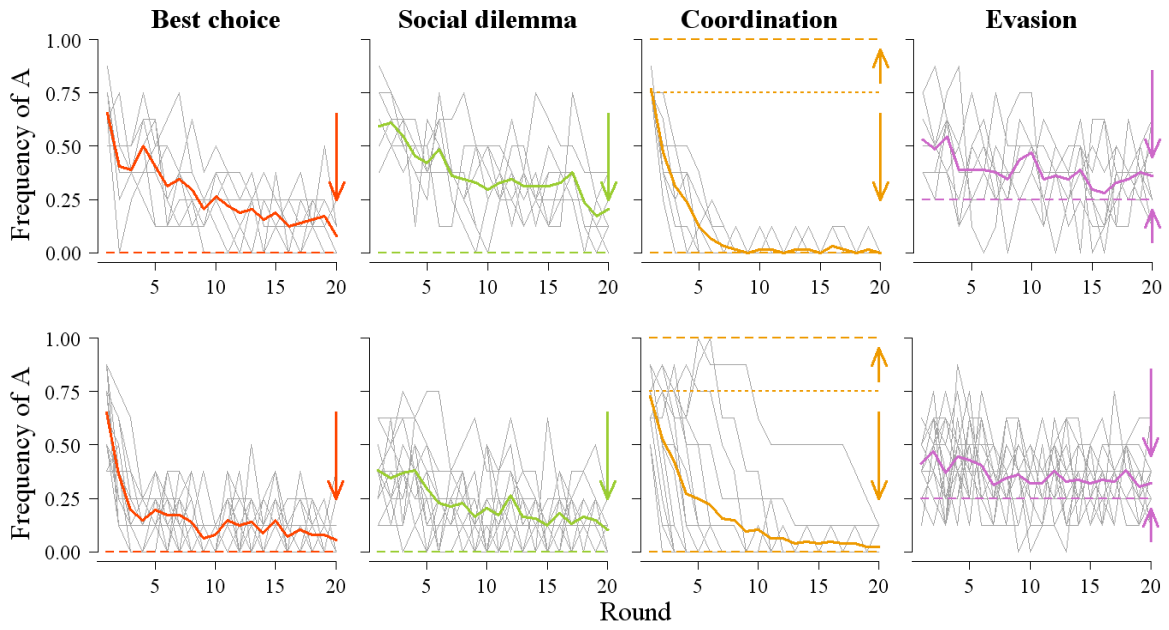
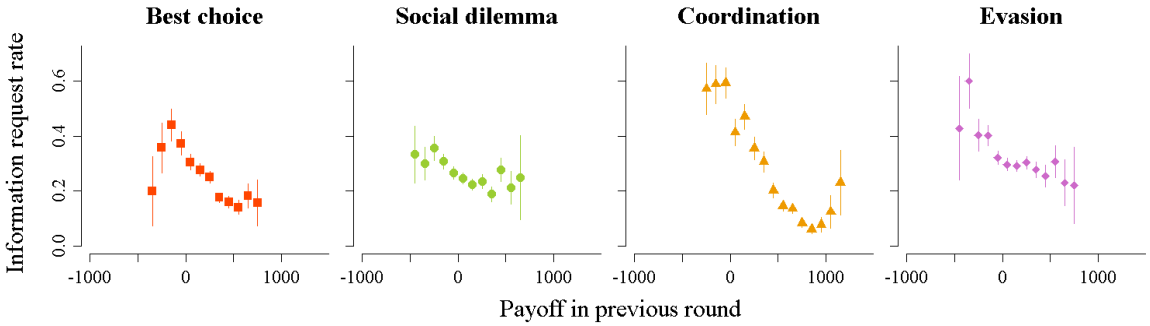
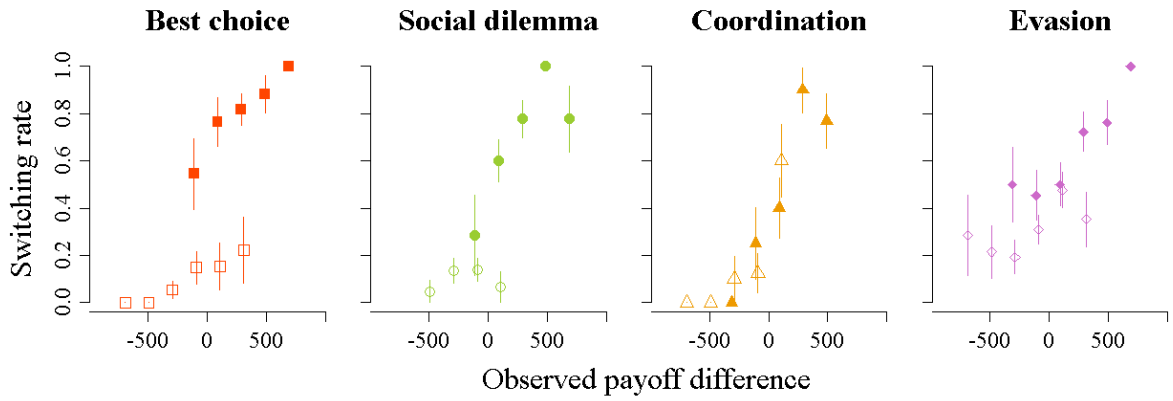
Supplementary Figures



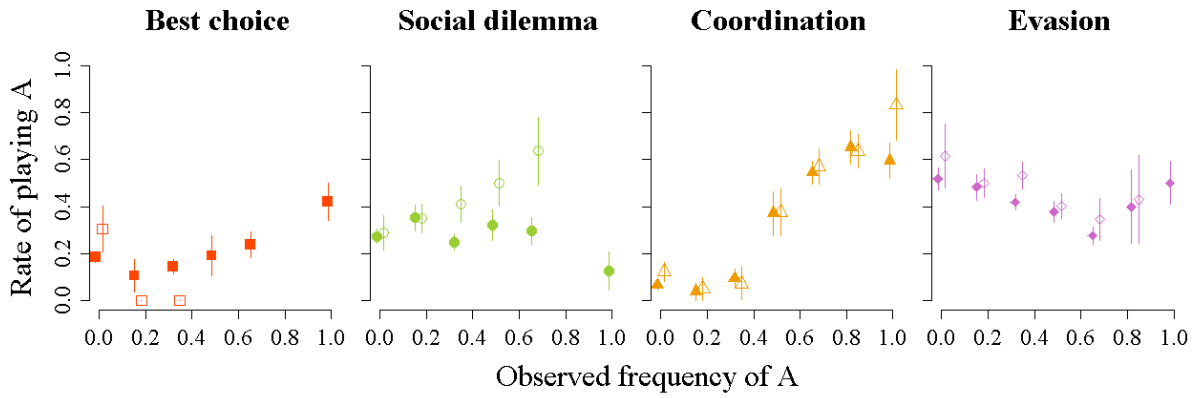
Supplementary Figure 1 | Decisions in the absence (top) and presence (bottom) of social information. To investigate how the presence of social information affected behaviour, we ran control sessions in which subjects could not collect information from their peers before making their decisions. Otherwise, the setup was identical to the setup of our primary experiment. We ran four independent sessions with sixteen subjects each (n=64), for a total of eight replicate groups of eight participants each. The top row of panels shows frequencies of choices for option *A* over twenty rounds of the four interaction in four control sessions, where social information was not available (eight replicate groups in grey, averages in colour). As a reference, the bottom row of panels shows the decision data from our primary experiment (where social information was available), also presented in Figure 1 of the main text. Over time, groups tended to approach the Nash equilibria of the one-shot version of the games (indicated by the arrows and the dashed lines). The presence of social information led to a more rapid convergence to the superior option in the best choice context and the dominant strategy in the social dilemma. In fact two separate binomial generalized linear mixed models (GLMM) with subject nested in group as random factor, and with 'information present/absent x round', 'previous decision x previous payoff' and 'difference between payoffs in round t–1 and t–2' as fixed factors, detected significant effects of the presence of information on the probability that an individual chooses option *B* (n=3840 decisions for each model; t-test, P<0.001 and P=0.047).



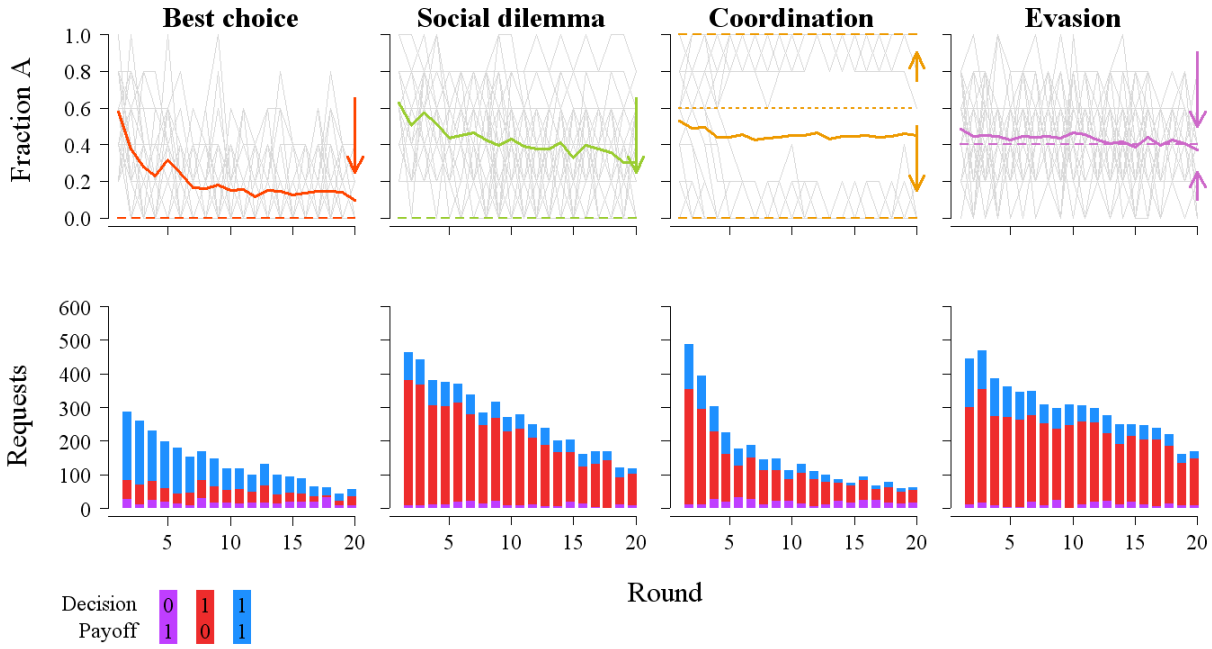
Supplementary Figure 2 | Information request rates as a function of individual payoff in the previous round. In line with earlier findings[30], information request rates were negatively related to previous payoffs (n= 9728, t-test, $P < 0.001$), indicating that subjects tended to request social information when their current behaviour had unfavourable returns. Moreover, information request times decreased over time (n= 9728, t-test, $P < 0.001$), confirming the observations from Figure 1 of the main text. Data on previous payoffs were pooled in cohorts of 100 points to obtain the symbols reflecting request rates. Error bars represent 1 standard error of the mean (SEM). Total number of observations per context was 2432. Cohorts with fewer than five observations are not shown as a data point. P-values are based on a binomial GLMM relating the request probability to previous payoff, using 'subject' as random factor, and 'round' and 'interaction context' as fixed factors.



Supplementary Figure 3 | Effects of payoff information on switching behaviour. Panels show proportions of decisions to switch behaviour as a function of the difference in own previous payoff and the observed average payoff of the alternative option. We consider only those decisions that were preceded by requests for others' previous decisions *and* previous payoffs. Open and solid symbols represent rates of switching (±1SEM) from *B* to *A* and from *A* to *B*, respectively. Payoff differences were pooled in cohorts of 200 points to obtain the data points shown (cohorts with fewer than five observations were omitted). In all cases, a subject's probability to switch behaviour is positively associated with the difference between the payoff of another player and the subject's own payoff. This conclusion, which is a clear indication of payoff-based learning, is based on four separate binomial GLMM (one per context) and an overarching GLMM. Each separate GLMM (where subject nested in group was treated as a random factor, while 'round' and 'previous decision x previous payoff x observed payoff difference' were entered as fixed factors) revealed a significant effect of the observed payoff difference on the probability of switching ($P<0.001$ in BC, CO and EV, and $P=0.023$ in SD). Similarly, the overarching GLMM (with the same error but 'interaction context' being added to the interaction terms) leads us to conclude that the observed payoff difference has a highly significant effect ($P<0.001$) on switching behaviour.



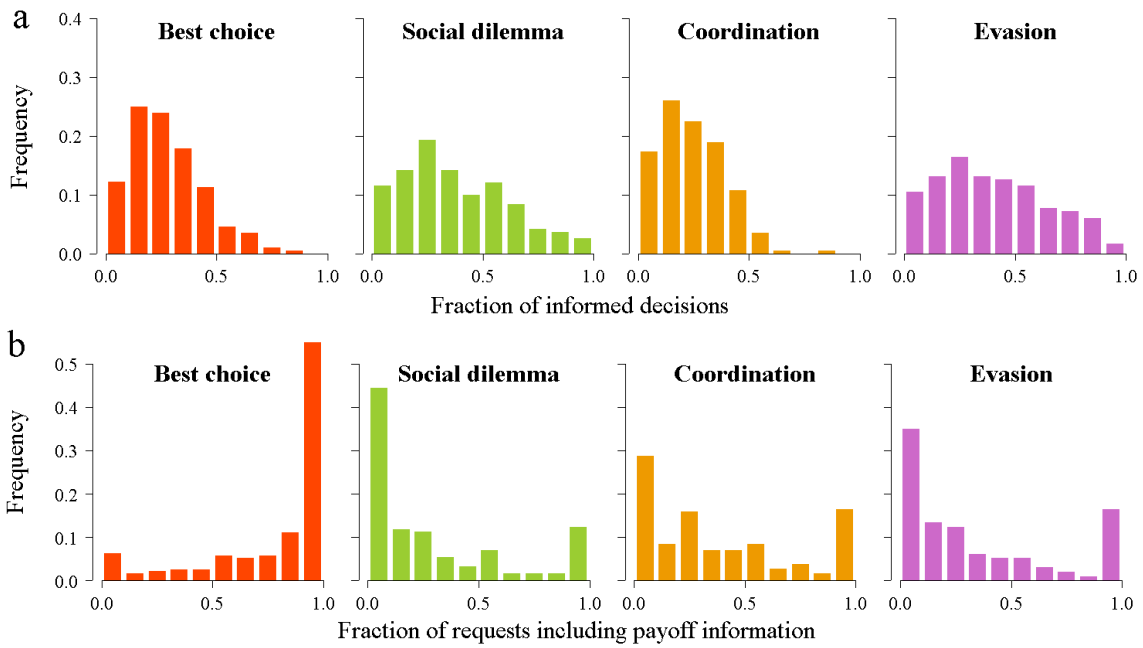
Supplementary Figure 4 | Effects of observing behaviour of group members on own behaviour. Panels show proportions of decisions (± 1SEM) for option *A* as a function of the observed frequency of *A*. Data points were pooled in cohorts of 1/6; cohorts with fewer than five observations are not shown. The solid symbols show the information requests where information about at least one decision of a group member was requested (so information about payoffs may have also been requested within the same information request). The open symbols show decisions that were preceded by requests purely for the decisions of fellow group members (so, no payoff information was requested). For each interaction context separately, we fitted binomial GLMMs to all decisions in which any decision information was requested (data corresponding to the solid symbols). In these regression models, we used 'subject nested in group' as random factors, and 'round', 'previous payoff' and 'previous decision x observed frequency of *A*' as fixed factors. In BC, SD and CO we find positive effects of the observed frequency of *A* on the probability of also choosing that option (*t*-tests: n=523, P=0.002; n=587, P=0.039; n=472, P<0.001, for BC, SD and CO, respectively). The GLMM fitted to the SD data detected a strong negative effect of 'round', which may explain why the slightly positive influence of observing *A* (cooperation) on the probability of also choosing *A* is not visible from the solid dots. In the evasion game, we find negative effects of the observed frequency of *A* on the probability to choose it (*t*-test: n=722, P=0.009), suggesting that subjects reacted to the underlying structure of the game. For SD, CO and EV, the findings of the models were confirmed by a GLMM fitted to data in which *only* decision information was requested (data corresponding to the open symbols). In these cases, we observe in the SD a stronger positive influence of the observed frequency of cooperation on the probability of an individual to also cooperate (*t*-test: n=175, P=0.002).



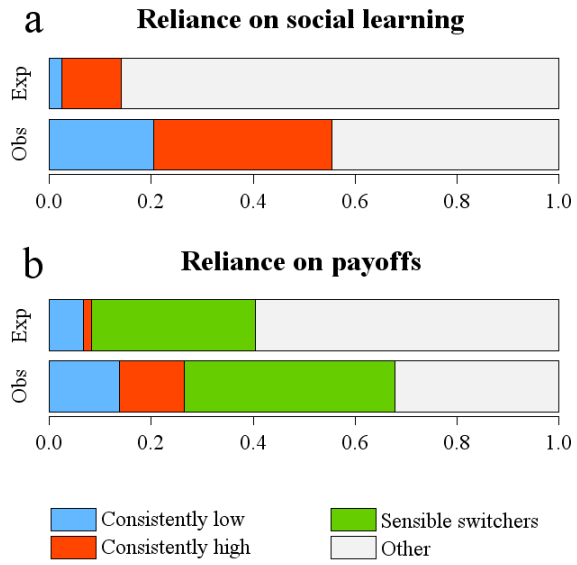
Supplementary Figure 5 | Follow-up experiment: decisions and information requests in four interaction contexts. This figure corresponds to Figure 1 in the main text, which presents the corresponding outcomes in the primary experiment.

The top row of panels shows the fraction of subjects choosing *A* over twenty rounds of the four contexts (40 replicate groups of five subjects in grey, averages in colour). Dashed coloured lines indicate Nash equilibria of the one-shot version of the four games. In the best choice context, dynamics were similar to the primary experiment. In the social dilemma, levels of cooperation tended to be higher than in the primary experiment, possibly due to the smaller group size. In the coordination game, a higher proportion of groups ended up at the equilibrium where all group members choose *A*, which is Pareto inferior to the equilibrium where all group members play *B* (*i.e.* all players would obtain higher payoffs in the latter equilibrium). This may be explained by the fact that the Pareto inferior equilibrium had a larger basin of attraction in the follow-up experiment, since the unstable internal equilibrium was closer to 0.5 than in the primary experiment. In the evasion game, groups were attracted to the internal equilibrium (as occurred in the primary experiment).

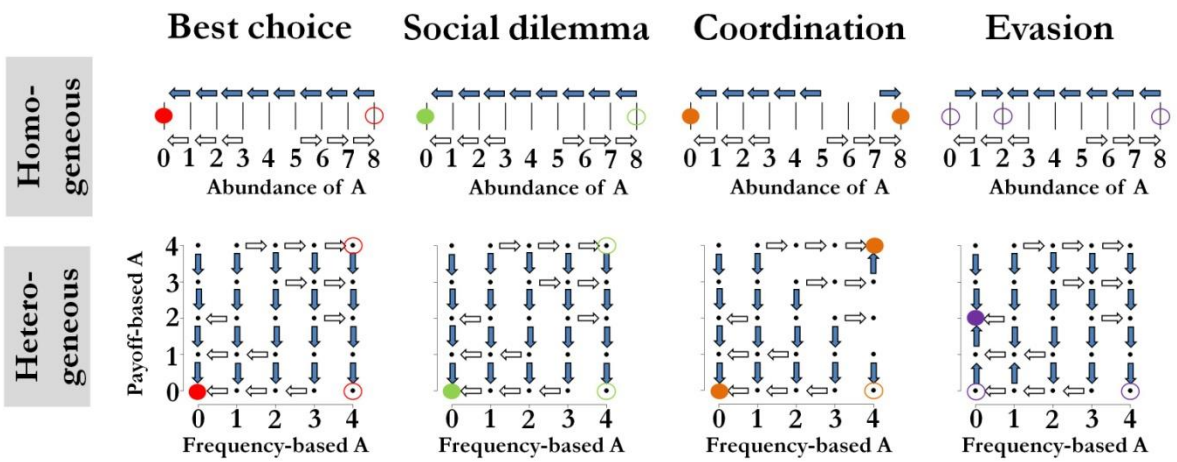
The bottom row of panels shows counts of combinations of information types requested in corresponding rounds. Colours indicate specific (combinations of) information requested: only the payoff (green), only the decision (red), or both (blue). Information requests tended to decrease over time in each of the contexts. We observe that fewer decisions were based on social information in the best choice context, compared to the other interaction contexts. Also, we find that most requests were for either a combination of previous decision *and* associated payoffs (blue bars), or for previous decisions only (red bars). In the best choice context, requests for both decision and payoffs are the dominant type of information request. In the other contexts, most requests were for previous decisions only. These observations are in line with those from the primary experiment.



Supplementary Figure 6| Follow-up experiment: social learning strategies vary between individuals. This figure corresponds to Figure 2 in the main text, which presents the corresponding outcomes in the primary experiment. **a,** Distribution of subjects' reliance on social learning in decision making, measured as the fraction of decisions in which subjects chose to collect peer information. **b,** Distribution of subjects' reliance on payoffs in social learning, measured as the fraction of requests for group member information that included at least one request for previous payoffs. The strong overrepresentation of extreme strategies – some individuals always requested payoff information, and others disregard it altogether – confirms our findings in the primary experiment.



Supplementary Figure 7| Follow-up experiment: social learning strategies tend to be consistent across contexts. This figure corresponds to Figure 3 in the main text, which presents the corresponding outcomes in the primary experiment. **a,** The observed fraction of individuals with consistent (high or low) reliance on social learning is significantly larger than expected on the basis of independence across contexts (χ^2 = 29.551, n=200, d.f.=1, P < 0.001). **b,** The observed fraction of individuals with consistent reliance on payoffs is significantly larger than expected on the basis of independence across (χ^2 = 13.673, n=200, d.f. = 1, P < 0.001). 'Sensible switchers' (that tended to rely on payoff information in the best choice context only) occur at a high frequency, but this frequency is not significantly larger than expected on the basis of independence (χ^2 = 2.801, n=200, d.f. = 1, P < 0.094). For each of the four contexts, we categorized subjects' reliance on social learning as 'low' or 'high' (requesting information in <25%, or >25% of the rounds, respectively). Similarly, subjects' reliance on payoffs in social learning was categorized as 'low' or 'high' (with <50%, or >50% of the requests including payoff information, respectively) for each of the contexts separately. Next, we calculated the expected percentage of consistent individuals, assuming independence of behaviour in the different contexts.



Supplementary Figure 8 | Effects of individual variation in social learning strategies on the dynamics of cultural evolution. Arrows indicate the expected direction of change due to social learning in a group of eight individuals. Individuals learn based on frequencies (white arrows) or based on payoffs (blue arrows). The top row of panels represents the possible dynamics in a group that is homogeneous with respect to social learning strategies: each individual updates its behaviour based on either frequency- or payoff-based learning with a 50-50 probability. The bottom row of panels represents the possible dynamics in a group that is heterogeneous with respect to individual social learning strategies: four individuals always learn based on frequencies, and the other four always learn based on payoffs. Closed and open circles indicate dynamically stable and unstable equilibria of the systems, respectively. See Supplementary Note 1 for a detailed description of the simulation model.

Supplementary Notes

Supplementary Note 1. Implications of individual variation in social learning strategies for cultural evolution: details of the simulation model

Here we give more detailed information about the simulation model presented in the main text. This model is a conceptual 'toy model', providing a first theoretical investigation on whether individual variation in social learning strategies influences the dynamics and outcome of cultural evolution. We do not intend to make any testable predictions or accurately mimic reality. Rather, we aim to provide a proof of principle that individual differences in social learning strategies potentially have a strong effect on the outcome of cultural evolution.

For simplicity, we focus on two different kinds of learning: payoff-based learning (in which individuals tend to imitate individuals with higher payoffs) and frequency-based learning (in which individuals tend to imitate the majority of the group).In our simulations, we compare groups that are *homogeneous* with respect to social learning strategy (*i.e.* all individuals employ the same mixed learning strategy that consists of both payoff-based learning and frequency-based learning) with groups that are *heterogeneous* in this respect (*i.e.* each individual uses either only payoff-based learning or only frequency-based learning, but individuals differ with respect to which type of learning they use).

The basic structure of the model closely follows the setup of our primary experiment. We consider groups in which individuals are involved in social interactions. Individuals have a trait (*A* or *B*) that determines their behaviour in these interactions. Payoffs of *A* and *B* are $pa+(1-p)b$ and $pc+(1-p)d$, respectively, where p denotes the fraction of the group playing *A*. Individuals have pure strategies (either playing *A* or *B*), but their strategy can change over the course of time due to social learning.

For each of the four interaction contexts presented in Figure 4 of the main text, we simulated a process of cultural evolution in 10^5 replicate groups. In each of these replicates, individuals interacted in groups of eight and obtained payoffs according to the same payoff matrices as were used in the primary experiment (see Methods section of the main text). After each interaction, one individual was randomly drawn to update her strategy (see below). This cycle was repeated for 160 iterations (loosely according to the number of potential strategy updating events in the experiment, in which eight individuals could update their strategy in each of 20 rounds). At the beginning of each simulation, seven individuals played *A* and one played *B*. This setup of our simulation mimics the introduction of a deviant behaviour in a resident group of which all individuals have adopted the same behaviour. Simulations were programmed in C++ and the code is available upon request from the corresponding author.

We implemented updating of traits through social learning is such a way that individuals used the same amount of information as was allowed in the primary experiment. In the event of frequency-based learning, behavioural updating was based on the previous decisions of a sample of six group members. The behaviour was updated to the behaviour exhibited by the majority of the sampled group members (if both behaviours were observed in equal proportions, the behaviour was not

changed). For payoff-based learning, behavioural updating was based on both the previous decision and the previous payoff of a sample of three group members. The behaviour was updated to the behaviour that was associated with the highest payoff in this sample (but only if this payoff exceeded own payoff). In the simulation model, we do not allow for spontaneous switching between *A* and *B* (cf. mutation or innovation), neither do we add a stochastic component to the payoffs. These assumptions are made for simplicity. We documented the number of simulation runs that fixated for *A* and for *B* (recall that there is no spontaneous strategy switching in our model, so a fixated group cannot change anymore).

In Supplementary Figure 8, we illustrate how frequency- and payoff-based learning affect the expected change in the abundance of *A* and *B* in each of the four interaction contexts, in groups that are homogeneous and heterogeneous with respect to their social learning strategies. This graph aims to give insight into the dynamics that lead to the outcomes of cultural evolution presented in the panels in Figure 4 of the main text. In the best choice context and the social dilemma, the individually inferior behaviour *A* is more likely to fixate in a group in the absence of individual variation. The arrows in Supplementary Figure 8 offer an intuition for this result: if a payoff-learner in the heterogeneous group has adopted the individually superior behaviour *B*, this individual will never switch back to *A*. This guarantees the eventual fixation of *B*. However, in heterogeneous groups, fixation in *A* always remains possible, particularly when *B* is initially rare and is disfavoured by frequency-based learning. In the coordination game, there are no differences between homogeneous and heterogeneous groups in terms of the outcome of cultural evolution. This is because any learning event (payoff-based learning or frequency-based learning) will lead to choosing *A*, so the fixation of *A* is the only possible outcome.

The most striking effects of individual variation in social learning strategies occur in the evasion game. The probability of fixation of either *A* or *B* is much higher in groups that are homogeneous. Note that in the evasion game, average payoffs are higher when *A* and *B* coexist (cf. the payoff matrices in the Methods). As illustrated by the bottom right panel of Supplementary Figure 8, the dynamics of cultural evolution in the heterogeneous group will lead to a group composition where all frequency-based learners play *B*, whereas the payoff-based learners play *A* and *B* in equal proportions. When at this equilibrium, frequency-based learning events can no longer lead to the fixation of the most common strategy (*B*), because the only individuals that are playing *A* are payoff-based learners. In contrast, in a homogeneous group in which two players are playing *A* and six are playing *B*, fixation of *B* (through two consecutive frequency-based learning events) is still possible, and even likely to occur at some point. The emerging role differentiation in the heterogeneous group, in which all frequency-based learners play *A* and the payoff-based learners play *A* or *B* with equal probability, ensures that the group retains behavioural polymorphism. In this particular interaction context, individual variation in social learning strategies increases average payoffs in a group.

Supplementary Note 2. Experimental instructions and screenshots

This Supplementary Note contains the instructions that participants received on paper. The on-screen instructions follow below.

Instructions on paper

Welcome!

This session will last for approximately 2 hours. During the session **it is not allowed to talk** or communicate with the other participants. If you have a question, please raise your hand and one of us will come to you to answer it. During this session you will earn money. The amount you earn depends on your decisions and (sometimes) the decisions of others. At the end of the session the amount you have earned, plus a show-up fee of 5 euros, will be paid to you in cash. These payments are anonymous; you will be paid individually in a separate room. Please **stay seated at the end of the session** until your desk number is called. We will not inform any of the other participants about your earnings. It is impossible for us to associate your desk number with your identity.

Setup

In the coming session you own an imaginary farm. You will run this farm in **4 different experiments** in which you can earn points. At the end of the session, these points will be translated into real money (1000 points = 1 euro). At the beginning of each experiment, you will be grouped with 7 randomly chosen other participants in the room and each is randomly labeled with a number 1-8. All experiments are **anonymous**; you cannot find out the real identity of the other Farmers in your group, and they cannot find this out about you. During each experiment, your group consists of the same 8 participants, but when a new experiment starts, new groups are formed. At the beginning of each experiment, you will receive 2500 points to start with. The experiments are completely separate – what you earn in one experiment does not influence what you can earn in the other experiments. Also, it is likely that your earnings will vary between experiments.

Making decisions

Each experiment lasts for **20 Seasons**. In each Season, you will make a decision how to use the land on your farm. In each experiment, you will decide between two different options. All Farmers in your group will make this decision at the same time.

Once all Farmers have made their decision, all Crops of all Farmers in your group are sold on the Market. This procedure is automated in the computer program. The number of points you earn (your 'revenue') depends on three things:

1. **Crop value**: one of the two Crops is worth more points. This Crop yields higher revenues on the Market.

2. **Your land production**: the number of Crops you produce in a Season is affected by things that you don't control, such as crop diseases or pests. In each Season, Farmers that make the same decision, will not receive exactly the same revenues. It will sometimes occur that a Farmer that chose the Crop with the higher value, will nevertheless get less revenues than a Farmer that chose the Crop with the lower value. However, on average, the Crop with the higher value will give higher revenues. The variation in Crop production is fully determined by chance.
3. **Decisions of other Farmers**: in experiment 1, your revenues are independent of the decisions of the other Farmers in your group. In experiments 2, 3 and 4, your revenues **also** depend on the decisions of other Farmers. *How* your revenues depend on your decisions and the decisions of the other Farmers in your group, is given by the 'Market Rules'. These Rules are different in each of the experiments. The Market Rules will always be explained by the computer program at the beginning of each experiment. During an experiment, the Market Rules remain the same. Before the Seasons of an experiment start, you will fill out a short Quiz to check if you understand the Market Rules.

At the end of each Season, you are informed about the number of points you received for your Crops. It is also possible to lose points. The points you earn are stored in the computer memory.

Collecting information

Before making your decision, you can **collect information** about the decisions and revenues of the Farmers in your group. Three kinds of information are available about each Farmer:

i) Decision in the previous Season
ii) Revenue in the previous Season
iii) Total points in the current experiment

At the start of each Season, you can indicate whether or not you want to collect information. Collecting information costs **25 points**. For those 25 points, you can collect up to 6 pieces of information in total. Remember that your revenues only *partly* depend on your land production in a Season. Collecting information about the decisions and revenues of other Farmers can be useful in finding out which Crops the other Farmers are producing, and which Crop has the higher value.

If you indicated not to collect information, no costs will be charged. In the first Season of each experiment, no information is available yet. At the start of the session, a **test trial** will make you familiar with the decision making environment on the computer screen. This test trial lasts for 5 Seasons. You cannot earn anything; the points are not worth any money. After the test trial, the experiments will start.

Time limits

In each Season (except the first of each experiment), you have **10 seconds** to decide if you want to collect information about the Farmers in your group. If you have not decided within this time

period, you will not get the opportunity to collect information. If you indicated 'Yes' (and confirmed), you have **20 seconds** to decide which pieces of information you wish to collect. In the Decision screen the information you collected is shown to you, and you can make your decision how to use your land. You have **20 seconds** to make this decision. If you have not decided (and confirmed your decision) within this time period, the computer program will make a random choice for you. When the Season has finished, a Results screen shows you the number of points you earned in this Season. This screen is shown to you for **10 seconds**. After the Results screen, a new Season will start.

After 20 Seasons, the experiment is over, and a new experiment will be started by us. At the beginning of this new experiment, you will again be grouped with 7 randomly chosen other participants in the room and each participant is randomly labeled with a number 1-8, and the new Market Rules are explained in the computer program. After 4 experiments, the session ends.

End of the session

At the end of the series of 4 experiments, you are asked to fill out a short Questionnaire. Once you have filled out this Questionnaire, you will be paid in a separate room. Please stay seated until we call your desk number.

On-screen instructions

Before a new environment started, a screen was displayed to the subjects, announcing:

A new experiment starts now.
You are grouped with 7 randomly chosen other participants.
You receive 2500 points to start with.

Subsequently, the 'Market rules' of the upcoming environment were displayed. There rules were specific of each of them, followed by a quiz to check the subject's understanding of the market rules. The quiz consisted of four statements, and subjects had to indicate if they were true or false. The rounds (or 'seasons') of the environment were started once all subjects had correctly filled out the quiz.

Best choice

Market rules

In this experiment, two Crops are available: **Potatoes** and **Wheat**.
One of the two Crops (Potatoes or Wheat) has a higher value than the other Crop.
The revenues of your Crops **do not** depend on the decisions of the other Farmers in your group.

Before the Seasons of this experiment start,
a short Quiz will check if you understood the Market Rules.

Quiz

1.In each Season of this experiment, I can decide to produce Potatoes or Wheat.

2. Before I make my decision, I can collect information about the decisions of the other Farmers in my group.

3. The revenues of my Crop production also depend on the decisions of other Farmers in my group.

4. My group consists of the same Farmers during this experiment. When the experiment has finished, new groups will be formed randomly.

Social dilemma

Market rules

In this experiment, Apples are the only Crop available.
Each Season, you can decide if you want to **spray Pesticides** during the production of your Apples.
Using Pesticides will increase your land production.
However, the **more** Farmers decide to use Pesticides,
the **lower** the value of **all Apples** on the Market
(both Sprayed Apples and Unsprayed Apples).

Quiz

1. New groups have been formed randomly.

2. In each Season of this experiment, I can decide to use Pesticides to produce my Apples.

3. The value of all Apples is higher when I choose to use Pesticides.

4. The value of all Apples is highest, if all Farmers produce their Apples without using Pesticides.

Coordination

Market rules

In this experiment, two Crops are available: **Melons** and **Pumpkins**.
In principle, one of the two Crops (Melons or Pumpkins) has a higher value than the other Crop.
Yet, the Supplier of the Crop seeds will give more discount
if more Farmers decide to produce the same Crop.

This means that the revenues of your Crops **also** depend on the decisions of the other Farmers in your group:
the **more** Farmers decide to produce a specific Crop, the **higher** its revenues.

Quiz

1. In this experiment, the revenues of my decision depend on the decisions of the other Farmers.

2. The more Farmers decide to produce a certain Crop, the more discount the Supplier will give. This will increase the revenues of this Crop.

3. The more Farmers decide to produce Melons, the higher the revenues of Melons will be.

4. The more Farmers decide to produce Pumpkins, the higher the revenues of Pumpkins will be.

Evasion

Market rules

In this experiment, two Crops are available: **Carrots** and **Onions**.
In principle, one of the two Crops (Carrots or Onions) has a higher value than the other Crop.
Yet, if a Crop is scarce on the Market, its value increases.
This means that the revenues of your Crops **also** depend on the decisions of the other Farmers in your group:
the **more** Farmers decide to produce a specific Crop, the **lower** its value.

Quiz

1. In this experiment, the revenues of my decision depend on the decisions of the other Farmers.

2. The more Farmers decide to produce a certain Crop, the lower its value.

3. The more Farmers decide to produce Carrots, the higher the revenues of Carrots will be.

4. The more Farmers decide to produce Onions, the higher the revenues of Onions will be.

Collecting information and making decisions

Before making their decisions in a Season, subjects could collect information from their fellow group mates. On the screen, subjects were asked the following:

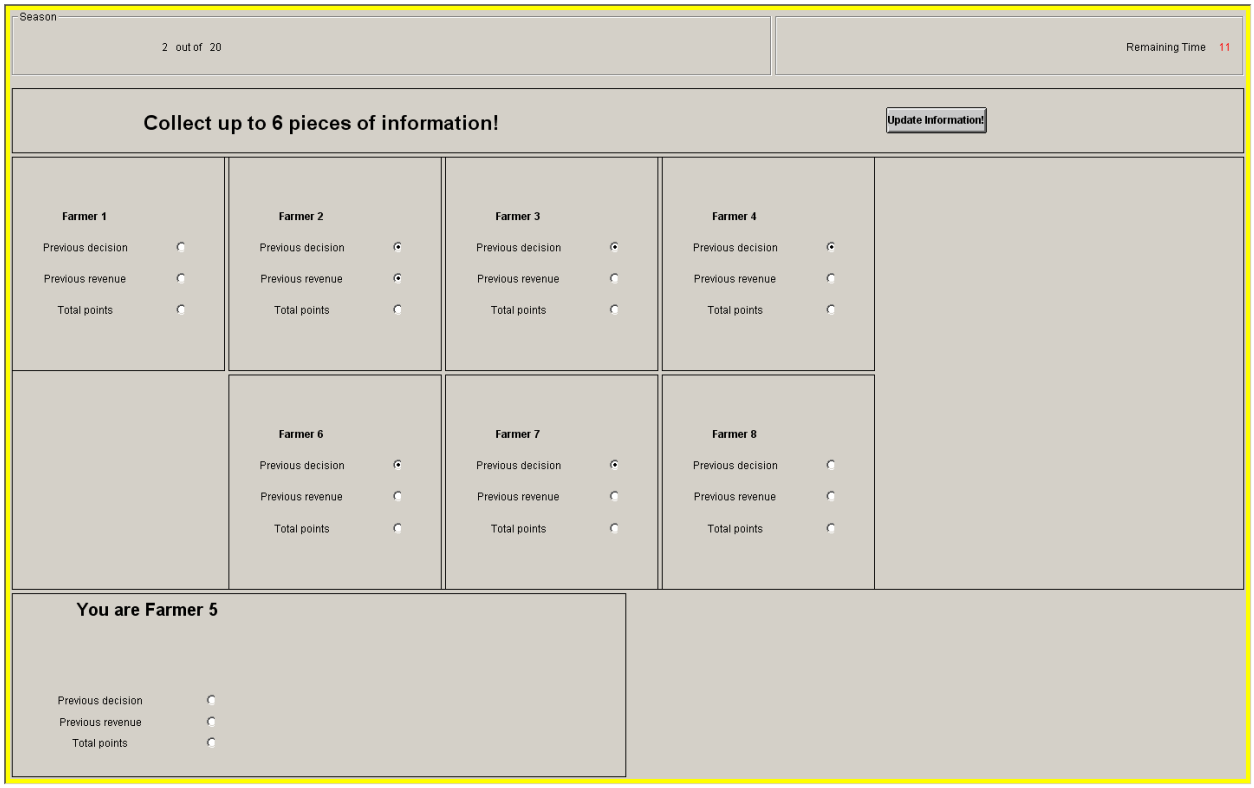
Do you want to collect information before making your decision in this Season?

o yes

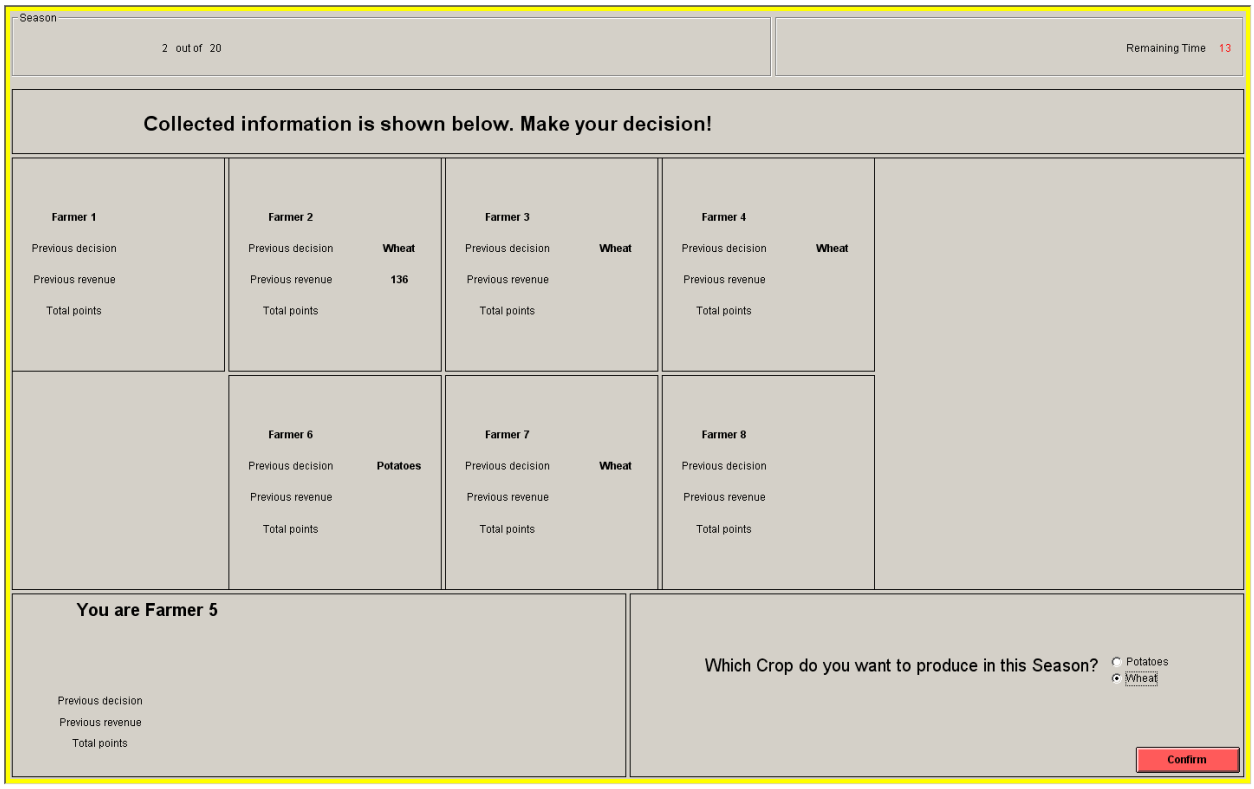
o no

If a subject ticked 'yes', a screen appeared in which he could collect information about his fellow group members (see below for screenshots). If subjects indicated 'no', they immediately went on to the decision screen. After all subjects made their decisions, a Results screen was shown, giving information about the payoffs acquired in the current Season. After 20 Seasons, subjects were also shown their total number of points collected in the current environment.

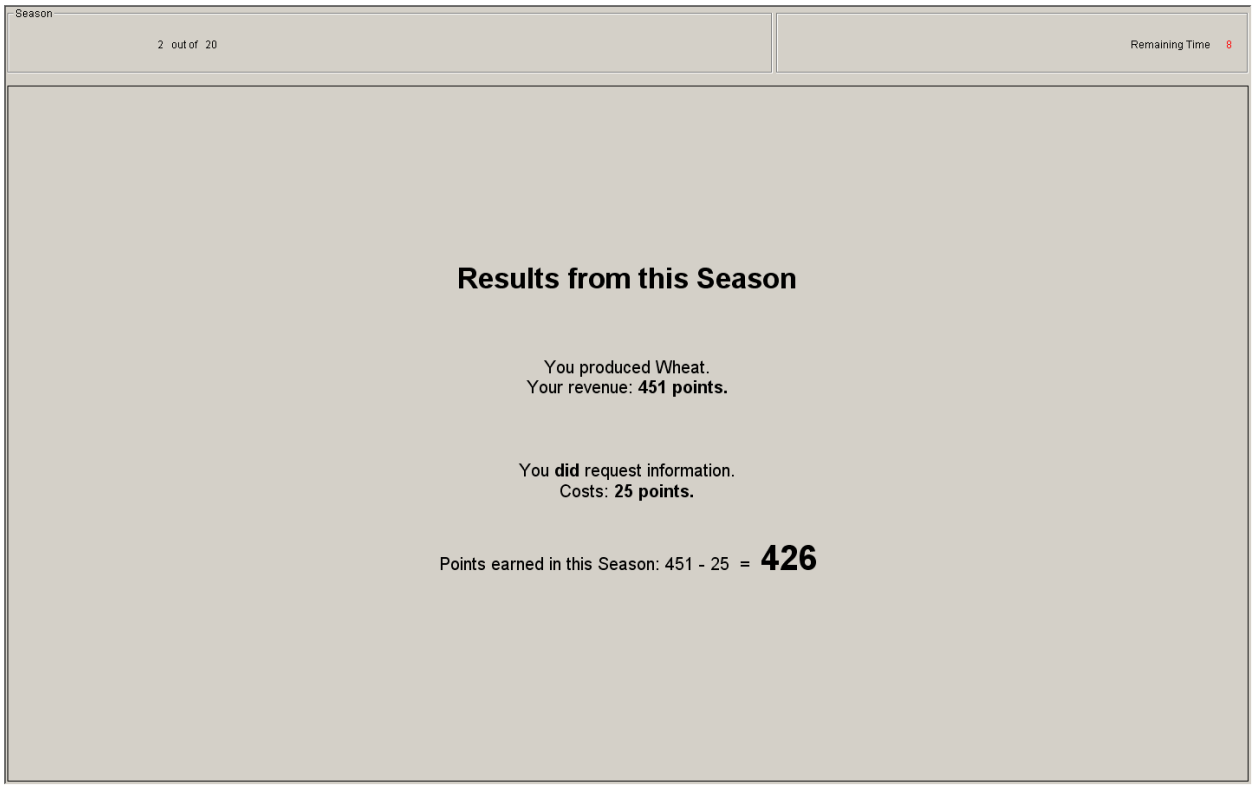
Screenshots



Collecting information of fellow group members. Subjects could check radio buttons to collect up to six pieces information of each other Farmer in their group separately. After checking the desired information, they could click the button 'Update information!'.



View of the collected information and decision making. The program shows the requested information. In the bottom right part of the screen, subjects could make their planting decision for that Season.



Screen showing the results of the current Season. After making their planting decision, subjects were informed about their revenues. If information was requested, the costs were subtracted.