as conservative ones and those with $\phi < 0.5$ as open ones. In particular, individuals with $\phi = 1$ are ultra-conservative and never accept others' solutions. Individuals with $\phi = 0$ are ultra-open and always copy others' solutions directly.

Next, we quantify the conservativeness of a group in terms of the SQB distribution of all individuals in the group. According to Rogers' research (Rogers, 2010), we classify groups into three types, namely the normal group, the conservative group, and the open group. The number of conservative and open individuals is equal in the normal group, and most individuals are biased to neutrality (i.e., the average SQB of all individuals is close to 0.5). The conservative group has a majority of conservative individuals (over 50%), with a mean SQB greater than 0.5. In contrast, the open population has less than 50% of conservative individuals with a mean SQB lower than 0.5. Here, we use mathematical tools to precisely control the SQB distribution. Specifically, we use the normal distribution, the left-skewed distribution (LSD), and the right-skewed distribution (RSD) to simulate the SQB distributions of the normal, conservative, and open groups, respectively (Supplementary Fig. 5). LSD and RSD are generated by the skew-normal distribution with a skewness parameter (Azzalini and Capitanio, 1999). We can control the proportion of conservative individuals in a group by varying the skewness parameter (Supplementary Table 1). Here, we study the effect of group conservativeness by varying the proportion of conservative individuals from 10 to 90%.

Social learning mechanism. Each individual in a group is randomly assigned an initial solution with a payoff sampled from the task environment. Most individuals have low payoffs, so they have to explore new solutions with high payoffs by learning from their neighbors. In each step, individuals find new solutions by applying the social learning mechanism consisting of search, decision-making, and imperfect imitation. First, they search among their contacts in the network and select m solutions of their neighbors randomly. Next, they have to decide which one is chosen from the m solutions. There are three different strategies available for them, that is, the best strategy (selecting the one with the highest payoff), the conformity strategy (selecting the one used by the majority of neighbors), and the random strategy (selecting one randomly). Note that individuals can use only one strategy in each experiment. After that, one individual observes whether the payoff of the selected solution is higher than itself. If yes, the individual imitates the solution and obtains a new solution fused by its solution and the socially acquired solution; otherwise, the individual holds its original solution. The fusion rule for generating a new solution is as follows (Crawford, 1995; Friedkin and Johnsen, 1990):

$$s_i(t+1) = \phi_i s_i(t) + (1 - \phi_i) s_j(t)$$

where ϕ_i is the SQB of individual i , $s_i(t)$ and $s_j(t)$ are the current solution of individual i and the individual to be imitated, $s_i(t+1)$ is the new solution that will be adopted by individual i at next step. It can be seen that conservative individuals' new solutions will be closer to their original solutions, while open individuals will be biased towards others' solutions.

Simulation procedure. Each simulation experiment consists of the following steps. First, we generate a group of N individuals with a specific network structure and a specific SQB distribution. Then, we assign to this group one task with a specific complexity C and a specific granularity ω . After all individuals in this group are randomly assigned an initial solution, they select one strategy

to find new solutions with the above social learning mechanism. We record the average payoff of all individuals as the group performance on each step, and these steps are repeated 500 times. We perform multiple experiments for each combination of task features, group features, and strategy-related variables. Here, we construct a basic combination where the group consisting of the FC network has $N = 100$ individuals with the normal SQB distribution; the task environment is a hard-to-easy task with $C = 5$ and $\omega = 4$; individuals learn 500 steps using the best strategy with sample size $m = 3$. A group can achieve optimal performance under the basic combination. Unless otherwise specified, we use this basic combination to study the effect of a variable by adjusting only its value. The final results reported are averaged across ten repetitions.

Decision trees. We use two popular decision tree algorithms to evaluate the effect of one variable on group performance, that is, Classification and Regression Tree (CART) (Li et al., 1984) and Random Forest (RF) (Breiman, 2001). CART is a classic and well-explained algorithm, while RF is an ensemble learning algorithm that combines the output of multiple decision trees to reach a single result and thus has better prediction accuracy. In particular, we use CART and RF to build machine learning models $y = f(X)$ to predict the group performance y given the values of input variables X . Here, the task granularity, task types, task complexity, group conservativeness, and social learning strategies are selected for the input variables of this model. We use the popular machine learning library based in Python, *scikit-learn* (Decision Trees—Scikit-Learn 1.1.0 Documentation, 2022), to implement CART and RF. Since the *scikit-learn* implementation does not support categorical variables for now, we employ one-hot encoding to convert categorical variables such as task types and social learning strategies into binary variables. To simplify the representation in the main text, we use T to represent task types ($T = 0$ for the hard-to-easy task; $T = 1$ for the easy-to-hard task), and S to represent social learning strategies ($S = 0$ for the best strategy; $S = 1$ for the conformity strategy; $S = 2$ for the random strategy).

We consider all 7560 parameter combinations of the above five variables: $T = 0, 1$; $\omega = 1, 2, \dots, 10$; $C = 1, 2, \dots, 14$; $\rho = 10, 20, \dots, 90$; $S = 0, 1, 2$. We perform 10 simulation experiments for each parameter combination and record the final group performance y . Then, we obtain the full data consisting of 75,600 samples. We use 80% of the data as the training set to train decision tree models and the remaining 20% as the test set to validate the accuracy of the models. We chose mean square error (MSE) as the criterion to measure the quality of a split. MSE and root MSE (rMSE) are used to evaluate the prediction performance of the models on the test set. The prediction performance of decision tree models improves as the depth of the tree increases (Supplementary Table 2). Here, the depth of a tree refers to the maximum distance between the root node and any leaf node. A decision tree with a large depth may lead to overfitting the data. To avoid this problem, we stop splitting when the reduction in rMSE is below 0.01 and then obtain a four-layer binary regression tree, as shown in Fig. 5a. Note that the results for CART and RF are similar on our simulation data. The prediction error for RF is only slightly lower than that for CART, but they generate the same tree (see Supplementary Figs. 20 and 21).

A great advantage of decision trees is their ability to evaluate the importance of variables to the model. In a decision tree model, the importance of a variable is related to two indicators. The first is the MSE, which measures the predictive accuracy of a variable on the group performance. A lower MSE indicates that this variable has good predictive performance, and thus it has a

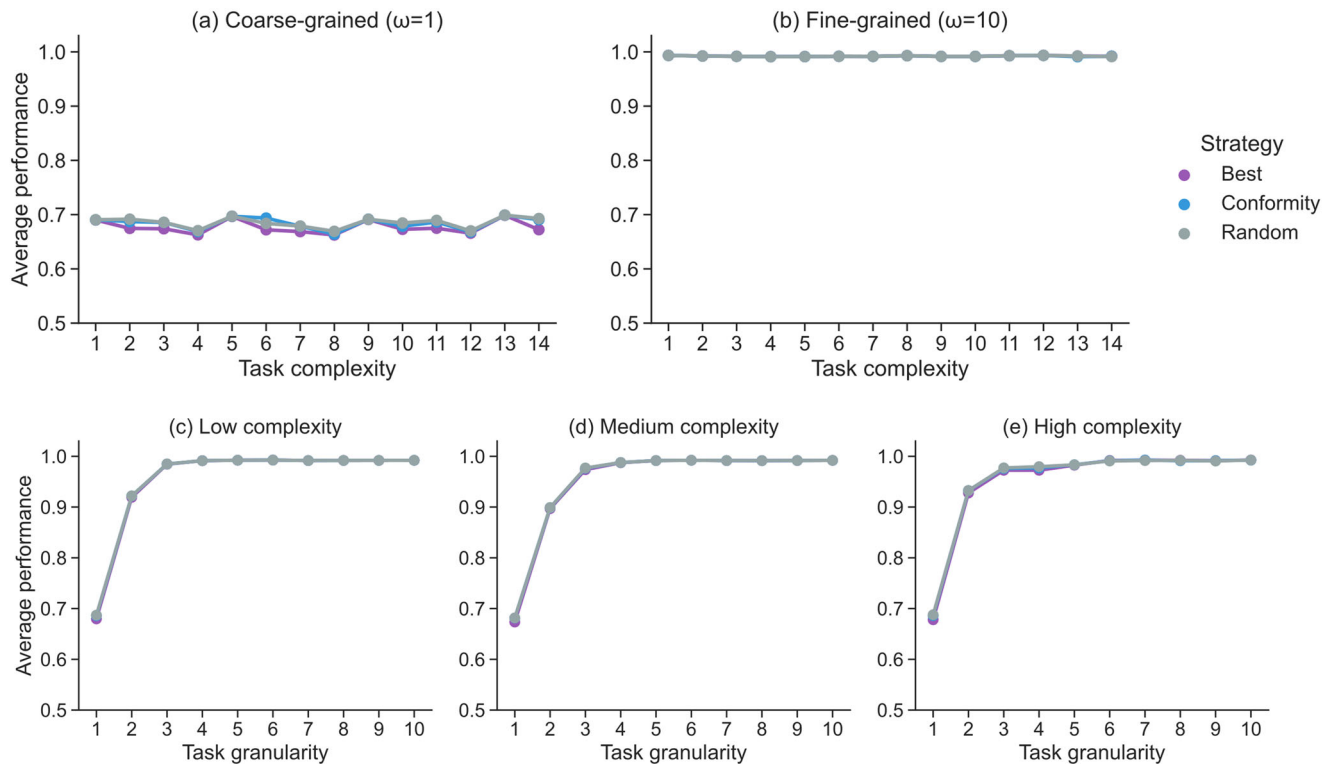


Fig. 2 Performance of groups using different strategies in hard-to-easy tasks. Two panels in the first row show the results in **a** coarse-grained and **b** fine-grained hard-to-easy tasks with different complexity. Three panels in the second row show the results in hard-to-easy tasks with **c** low, **d** medium, and **e** high complexity. Purple lines: the best strategy. Blue lines: the conformity strategy. Gray lines: the random strategy. Based on a fully connected network with $N=100$ individuals and the normal SQB distribution.

higher importance score. The second one is the sample size of the node where this variable is located. The greater the number of samples that rely on this variable for prediction, the more important this variable is. CART and RF can output the importance scores of all variables automatically, and the higher the score, the more important the variable. The importance scores of all variables sum to 1. The results for CART and RF are similar on our simulation data (see Supplementary Tables 2 and 3).

Results

The effects of task features. Figures 2 and 3 show the average performance achieved by each strategy for two distinct tasks: a hard-to-easy one (Fig. 2) and an easy-to-hard one (Fig. 3). We first study the role of task complexity on group performance under the coarsest granularity ($\omega=1$) and finest granularity ($\omega=10$). Next, we investigate the effect of reducing granularity on group performance. Here we focus on the average performance of different social learning strategies (with sample size $m=3$) in a fully connected network with normal SQB distribution.

Performance in hard-to-easy tasks. In the hard-to-easy tasks, all strategies perform equally well since the average performance achieved by each strategy reaches 1 across complexities (Fig. 2a, b). The result suggests that no matter how complex the hard-to-easy task is, a group can find the global optimum by any of the social learning strategies. This is mainly due to the characteristics of the hard-to-easy task itself. The hard-to-easy task is modeled by the Ackley function, in which the trend term plays a larger role than the fluctuation term (Supplementary Note 1). The increase in complexity will not lead to a great change in the landscape near the global optimum (Fig. 1a and Supplementary Fig. 2). Consequently,

as an individual's solution gets closer to the optimal solution, it is more influenced by the trend term and less likely to get stuck in local optima. Once individuals come out of the initial hard learning phase in this type of task, they pick up later and make great progress very quickly. As a result, as soon as one individual in the group finds the optimal solution, the other individuals can quickly follow, regardless of the social learning strategy.

As for the task granularity, Fig. 2c–e show that groups relying on social learning have difficulty in coarse-grained tasks ($\omega < 3$) since the performance of all strategies drops significantly when the task granularity decreases below 3. This occurs because the number of solutions becomes small, and the variation between solutions becomes large in coarse-grained tasks, making it more difficult for individuals to learn new solutions. As a result, many individuals fail to keep up with better solutions and stay at a low level of exploration. This phenomenon is evident in conservative individuals who stop learning prematurely and do not accept any new solution. For example, the average performance of conservative individuals stays below 0.4 after step 10 and cannot be improved in hard-to-easy tasks ($\omega=1$) (see Supplementary Fig. 7). We call it learning bottleneck and explain this phenomenon in detail with examples (see Supplementary Note 2).

Performance in easy-to-hard tasks. In fine-grained easy-to-hard tasks, these strategies perform well in low-complexity tasks ($1 \leq C \leq 5$), but as the complexity increases ($C > 5$), they can no longer maintain the original performance (Fig. 3b). However, group performance fluctuates as the increase of complexity in coarse-grained tasks (Fig. 3a). As the easy-to-hard task is generated by the Rastrigin function where the fluctuation term plays a larger role than the trend term in the surroundings of the global optimum (Fig. 1b). In the fine-grained case, numerous peaks of

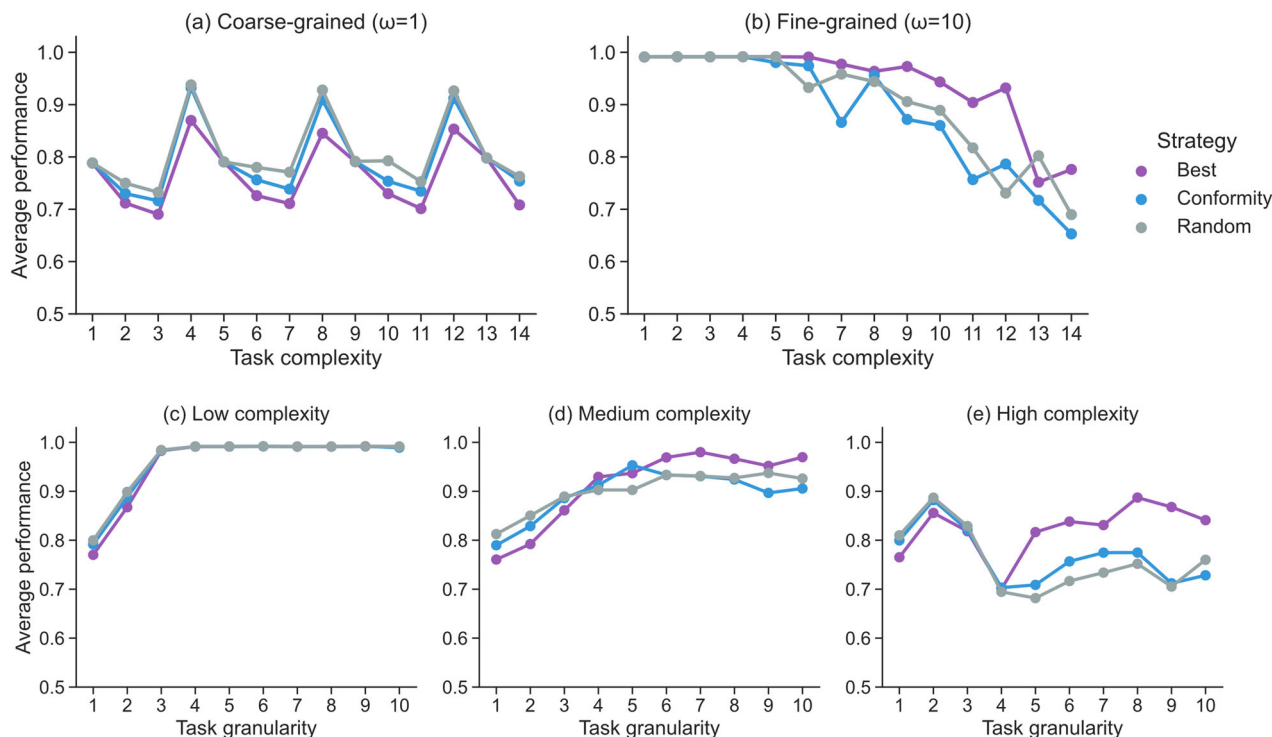


Fig. 3 Performance of groups using different strategies in easy-to-hard tasks. Two panels in the first row show the results in **a** coarse-grained and **b** fine-grained easy-to-hard tasks with different complexity. Three panels in the second row show the results in easy-to-hard tasks with **c** low, **d** medium, and **e** high complexity. Purple lines: the best strategy. Blue lines: the conformity strategy. Gray lines: the random strategy. Based on a fully connected network with $N = 100$ individuals and the normal SQB distribution.

this Rastrigin function are represented in the easy-to-hard task, so individuals face a greater risk of trapping in locally optimal solutions as the complexity increases. In the coarse-grained case (there are rare solutions), the environments of the easy-to-hard tasks with different complexity vary greatly (see Supplementary Fig. 3), leading to the fluctuation of group performance. The above results are in line with the characteristics of the easy-to-hard tasks and also highlight the role of task types. Interestingly, strategies differ in the reduction of group performance in high-complexity tasks, and the best strategy outperforms the other two strategies in fine-grained tasks. This result shows that individuals relying on the best strategy can perform better across various tasks (Zheng et al., 2020).

As shown in Fig. 3c–e, reducing granularity also decreases the performance of all strategies in easy-to-hard tasks of low and medium complexity. However, this reduction in granularity shows a different pattern in easy-to-hard tasks of high complexity. Groups perform a little better in coarse-grained tasks than in fine-grained tasks. First, as mentioned above, there are many local optima near the global optimum in high-complexity tasks. Conservative individuals tend to fall into local optima, making the group perform poorly. In particular, the number of solutions in fine-grained tasks becomes larger, further increasing the number of local optima near the global optimum, making it more difficult for conservative individuals to explore the solution space. The two reasons lead to poorer group performance in fine-grained tasks of high complexity. Second, the decline in granularity greatly reduces the number of solutions in the task environment and decreases local optima near the global optimum. This reduction of local optima, in turn, alleviates the interference of individual exploration and increases the conservative individuals’ chances of finding the globally optimal solution. Thus, we see a slight improvement in the group performance in coarse-grained tasks.

Group conservativeness and task granularity interact. We have seen that different task features can lead to remarkably different performances within the same group. Here, we focus on the effect of various group features, including the group size, network structure, network density, and group conservativeness. Numerous experiments show that only group conservativeness significantly affects group performance (Supplementary Note 4). In both types of tasks, the more conservative individuals in a group, the worse the average performance of the group (Supplementary Fig. 6). Conservative individuals hit the learning bottleneck because they tend to maintain the status quo rather than follow new solutions, leading to inferior performance (Supplementary Note 2). This fact can be seen in Fig. 4 and Supplementary Fig. 9, which show that open individuals always find the optimal solution and perform better than conservative individuals.

However, Fig. 4 also suggests that increasing task granularity can reverse this disadvantage for the conservative-biased group. This occurs because conservative individuals have better performance in fine-grained tasks than coarse-grained ones. In particular, the number of solutions becomes large, and the difference between solutions becomes small in fine-grained tasks. Conservative individuals can make little progress by accepting solutions that differ less from their solutions in each step. Therefore, they can approach what open individuals achieve in a short time through long-run slow changes (Supplementary Figs. 7 and 8). A rigorous mathematical analysis also demonstrated that task granularity and group conservativeness interact to influence the performance of individuals exploring new solutions (Supplementary Note 3).

Feature importance. Controlled simulation experiments allow us to quantify the effect of each variable on group performance. Ultimately, we identify five variables affecting group performance

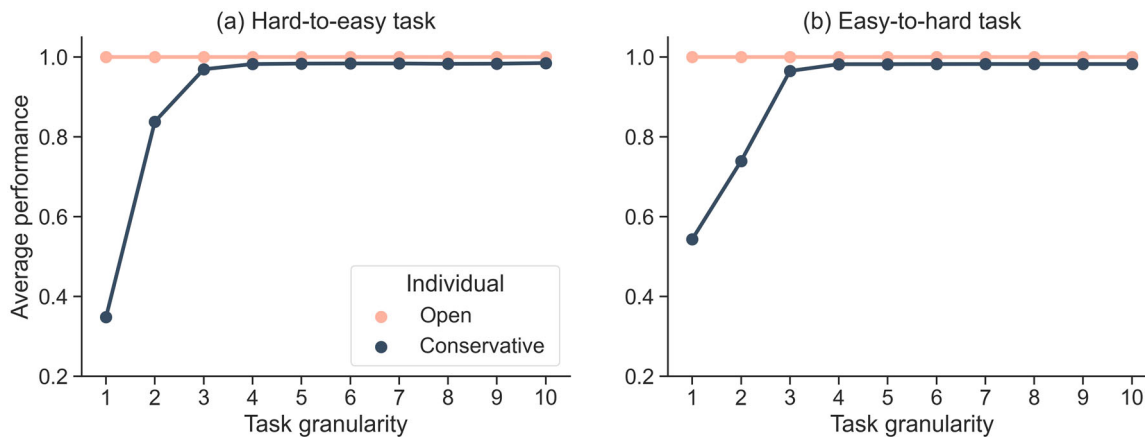


Fig. 4 Performance of open and conservative individuals in two tasks. Average performance of open and conservative individuals in **a** hard-to-easy task and **b** easy-to-hard tasks under different granularity. All individuals use the best strategy with sample size $m = 3$. Based on a fully connected network with $N = 100$ individuals.

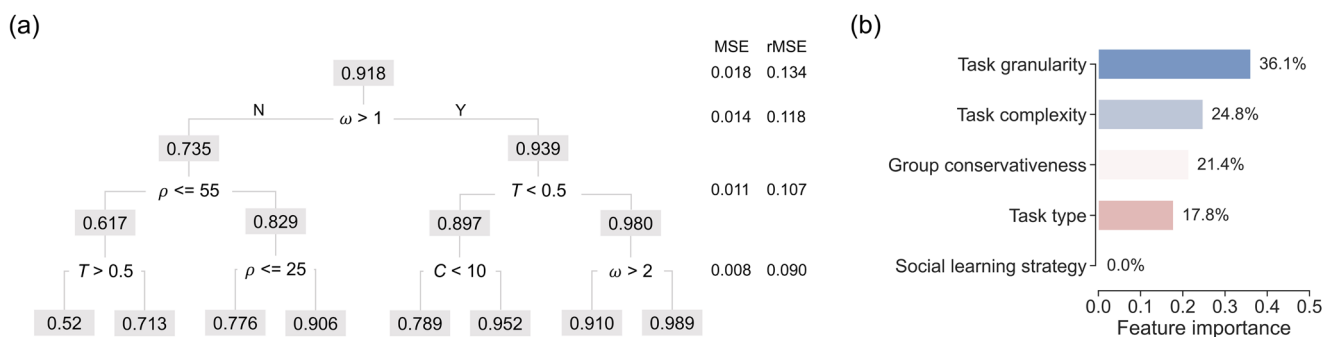


Fig. 5 The results of decision tree models learned on the simulation data. **a** A four-layer binary regression tree given by CART. Each gray box with the below conditional statement is a root node; these gray boxes without conditional statements are leaf nodes. The number in the gray box denotes the predicted value of group performance, and the conditional statement below the gray box represents the split point on a variable. N (the left of the conditional statement) and Y (the right of the conditional statement) represent No and Yes, respectively. For simplicity, N and Y are omitted next to other conditional statements. **b** Feature importance scores given by Random Forest.

significantly: three task features, group conservativeness, and social learning strategies (Supplementary Note 4). Which of these variables has the greatest effect? Simulation experiments cannot tell us this result. Therefore, we use CART and RF, commonly used decision tree algorithms, to evaluate the importance of the above five variables. In particular, CART and RF are used to learn a model that predicts the group performance given the values of the five variables. The learned model is represented as a binary tree where each root node represents a single input variable and a split point on that variable with a prediction of the target value. Besides, this tree model gives the importance score of all input variables.

Figure 5a shows the binary regression tree learned by CART on the training data. Note that RF generates the same tree as CART (see Supplementary Figs. 20 and 21). In Fig. 5a, the number in the gray box denotes the predicted value of group performance, and the conditional statement below the gray box represents the split point on a variable. These predictions use the average value \bar{y} within each subset, which is selected to minimize the mean square error, $MSE = \sum_i (\bar{y} - y_i)^2/n$. For example, the group performance can be trivially estimated by its average across the full training sample, $\bar{y} = 0.918$ with $MSE = 0.018$. The MSE can be lowered by splitting the sample into two subsets based on the value of one input variable and minimizing $n_1MSE_1/n + n_2MSE_2/n$, the weighted average of MSE of both subsets. Doing this, we find that task granularity is chosen as the variable for the first

split, and the boundary at $\omega = 1$ yields the lowest MSE value. This split forms the first branching in the binary tree, and the newborn left and right nodes below 0.918 represent two heterogeneous subsets with homogeneous values of the group performance inside. New predictions are made by the average value of the group performance of the sample in each subset. For instance, the predicted value of group performance for the left subset ($\omega \leq 1$) is 0.735 and for the right subset ($\omega > 1$) 0.939. This result indicates that task granularity greatly affects group performance, and groups perform poorly in coarse-grained tasks, consistent with the previous findings.

Once the first split is chosen, each of the two subsets is split again using the same approach, and the process continues iteratively. As we split the regression tree, we get more nodes as well as a lower MSE (with relative rMSE). To avoid overfitting the data, we stop splitting when the reduction in rMSE is below 0.01 and then obtain a four-layer binary regression tree as shown in Fig. 5a (see Supplementary Fig. 17 for the full tree; Supplementary Table 2 for the results of other depths). The tree goes through a total of seven splits. Among them, task granularity is chosen for two splits ($\omega > 1$ and $\omega > 2$), task type for two splits ($T < 0.5$ and $T > 0.5$), group conservativeness also for two splits ($\rho \leq 55$ and $\rho \leq 25$), and task complexity for once ($C \leq 10$).

Figure 5b shows the importance score of the five input variables given by RF. As mentioned above, the results for CART and RF are similar on our simulation data (Supplementary Tables

2 and 3). We obtain two striking results from Fig. 5b. First, the three task features together account for 78.6% of the importance, with task granularity alone contributing 36.1% of the importance (24.7% for task complexity and 17.8% for task type). This result highlights the role of tasks in social learning. Especially, task granularity, a previously neglected variable, plays the largest role. Second, group conservativeness shares the remaining 21.4% of importance though its effect is significant only in the coarse-grained tasks (Fig. 4). Surprisingly, the importance of the social learning strategy is 0 since it is absent from the regression tree (Fig. 5a). It is worth noting that this result does not imply that social learning strategies have no effect on group performance. If we focus on the group performance in the high-complexity easy-to-hard task (Fig. 3e), the three social learning strategies show a slight difference. The result only suggests that the effect of social learning strategies on group performance is negligible compared to task features and group conservativeness.

Discussion

We have figured out three questions in this paper. First, what is the role of tasks in social learning? Our work provides a detailed analysis of task features. Previous studies on social learning have either ignored task types or confused task types with task complexity (Almaatouq et al., 2021; Barkoczi et al., 2016; Barkoczi and Galesic, 2016). For example, these studies dichotomized tasks into simple and complex ones by complexity (Zheng et al., 2020; Flynn et al., 2016; McElreath et al., 2005). We distinguished them for the first time and found the interaction between task types and task complexity. Generally, groups relying on social learning can achieve excellent performance in low-complexity tasks; however, they may fail to do this in high-complexity tasks due to different types of tasks. In the easy-to-hard tasks, the local environment near the globally optimal solution changes drastically, which means that there are many locally optimal solutions, leading to an increased difficulty for individuals to find the global optimum. Boosting the complexity can worsen this problem and bring down the group performance.

Disentangling task complexity from task types helps us understand many common learning activities in cultural evolution. For example, learning to ride a bike or drive a car can be regarded as a proxy for hard-to-easy tasks (Wulff et al., 2009). Such tasks are associated with procedural memory (Sun et al., 2001), which is difficult to acquire at the beginning of learning. Once individuals acquire procedural memory through tons of practice, they do not easily forget it even if the complexity of the task changes. For instance, driving a car is more complicated than riding a bicycle, but this does not prevent most people from obtaining a driver's license. In contrast to the tasks above, learning mathematics can be considered an easy-to-hard task. It requires learners to follow a strict sequence because the content of advanced courses is often a combination of content from primary courses (Zeps, 2009). For example, students should learn simple algebra before calculus and master single-variable calculus before learning multivariable calculus. For such tasks, the low-complexity ones tend to be easier than the high-complexity ones.

In addition, we introduce task granularity, a variable that has been overlooked in social learning but received extensive attention in other fields (Janssen et al., 2015; Rosà et al., 2019; Tran et al., 2016). Task granularity determines the perceived granularity of a group to one task. We found that groups perform poorly on coarse-grained tasks (with a small number of solutions), while increasing task granularity can greatly improve group performance. The surge in the number of solutions allows individuals to fully explore the solution space, which leads to higher group performance. This finding is supported by many cases in the evolution of human culture or technology. During the Ming Dynasty in China

(1368–1644), a beautifully crafted porcelain came from the Jingdezhen official kiln, called underglaze red, whose coloring required the temperature of fire to be controlled between 1280 and 1300 degrees Celsius (Nanjing Museum, 2022). This requirement was beyond the capacity of the artisans of the time, so there were very few qualified products. A recent example is that after the scientific and industrial revolution, Europeans were able to exactly calculate the proportions of the various fuels used in the steelmaking process through chemical analysis (Barraclough, 1990) (such as the Siemens-Martins method, born in 1856), which allowed the Europeans to surpass the Chinese who had been ahead for thousands of years. Another famous case is the Toshiba-Kongsberg incident, which also suggests the importance of granularity (i.e., precision) to manufacturing. Throughout the cold war, large-sized powers were looking to build nuclear-powered submarines that were quieter, and therefore more difficult to track and intercept. Soviet CNC (computer numerically controlled) machines could not meet the machining precision of cutting propellers, resulting in its submarines being consistently noisy. Subsequently, the Soviet government secretly imported highly advanced milling machines made by Toshiba because they cut smoother propellers, hence making the nuclear submarine propulsion quieter (Wrubel, 2011).

Second, how does individuals' status quo bias affect their performance in social learning? Our results show that conservative individuals are less likely to find the globally optimal solution due to their low preference for socially acquired solutions. This makes them perform inferior to open individuals who always achieve high-level performance. As a result, a group with more conservative individuals performs worse. However, this drawback of conservative individuals can be mitigated by increasing task granularity. The intuition underlying the result is the following. The proliferation of the number of solutions instead narrows the difference between solutions. Conservative individuals can accept solutions that are less different from their own and thus achieve slow progress. After a slow long-run effort, conservative individuals gradually catch up with open individuals. This finding also coincides with relevant studies in cognitive psychology. In these studies, fine-grained tasks are described as tasks with fast and immediate feedback (Epstein et al., 2002). Some individuals who have poor performance or negative emotions (e.g., resistance to learning) will achieve better learning outcomes if they receive immediate and positive feedback from their teachers (Muis et al., 2015).

Third, how do we quantify the effects of different variables on group performance? Few studies focused on this problem. On the one hand, there are so many variables that affect the performance of a group relying on social learning. It is difficult and expensive to conduct behavioral experiments to measure these variables' effects. On the other hand, previous studies and our work showed the interactions between variables, which pose great challenges in disentangling their effects. In this paper, we presented a method to comprehensively examine the effects of different factors on group performance in social learning. This approach consists of two phases. In the first stage, we conducted a simulation experiment to examine the effect of a single variable by changing its value and then observe the significance of its effect on group performance. Five variables that significantly impact group performance are figured out: three task features, group conservatism, and social learning strategies. Then, we used decision tree models to quantify the predictive performance of these variables on group performance. The feature importance scores given by the models show the impact of these variables. Our results suggest that task features have the highest influence on group performance, at over 78%, which reveals the decisive role of tasks in social learning. In particular, our method has two distinct advantages. One is brought by the agent-based simulation model, which gives good

reproducibility to the experimental results. The other is that decision tree models offer an opportunity to estimate each variable's importance. Computational sociologists can benefit a lot from this machine learning approach.

Notably, we obtained some results that are different from previous studies. For example, we found that group-related variables, such as network structure (Barkoczi and Galesic, 2016; Lazer and Friedman, 2007), network density (Shi et al., 2017), and network size (Ausloos, 2015), had low impacts on group performance. One possible reason is that previous studies did not take task granularity into account. In their experiments, the task granularity was generally designed to be low, i.e., the number of solutions available to individuals was small, which limited the group performance. They may have misattributed the observed changes in group performance to other variables. In addition, all three strategies achieved high performance in most cases, with little difference between them. This may be because we regard social learning as imperfect imitation, which allows individuals to explore new solutions by a trade-off between their solutions and others' solutions (Posen and Martignoni, 2018). This differs from studies (Fang et al., 2010; Lazer and Friedman, 2007) in which individuals used two learning patterns simultaneously: individual learning for exploring new solutions and social learning for disseminating them. However, it is hard to distinguish the roles of the two learning modes in these studies.

It is worth mentioning that our results do not negate the role of social learning strategies but show that there is little difference between these strategies. Frankly, individuals can improve their payoffs as long as they consistently learn from individuals who are better than them; it is less important which strategy is used to choose these individuals. It is undeniable that one strategy may outperform others in some cases. For example, the best strategy is better than the other two strategies in fine-grained easy-to-hard tasks of high complexity. This phenomenon is particularly evident in higher education, where top scientists tend to promote the rapid development of the discipline and identify and develop outstanding talent. For example, Leonhard Euler, the Swiss mathematician recognized as the most brilliant in the 18th century, made considerably high-level research in Saint Petersburg. He played a huge role in advancing the development of Russian mathematics and got the highest possible recognition from Russia (Schulze, 1985).

Our study shows that social learning enables groups to achieve high-level performance (Posen et al., 2013). However, individuals can also learn from their own trial-and-error experience, an important component of human decision-making. In this study, we assume that individuals rely only on social learning without independent exploration, which is a limitation of the present study. Future work can explore the mixture of the two different learning styles and examine which plays a greater role in human decision-making (Zhang and Gläscher, 2020). In addition, the tasks studied in this paper have only one independent variable, whereas real-world tasks may include multiple independent variables. Future research can investigate the effect of task dimensionality on group performance.

Taken together, our results provide new insights into social learning. First, the result that task features contribute nearly 80% to predicting group performance suggests that tasks almost determine the performance a group can achieve via social learning. Second, increasing task granularity helps conservative individuals avoid hitting the learning bottleneck, which means that groups can achieve the performance they otherwise would not by changing task features. These results have broad implications for cultural transmission and evolution. For example, empirical studies by cultural anthropologists have shown that new solutions with high payoffs are resisted by conservatives when they are disseminated within groups (Rogers, 2010). Promoters of these new solutions typically employ

psychological methods to persuade conservative individuals, although these methods are laborious. Our work offers an alternative approach. Promoters preferably split the solution they spread and recommend only a portion of them to conservative individuals at a time. In this way, conservative individuals will not show a strong tendency to resist and will gradually accept the new solution.

Data availability

Code and data are available in an open repository (<https://github.com/networkanddatasciencelab/social-learning>).

Received: 25 January 2022; Accepted: 5 August 2022;

Published online: 20 August 2022

References

- Acerbi A, Tennie C, Mesoudi A (2016) Social learning solves the problem of narrow-peaked search landscapes: experimental evidence in humans. *R Soc Open Sci* 3(9):160215
- Almaatouq A, Alsobay M, Yin M (2021) Task complexity moderates group synergy. *Proc Natl Acad Sci USA* 118(36):e2101062118
- Ausloos M (2015) Slow-down or speed-up of inter- and intra-cluster diffusion of controversial knowledge in stubborn communities based on a small world network. *Front Phys* 3:43
- Azzalini A, Capitanio A (1999) Statistical applications of the multivariate skew normal distribution. *J R Stat Soc Ser B: Stat Methodol* 61(3):579–602
- Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512
- Barkoczi D, Analytis PP, Wu C (2016) Collective search on rugged landscapes: a cross-environmental analysis. In: *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Cognitive Science Society. pp. 918–923
- Barkoczi D, Galesic M (2016) Social learning strategies modify the effect of network structure on group performance. *Nat Commun* 7(1):1–8
- Boyd R, Henrich J (2002) On modeling cognition and culture: why cultural evolution does not require replication of representations. *J Cogn* 2(2):87–112
- Boyd R, Richerson PJ, Henrich J (2011) The cultural niche: why social learning is essential for human adaptation. *Proc Natl Acad Sci USA* 108(Suppl 2):10918–10925
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
- Cantor M, Shoemaker LG, Cabral RB, Flores CO, Varga M, Whitehead H (2015) Multilevel animal societies can emerge from cultural transmission. *Nat Commun* 6(1):1–10
- Crawford V (1995) Adaptive dynamics in coordination games. *Econometrica* 63(1):103–143
- Csaszar F, Science NS (2010) How much to copy? Determinants of effective imitation breadth. *Organiz Sci* 21(3):661–676
- Scikit-learn (2022) Decision Trees—scikit-learn 1.1.0 documentation. Retrieved May 19, 2022, from <https://scikit-learn.org/stable/modules/tree.html>
- Derex M, Boyd R (2016) Partial connectivity increases cultural accumulation within groups. *Proc Natl Acad Sci USA* 113(11):2982–2987
- Derex M, Feron R, Godelle B, Raymond M (2015) Social learning and the replication process: An experimental investigation. *Proc R Soc B: Biol Sci* 282(1808):20150719
- Dietrich E, Markman AB (2003) Discrete thoughts: Why cognition must use discrete representations. *Mind Lang* 18(1):95–119
- Epstein ML, Lazarus AD, Calvano TB, Matthews KA, Hendel RA, Epstein BB, Brosvic GM (2002) Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *Psychol Record* 52(2):187–201
- Ethiraj SK, Levinthal D, Roy RR (2008) The dual role of modularity: Innovation and imitation. *Manag Sci* 54(5):939–955
- Fang C, Lee J, Schilling MA (2010) Balancing exploration and exploitation through structural design: the isolation of subgroups and organizational learning. *Organiz Sci* 21(3):625–642
- Flinn MV (1997) Culture and the evolution of social learning. *Evol Hum Behav* 18(1):23–67
- Flynn E, Turner C, Giraldeau LA (2016) Selectivity in social and asocial learning: Investigating the prevalence, effect and development of young children's learning preferences. *Philos Trans R Soc B: Biol Sci* 371(1690):20150189
- Fogarty L, Rendell L, Laland KN (2012) Mental time travel, memory and the social learning strategies tournament. *Learn Motiv* 43(4):241–246
- Friedkin NE, Johnsen EC (1990) Social influence and opinions. *J Math Sociology* 15(3–4):193–206

- Haviland W, Prins H, McBride B (2013) *Anthropology: The human challenge*. Cengage Learning, Boston
- Janssen F, Grossman P, Westbroek H (2015) Facilitating decomposition and recomposition in practice-based teacher education: The power of modularity. *Teach Teach Educ* 51:137–146
- Jordà S (2004) Digital instruments and players: part I—efficiency and apprenticeship. In: *Proceedings of the 2004 Conference on New Interfaces for Musical Expression*. National University of Singapore. pp. 59–63
- Barracough K (1990) *Steelmaking 1850–1900*. Woodhead Pub Limited, Sawston
- Kauffman S, Levin S (1987) Towards a general theory of adaptive walks on rugged landscapes. *J Theor Biol* 128(1):11–45
- Kendal RL, Boogert NJ, Rendell L, Webster M, Jones PL (2018) Social learning strategies: Bridge-building between fields. *Trend Cogn Sci* 22(7):651–665
- Kurt S, Ehret G (2010) Auditory discrimination learning and knowledge transfer in mice depends on task difficulty. *Proc Natl Acad Sci USA* 107(18):8481–8485
- Laland KN (2004) Social learning strategies. *Anim Learn Behav* 32(1):4–14
- Lamberson PJ (2010) Social learning in social networks. *B.E. J Theor Econ* 10(1):36
- Lazer D, Friedman A (2007) The network structure of exploration and exploitation. *Admin Sci Q* 52(4):667–694
- Levinthal DA (1997) Adaptation on rugged landscapes. *Manag Sci* 43(7):934–950
- Li B, Friedman J, Olshen R, Stone C (1984) Classification and regression trees (CART). *Biometrics* 40(3):358–361
- Lieberman MB (1987) The learning curve, diffusion, and competitive strategy. *Strateg Manag J* 8(5):441–452
- Mason W, Watts DJ (2012) Collaborative learning in networks. *Proc Natl Acad Sci USA* 109(3):764–769
- McElreath R, Lubell M, Richerson PJ, Waring TM, Baum W, Edsten E, Efferson C, Paciotti B (2005) Applying evolutionary models to the laboratory study of social learning. *Evol Hum Behav* 26(6):483–508
- Mesoudi A (2008) An experimental simulation of the “copy-successful-individuals” cultural learning strategy: adaptive landscapes, producer–scrounger dynamics, and informational. *Evol Hum Behav* 29(5):350–363
- Mesoudi Alex, Thornton A (2018) What is cumulative cultural evolution? *Proc R Soc B: Biol Sci* 285(1880):20180712
- Molleman L, Van den Berg P, Weissing FJ (2014) Consistent individual differences in human social learning strategies. *Nat Commun* 5(1):1–9
- Moore BCJ (1974) Relation between the critical bandwidth and the frequency-difference limen. *J Acoust Soc Am* 55(2):359
- Morgan TJH, Rendell LE, Ehn M, Hoppitt W, Laland KN (2012) The evolutionary basis of human social learning. *Proc R Soc B: Biol Sci* 279(1729):653–662
- Morin O, Jacquet PO, Vaesen K, Acerbi A, Morin O, Jacquet PO, Vaesen K, Acerbi A (2021) Social information use and social information waste. *Philos Trans R Soc B* 376(1828):20200052
- Muis K, Ranellucci J, Trevors G, Duffy MC (2015) The effects of technology-mediated immediate feedback on kindergarten students’ attitudes, emotions, engagement and learning outcomes during literacy skills. *Learn Instruct* 38:1–13
- Nanjing Museum (2022). Jin State. Retrieved May 16, 2022, from <http://www.njmuseum.com/zh/generalDetails?id=556>
- Pil FK, Cohen SK (2006) Modularity: Implications for imitation, innovation, and sustained advantage. *Acad Manag Rev* 31(4):995–1011
- Posen HE, Lee J, Yi S (2013) The power of imperfect imitation. *Strateg Manag J* 34(2):149–164
- Posen Hart E, Martignoni D (2018) Revisiting the imitation assumption: why imitation may increase, rather than decrease, performance heterogeneity. *Strateg Manag J* 39(5):1350–1369
- Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, Ghirlanda S, Lillcrap T, Laland KN (2010) Why copy others? Insights from the social learning strategies tournament. *Science* 328(5975):208–213
- Rogers AR (1988) Does biology constrain culture? *Am Anthropol* 90(4):819–831
- Rogers E (2010) *Diffusion of innovations*. Simon and Schuster, New York, NY
- Rosà A, Rosales E, Binder W (2019) Analysis and optimization of task granularity on the Java virtual machine. *ACM Trans Prog Lang Syst* 41(3):1–47
- Schulze L (1985) The Russification of the St. Petersburg Academy of Sciences and Arts in the eighteenth century. *Br J Hist Sci* 18(3):305–335
- Shi Y, Zeng Y, Engo J, Han B, Li Y, Energy RM, Muehleisen RT (2017) Leveraging inter-firm influence in the diffusion of energy efficiency technologies: An agent-based model. *IEEE Trans Cybernetics* 48(6):1733–1746
- Shore J, Bernstein E, Lazer D (2015) Facts and figuring: An experimental investigation of network structure and performance in information and solution spaces. *Organ Sci* 26(5):1432–1446
- Sperber D (1996) *Explaining culture: a naturalistic approach*. Blackwell, Oxford
- Sun R, Merrill E, Peterson T (2001) From implicit skills to explicit knowledge: a bottom-up model of skill learning. *Cogn Sci* 25(2):203–244
- Tran B, Xue B, Zhang M (2016) A new representation in PSO for discretization-based feature selection. *IEEE Trans Cybernetics* 48(6):1733–1746
- Van Leeuwen EJ, Cohen E, Collier-Baker E, Rapold CJ, Schäfer M, Schütte S, Haun DB (2018) The development of human social learning across seven societies. *Nat Commun* 9(1):1–7
- Watts DJ, Strogatz SH (1998) Collective dynamics of ‘small-world’ networks. *Nature* 393(6684):440–442
- Wisdom TN, Song X, Goldstone RL (2013) Social Learning Strategies in Networked Groups. *Cogn Sci* 37(8):1383–1425
- Wrubel WA (2011) The Toshiba-Kongsberg incident: shortcomings of COCOM, and recommendations for increased effectiveness of export controls to the East Bloc. *Am Univ J Int Law Policy* 4(1):241–273
- Wulff P, Schonewille M, Renzi M, Viltono L, Sassoè-Pognetto M, Badura A, Gao Z, Hoebek FE, Van Dorp S, Wisden W, Farrant M, De Zeeuw CI (2009) Synaptic inhibition of Purkinje cells mediates consolidation of vestibulo-cerebellar motor learning. *Nat Neurosci* 12(8):1042–1049
- Zeps D (2009) Four levels of complexity in mathematics and physics: quantum distinctions. *CiteSeerX*
- Zhang L, Gläscher J (2020) A brain network supporting social influences in human decision-making. *Sci Adv* 6(34):eabb4159
- Zheng J, Guo N, Wagner A (2020) Selection enhances protein evolvability by increasing mutational robustness and foldability. *Science* 370(6521):eabb5962

Acknowledgements

This work was supported in part by the National Social Science Foundation of China (Grant No. 21BXW008), and in part by Science and Technology Commission of Shanghai Municipality (Grant No. 19JG0500700).

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1057/s41599-022-01305-2>.

Correspondence and requests for materials should be addressed to Yunlong Ma.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022