

Article

Social Learning between Groups: Imitation and the Role of Experience

Karl H. Schlag 

Economics Department, University of Vienna, 1090 Vienna, Austria; karl.schlag@univie.ac.at

Abstract: Social learning often occurs between groups with different levels of experience. Yet little is known about the ideal behavioral rules in such contexts. Existing insights only apply when individuals learn from each other in the same group. In this paper, we close this gap and consider two groups, novices and experienced. Experienced should not learn from novices. For novices learning from experienced, a particular form of probabilistic imitation is selected. Novices should imitate any experienced who is more successful, and sometimes but not always imitate an experienced who is less successful.

Keywords: imitation; social learning; experience; improving; reverse proportional imitation

JEL Classification: D83; C73

1. Introduction

Cultural transmission is recognized as being important for the evolution of culture [1]. Cues help to identify groups who are likely to possess better than average information, in particular when information about success is less available or noisier. Experience or age is a popular cue [2]. To copy or imitate the behavior of those possessing such cues saves the cost of learning and helps pass on the knowledge acquired by others. Imitation seems to be the only natural behavior for anyone who has no information about the environment. Yet, it does not automatically make sense to assume that those learning imitate blindly. Typically not all those with the cue have superior performance. Moreover, newcomers also have information of their own. They have limited experience in the new environment and possibly have some information about the success of those possessing such a cue. We wish to investigate how limited own information can be combined with knowledge (or belief) that others are more experienced when subjects choose their actions. The importance of when to “copy” is acknowledged in the field both among humans [3] and among animals [4]. Yet, our paper differs from other papers on cultural transmission that either assume that those belonging to one group blindly copy those of the other [1] or assume that the older generation can impose their behavior on the new generation [5].

Schlag [6] presents a simple formal model to help understand when to copy and when not, in view of aggregating information within a population. It has been used to explain behavior in a large variety of environments, ranging from foraging behavior of fish, to learning decisions of human beings (e.g., see [7]). The downside of the original paper is that it only applies to subjects learning within the same group, not to learning between groups when all know that subjects in one group are on average better informed than in the other. Yet many of the applications have precisely these features. The objective of this paper is to close this gap.

There are two challenges. A priori it is an open question whether information can be aggregated as in homogeneous populations. In homogeneous populations, it is as if individuals observe each other. One can then design a rule such that the probability to switch to the action of the observed is greater than the probability that the observed



Citation: Schlag, K.H. Social Learning between Groups: Imitation and the Role of Experience. *Games* **2022**, *13*, 60. <https://doi.org/10.3390/g13050060>

Academic Editors: Ulrich Berger and Heinrich H. Nax

Received: 13 October 2021

Accepted: 31 August 2022

Published: 6 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

switches to your action if and only if the observed is choosing a better action. When all use this rule average payoffs increase. However, in the setting of this paper, when computing the change in payoffs among the novices, one is not interested in the behavior of the individuals (experienced) that the novices observe. So the intuition from the homogeneous population does not carry over. The second challenge is whether differences in experience persist. If they do not necessarily persist then it is not clear how a common understanding of the difference in experience can emerge in the first place.

For our analysis, we start with the model of [6] where subjects have to choose between two actions that each give a random payoff. Subjects have very little a priori information about how these payoffs are generated. They do not choose a behavioral rule that is best according to some specific prior. Instead, they choose a rule under which they are able to learn in the sense of increasing their payoffs from one round to the next, regardless of the environment, provided all use this rule.

The new feature in the present paper is that subjects observe someone else who has a different experience and that subjects know whether the subject they observe is more or less experienced than they are. Specifically, we consider two different groups, experienced and novices. Experience is captured as follows. It is known to all that the proportion among the experienced choosing, the better action is higher than it is among the novices. However, it is unknown how large this difference in experience between the groups is. Given that there are two types of individuals we need to select a rule for a novice who observes an experienced and for an experienced who observes a novice. How to behave when observing someone from the same group follows from [6].

Behavior is easy to select when experienced observe novices. Fear that the novices are all choosing the worse action leads the experienced to never switch. A single observation of the payoff of each action is not enough to learn which action is better as payoffs are random.

What about the novices observing experienced? Given they observe a more experienced subject, one may think that it is best for the novices to use the rule Always Imitate (AI), to always imitate the experienced regardless of the payoffs. This is indeed one type of behavior that will make the inexperienced always better off as they adapt the behavior of the experienced from the previous round. Such behavior is consistent with the observance of taboos. However, this behavior is not very effective if the difference in experience is negligible. In that case, under this rule, their behavior hardly changes. An alternative is for all to use the Proportional Imitation Rule (PIR) selected by [6] for learning within groups. Under this rule, never imitate someone else who has a lower payoff and imitate someone who was more successful with a probability that is proportional to how much more successful they were. It turns out that this rule will lead to an increase in payoffs of the novices.

In fact, any rule that increases payoffs under within-group learning will also do so for the novices when they observe the experienced group. The intuition is no longer based on net switching behavior as in homogeneous populations. It turns out that the difference in experience leads to a stronger increase in payoffs as compared to the case where there is no difference in experience. More choosing the worse action see the better action which leads to more switching to the better action. At the same time, less choosing the better action will see the worse action and be tempted to switch away from the better action.

However, PIR is not a desirable rule for novices learning from experienced. The reason is that it switches too seldom. In this paper, we discover a new rule called the Reverse Proportional Imitation Rule (RPIR) that dominates PIR. The RPIR is defined as follows. Always imitate those that perform better, and sometimes imitate those that perform worse where the probability of not switching to an experienced performing worse is proportional to the difference in payoffs. It is as if the roles of the switching individual and the observed individual are swapped and then the PIR is applied. The probability of a novice imitating an experienced with a low payoff is only small if the payoff of the novice is high. In particular, novices do not ignore their own performance when they are successful and

the experienced individual observed is not. RPIR does as well as PIR when there is no difference in experience but outperforms it for any degree of difference in experience.

Our above discussion reveals two candidate rules, RPIR and AI. How do they compare? RPIR is best when experienced difference is small while AI is best when it is large. To select rules among the rules that always lead to an increase in payoffs we evaluate the performance of a rule in terms of loss of not having chosen the best one given the true difference in experience. We find that RPIR performs much better in terms of loss than AI. AI is too specialized for the case where experienced know what is best. On the other hand, RPIR involves lots of switching and hence does not perform badly in that case. The most preferred rule involves mixing between these two rules, putting the most weight on RPIR.

We also consider the case where not both payoffs are observable. Here we select the rule that mixes equally likely between Always Switch and the rule that is selected in [6] when there is no difference in experience.

Finally, we investigate whether differences in experience persist over time when both payoffs are observable. We find that it will persist if individuals can choose which group to learn from. This is intuitive as everyone wants to learn from the experienced and the experienced have a head start given their experience. However, it will not necessarily persist if individuals observe each other. To see this assume that the difference in experience is very small. Then novices are learning as if they are observing other novices while experienced are not learning when they see novices as they do not switch in that case. Consequently, experienced learn slower than novices and can be overtaken.

We proceed as follows. In Section 2 we introduce the model with two actions. In Section 3 we analyze the special case where payoffs are either 0 or 1. The more general case where payoffs lie between 0 and 1 is investigated in Section 4. In Section 5 we consider the case where payoffs of others cannot be observed. Section 6 contains the analysis of how the difference in experience changes over time. In Section 7 we conclude.

2. The Model

Individuals belonging to a large (essentially infinite) population repeatedly and independently face the following decision problem. There are two actions, denoted by A and B , from which a subject can choose. We use C to denote an element of $\{A, B\}$. Choice of action C in round t ($t = 1, 2, \dots$) yields a random payoff that belongs to a common bounded interval which we assume for simplicity to be $[0, 1]$.¹ Payoffs are realized independently of other events. Let a and b be the corresponding expected payoffs of actions A and B . All subjects prefer the action that yields the higher expected payoff, which we then also call the better action. Let P_C define the distribution of payoffs realized when choosing action C .

Individuals belong to one of two large groups, called novice and experienced. Fix t . Let $x_{t,C}$ and $y_{t,C}$ denote the fraction of novices and experienced respectively choosing action C in round t , so $x_{t,C}, y_{t,C} \geq 0$ and $x_{t,A} + x_{t,B} = y_{t,A} + y_{t,B} = 1$. The terminology novice and experienced comes from the fact that we assume at time t that the fraction of experienced choosing the better action is larger than the fraction of novices choosing this action. So we assume $y_{t,A} \geq x_{t,A}$ if $a > b$ and $y_{t,B} \geq x_{t,B}$ if $a < b$. This is equivalent to assuming that $(y_{t,B} - x_{t,B})(b - a) \geq 0$. Note that this is also equivalent to assuming that the average payoff among the experienced is higher than among the novices, i.e., $y_{t,A}a + y_{t,B}b \geq x_{t,A}a + x_{t,B}b$. Individuals know which group they belong to but do not know anything about the distribution of payoffs except that payoffs belong to $[0, 1]$. In particular, they do not know which action is better. They do not know the fractions of the actions chosen in the population. However, they do know that the fraction of experienced choosing the better action is larger than the fraction of novices choosing this action.

After choosing an action and realizing a payoff each subject observes the group identity, the action chosen, and the payoff realized of some randomly chosen subject. Based on their own experience in the last round and on this information the subject then chooses an action for the next round. It is assumed that all subjects belonging to the same group use the same rule for revising their action.

We refer to learning within group when the observed subject belongs to the same group as the observant, otherwise, we talk about learning between groups. In the following, we investigate learning between groups which extends the methodology of learning within groups by Schlag [6].

Consider the first behavior of the novices. Their learning rule will be denoted by F and is given by a mapping from $\{A, B\} \times [0, 1] \times \{A, B\} \times [0, 1]$ into $[0, 1]$. $F(C, u, D, v)$ denotes the probability of not choosing action C in round $t + 1$ after choosing action C and receiving payoff u in round t and observing an experienced subject who chose action D and received payoff v in round t . *Never Switch* is the rule F that satisfies $F(C, u, D, v) = 0$ for $u, v \in [0, 1]$. A rule F is called *imitating* if $F(C, u, C, v) = 0$ for all $u, v \in [0, 1]$. *Always Imitate* (AI) is the imitating rule F that satisfies $F_{AI}(C, u, D, v) = 1$ for $C \neq D$. The Proportional Imitation Rule (PIR), selected in [6] for learning within groups, is the imitating rule F that satisfies $F(C, u, D, v) = \max\{0, v - u\}$ when $C \neq D$.

Let f_{CD} denote the expected probability of switching under rule F after choosing C and observing a subject choosing D , calculated ex-ante to receiving a payoff from action C . So

$$f_{CD} = \int_0^1 \int_0^1 F(C, u, D, v) dP_C(u) dP_D(v).$$

Using this we compute the change in the play of action B and find

$$\begin{aligned} x_{t+1,B} - x_{t,B} &= -(x_{t+1,A} - x_{t,A}) \\ &= x_{t,A}y_{t,A}f_{AA} + x_{t,A}y_{t,B}f_{AB} - x_{t,B}y_{t,A}f_{BA} - x_{t,B}y_{t,B}f_{BB}. \end{aligned} \tag{1}$$

In the following, we will search for rules F that are “improving”. A rule is called improving [6] if the average payoff within the same group increases between rounds when all use this rule. As there are only two actions, this means that the probability of switching to the better action is larger than the probability of switching to the worse action. Formally, F is called improving if

$$\begin{aligned} x_{t+1,B} - x_{t,B} &\geq 0 \text{ if } b > a \text{ (so } y_{t,B} \geq x_{t,B}), \\ x_{t+1,B} - x_{t,B} &\leq 0 \text{ if } b < a \text{ (so } y_{t,A} \geq x_{t,A}). \end{aligned}$$

In particular, the above has to hold when $x_{t,A} = y_{t,A}$. Notice that when $x_{t,A} = y_{t,A}$ then it is as if a subject is observing someone who belongs to the same group. So, if a rule is improving for learning between groups, then it is also improving for learning within the group. The findings in Schlag, 1998 [6] for learning within the group are as follows. F is improving for learning within the group if and only if (i) F is imitating and (ii) there exists $\sigma \geq 0$ such that

$$F(A, u, B, v) - F(B, v, A, u) = \sigma(v - u) \tag{2}$$

holds for all $u, v \in [0, 1]$.

Assume that F is improving (between groups). As F is improving within groups, it follows from (i) and (ii) above that $f_{AA} = f_{BB} = 0$ and there exists $\sigma \geq 0$ such that $f_{AB} - f_{BA} = \sigma(b - a)$. Let $\delta_t = y_{t,B} - x_{t,B}$. Then (1) becomes

$$\begin{aligned} x_{t+1,B} - x_{t,B} &= x_{t,A}x_{t,B}\sigma(b - a) + \delta_t(x_{t,A}f_{AB} + x_{t,B}f_{BA}) \\ &= x_{t,A}x_{t,B}\sigma(b - a) + \delta_t x_{t,A}\sigma(b - a) + \delta_t f_{BA} \\ &= x_{t,A}y_{t,B}\sigma(b - a) + \delta_t f_{BA}. \end{aligned} \tag{3}$$

Note that δ_t measures the difference in experience when B is the better action (so if $b > a$).

If there is no difference in experience, so if $\delta_t = 0$, then (3) is the same formula as obtained in Schlag, 1998 [6]. The additional term with δ_t results from the higher likelihood of seeing better action among the experienced than among the novices. More novices that

have chosen A get the possibility to observe B and fewer novices who have already been choosing the better action B are distracted by observing the performance of action A .

Following (3), as $\sigma, f_{AB}, f_{BA} \geq 0$, any improving rule for the case of a single group, so for $\delta_t = 0$, is also improving for novices learning from the experienced (so when $\delta_t \geq 0$).

Before summarizing this result, we consider the behavior of the experienced when observing a novice. The improving condition requires that experienced are on average better off in the next round even when currently all experienced are choosing the better action and all novices are choosing the worse action. So under these specific circumstances, no experienced should switch. However, as one sample of payoffs of each action provides no information on which action is better, an experienced should never switch when observing a novice. The same intuition is used in Schlag, 1998 [6] to show that improving rules for learning within the group are imitating.

We have thus proven the following result.

Proposition 1.

- (i) *A learning rule for novices observing experienced is improving if and only if it is improving for the case where novices observe novices.*
- (ii) *Never Switch is the only improving rule for experienced observing novices.*

So, an improving rule for learning between groups has to be also improving for learning within the group as the difference in experience might be negligible. On the other hand, the reason that an improving rule for learning within the group is also improving for learning between groups is as follows. The fact that the experienced choose the better action more likely increases learning due to those choosing the worse action and switching. Similarly, it fosters learning as fewer choose the worse action, and hence fewer of those choosing the better action switch to the worse action.

On the side, we obtain from (3) that payoffs increase more when novices observe experienced than when they observe other novices. Moreover, we find that switching is good, provided the rule is improving for learning within the group. This latter fact will influence our selection of which rules are best.

3. A Simple Environment

We wish to select among the improving rules for novices. To build some intuition we investigate environments where the payoff can be either 0 or 1. Note that a and b are now the probabilities of realizing payoff 1 with actions A and B respectively. Initially we restrict attention to rules where switching probability does not depend on the labelling of the actions, so where $F(C, u, D, v) = F(D, u, C, v)$ for all $u, v \in [0, 1]$ and $C, D \in \{A, B\}$. At the end of this section we argue why this restriction can be made without loss of generality.

Let $s_{ij} = F(A, i, B, j) = F(B, i, A, j)$ for $i, j \in \{0, 1\}$. Then

$$\sigma = s_{01} - s_{10}$$

and

$$f_{BA} = (1 - b)as_{01} + b(1 - a)s_{10} + (1 - a)(1 - b)s_{00} + abs_{11}$$

so

$$x_{t+1,B} - x_{t,B} = x_{t,A}x_{t,B}(s_{01} - s_{10})(b - a) + \delta_t x_{t,A}(s_{01} - s_{10})(b - a) + \delta_t((1 - b)as_{01} + b(1 - a)s_{10} + (1 - a)(1 - b)s_{00} + abs_{11}). \tag{4}$$

As rules do not depend on labels we can assume that B is better than A , so $b \geq a$. It follows from (4) that the increase in fraction of novices choosing the best action is largest if $s_{00} = s_{01} = s_{11} = 1$. Note that there is no best choice of s_{10} that does not depend on a, b, δ_t

and $x_{t,A}$ as increase in s_{10} raises f_{BA} which is good but decreases σ which is bad. We now wish to investigate the effects of s_{10} . Assuming $s_{00} = s_{01} = s_{11} = 1$ we obtain

$$\begin{aligned} x_{t+1,B} - x_{t,B} &= x_{t,A}x_{t,B}(1 - s_{10})(b - a) + \delta_t x_{t,A}(1 - s_{10})(b - a) \\ &\quad + \delta_t((1 - b)a + b(1 - a)s_{10} + (1 - a)(1 - b) + ab) \\ &= x_{t,A}x_{t,B}(1 - s_{10})(b - a) + \delta_t x_{t,A}(1 - s_{10})(b - a) \\ &\quad + \delta_t(1 - b(1 - a)(1 - s_{10})). \end{aligned} \tag{5}$$

So, we are left with the question of how to choose a value for s_{10} . What should a novice who has obtained the highest possible payoff and observes an expert who has obtained the lowest possible payoff do?

Consider best choices for extreme cases of δ_t . If there is no difference in experience, so when $\delta_t = 0$, then it is best to choose $s_{10} = 0$, following (5), which also follows from [6]. If experience is maximal, so when all experienced are choosing the best action ($y_{t,B} = 1$ and hence $\delta_t = x_{t,A}$), then $x_{t+1,B} - x_{t,B}$ is increasing in s_{10} and hence it is best for novices to choose $s_{10} = 1$, which results in the rule AI. This result is intuitive. If the difference in experience is small, then one should treat own and observed equally and maximize net switching by choosing $s_{10} = 0$ and hence $s_{01} - s_{10} = 1$. If instead, the difference in experience is large then one should ignore own payoff and focus as much on the observed by always imitating the observed, so choosing $s_{10} = 1$.

Now consider the performance of the rule for a given value of s_{10} across all values of δ_t . After all, we assume that x_t and y_t are not known, hence that δ_t is not known. The performance of a rule is measured in terms of change in payoffs as given by $(x_{t+1,A}a + x_{t+1,B}b) - (x_{t,A}a + x_{t,B}b) = (x_{t+1,B} - x_{t,B})(b - a)$. We consider loss in this performance that is generated from not knowing x_t and y_t (formally called regret [8], axiomatized by Milnor [9]). For a given value of s_{10} , this is the difference between the best performance overall s_{10} for given x_t and y_t and performance for the given value of s_{10} . As (5) is linear in s_{10} , the best performance under the benchmark where x_t and y_t are known is given by either $s_{10} = 0$ or $s_{10} = 1$. If best performance is obtained when $s_{10} = 0$ then using (5) loss in performance is given by

$$\begin{aligned} &(x_{t,A}x_{t,B}(b - a) + \delta_t x_{t,A}(b - a) + \delta_t(1 - b(1 - a)))(b - a) - (x_{t+1,B} - x_{t,B})(b - a) \\ &= (x_{t,A}x_{t,B}(b - a) + \delta_t x_{t,A}(b - a) - \delta_t b(1 - a))(b - a)s_{10} \\ &\leq x_{t,A}x_{t,B}(b - a)^2 s_{10} \\ &\leq \frac{1}{4}s_{10}. \end{aligned}$$

If instead best performance is obtained when $s_{10} = 1$ then loss in performance is given by

$$\begin{aligned} &\delta_t(b - a) - (x_{t+1,B} - x_{t,B})(b - a) \\ &= -(x_{t,A}x_{t,B}(b - a) + \delta_t x_{t,A}(b - a) - \delta_t b(1 - a))(b - a)(1 - s_{10}) \\ &\leq -(x_{t,A}x_{t,B}(b - a) + x_{t,A}x_{t,A}(b - a) - x_{t,A}b(1 - a))(b - a)(1 - s_{10}) \\ &= a(1 - b)x_{t,A}(b - a)(1 - s_{10}) \\ &\leq \frac{1}{27}(1 - s_{10}). \end{aligned}$$

We are interested in the maximal loss in performance where the maximum is taken over all appropriate values of x_t and y_t , so $(y_{t,B} - x_{t,B})(b - a) \geq 0$. By the above calculations we obtain that the loss in performance is bounded above by

$$\max\left\{\frac{1}{4}s_{10}, \frac{1}{27}(1 - s_{10})\right\} \tag{6}$$

where these calculations show that this bound is tight. Hence (6) represents the maximal loss in performance.

For instance, the rule with $s_{10} = 0$ has maximal loss equal to $1/27 \approx 0.037$ while AI (with $s_{10} = 1$) yields maximal loss $1/4$. So when comparing these two extremes we prefer $s_{10} = 0$ as it attains a smaller maximal loss. Maximal loss can be made smaller if one takes a convex combination between these two rules, by choose $s_{10} = 4/31 \approx 0.129$ as solution to $\frac{1}{4}s_{10} = \frac{1}{27}(1 - s_{10})$. This yields maximal loss equal to $1/31 \approx 0.032$ which is only slightly better than the value $1/27$ attained when $s_{10} = 0$.

To summarize, we find the most appealing improving rule satisfies $s_{00} = s_{01} = s_{11} = 1$ and $s_{10} = 0$. The simplicity of setting $s_{10} = 0$ is chosen to favor the slightly better rule that requires $s_{10} = 4/31$. In the next section we consider more general payoffs.

Before we proceed, we need to discuss our restriction to rules that do not depend on the labeling of actions. Intuitively, such rules will be chosen, given there is nothing known about the environment. In fact, this is true when aiming to minimize maximal loss. The formal reasoning is as follows. Consider some general improving rule F_1 that depends on the labeling of actions. Define a rule F by $F(C, x, D, y) = \frac{1}{2}F_1(C, x, D, y) + \frac{1}{2}F_1(D, x, C, y)$. Then F does not depend on the labeling of actions and it is easily seen that F is improving. Moreover, the maximal loss of F is bounded above by the maximal loss of F_1 . This is because the maximal loss of a sum is less than or equal to the sum of the maximal losses. Hence, according to our criteria for selecting rules, we can restrict attention to rules that do not depend on the labeling of actions.

4. Selecting Rules under General Payoffs

Above we selected a rule for environments where payoffs can be either 0 or 1. The aim was to minimize the loss in performance from not knowing the parameters. The selected rule satisfies $s_{00} = s_{01} = s_{11} = 1$ and $s_{10} = 0$. Clearly, maximal loss in more general environments is weakly larger. In the following, we present several rules that satisfy $s_{00} = s_{01} = s_{11} = 1$ and $s_{10} = 0$ where maximal loss is not larger, hence where maximal loss is equal to $1/27$.

There is a unique imitating rule with this property where switching probability is linear in payoffs, it satisfies $F(A, u, B, v) = 1 - u(1 - v)$. The following slightly more complicated rule maximizes the switching probability among all improving rules that satisfy $\sigma = 1$. It is the imitating rule that satisfies $F(C, u, D, v) = 1$ if $v \geq u$ and $F(C, u, D, v) = 1 - (u - v)$ if $u > v$ for $C \neq D$. This rule is called the Reverse Proportional Imitation Rule (RPIR) and can equivalently be characterized by being the imitating rule such that $F(C, u, D, v) = 1 - \max\{u - v, 0\}$ for $C \neq D$. The name comes from the fact that it is as if one switches roles with the observed and chooses the strategy that the observed would choose if the observed would be using the Proportional Imitation Rule. The property of maximal switching follows from the fact that

$$F(A, u, B, v) - F(B, v, A, u) = v - u$$

since F is improving with $\sigma = 1$, and hence

$$F(C, u, D, v) \leq v - u + F(C, v, D, u) \leq \min\{1, v - u + 1\} = 1 - \max\{u - v, 0\}. \tag{7}$$

This leads us to the following result. First, for completeness, we define maximal loss in performance of a rule F (given $x_{t,A}$ and $y_{t,A}$) among improving rules by

$$\max_{P_A, P_B} \left\{ \max_{F_0 \text{ improving}} \{ (x_{t+1,B}(F_0) - x_{t,B})(b - a) \} - (x_{t+1,B}(F) - x_{t,B})(b - a) \right\}$$

where $x_{t+1,B} = x_{t+1,B}(F_0)$ is given by (1). Once again this is the difference between the maximal performance of one of the desired rules, namely improving rules, and the actual performance of the rule.

Proposition 2. Consider novices observing experienced.

- (i) The Reverse Proportional Imitation Rule is the unique rule that increases payoffs at least as much as any other improving rule with $\sigma = 1$. In particular, RPIR minimizes maximal loss in performance among all improving rules with $\sigma = 1$ and have a maximal loss in performance equal to $1/27$.
- (ii) The rule F^* that is imitating and satisfies $F^*(A, u, B, v) = F^*(B, u, A, v) = 1 - \frac{27}{31} \max\{u - v, 0\}$ minimizes the maximal loss in performance among all improving rules. Its maximal loss in performance equals $1/31$.

Following up on our discussion after Proposition 1, we find that a higher switching probability is better when comparing improving rules with the same value of σ . Ignoring the value of σ , AI is the rule that switches most. This rule does best when the difference in experience is maximal but does worst when the difference in experience is negligible. On the other hand, RPIR does best when the difference in experience is small, as $\sigma = 1$, and also does well when the difference in experience is large. There is no rule that does best among the improving rules in all environments. Trade-offs have to be made and are captured by maximal loss. Our analysis of loss reveals that RPIR is preferred to AI. It is better to protect for a small difference in experience than to bet on the fact that experience difference is maximal. In fact, a maximal loss can be made slightly smaller than under RPIR by randomizing between these two extreme rules. The improving rule that minimizes maximal loss puts a probability of $27/31$ on RPIR and a probability of $4/31$ on AI.

Note that the Proportional Imitation Rule that was selected for learning within a group no longer performs that well for learning between groups. While it is improving for learning between groups, it is dominated by the RPIR. Its maximal loss is substantial as it equals $1/4$. So we find very different rules for learning between groups than for learning within the group.

Proof. Part (i). As shown above, the Reverse Proportional Imitation Rule maximizes switching probability among all improving rules with $\sigma = 1$. Together with (3) this then shows that RPIR increases payoffs at least as much as any other improving rule with $\sigma = 1$. In fact, (7) shows that it is the unique rule with this property.

Let F_1 be the imitating rule such that $F_1(C, u, D, v) = 1 - u(1 - v)$, which satisfies $s_{00} = s_{01} = s_{11} = 1$ and $s_{10} = 0$. Given linearity in u and v and our results in Section 3, the maximal loss in performance of F_1 equals $1/27$. Following (3), the maximal loss in performance of F_{RPIR} is weakly smaller. However, as the maximal loss of F_1 is attained when $u, v \in \{0, 1\}$, and as $F_1(C, u, D, v) = F_{RPIR}(C, u, D, v)$ when $u, v \in \{0, 1\}$ we obtain that the maximal loss in performance of F_{RPIR} also equals $1/27$.

Part (ii). Consider the linear rule F_2 that is imitating and defined by $F_2(A, u, B, v) = 1 - \frac{27}{31}u(1 - v)$. Note that F_2 is improving with $\sigma = \frac{27}{31}$ and satisfies $s_{00} = s_{01} = s_{11} = 1$ and $s_{10} = 4/31$. Following Section 3 this rule attains the smallest maximal loss in performance among all improving rules if payoffs are in $\{0, 1\}$. The linearity of F_2 implies that this statement also holds when payoffs belong to $[0, 1]$. As the rule F^* given in (ii) is identical to F_2 when payoffs belong to $\{0, 1\}$ but always switches more likely than F_2 it follows that maximal loss under F^* cannot be larger than under F_2 which completes the proof. \square

5. Limited Observability

Next, we revisit our analysis in the previous two sections for the case where fewer payoffs are observable. First, we consider the case where one cannot observe the payoffs of others. So, we now consider rules F that map $\{A, B\} \times [0, 1] \times \{A, B\}$ into $[0, 1]$ where $F(C, u, D)$ denotes the probability of not choosing action C in the next round after choosing action C and receiving payoff u and observing someone else who chose action D .

The rule selected by Schlag [6] for learning within the group is the Proportional Reviewing Rule (PRR). It is an imitating rule defined by $F(C, u, D) = 1 - u$ for $u \in [0, 1]$.

Our previous analysis shows that the experienced should not change behavior when observing novices.

We now select among improving rules for novices who observe experienced. Let F be an improving rule.

We first consider payoffs in $\{0, 1\}$. Let $s_i = F(A, i, B) = F(B, i, A)$ for $i \in \{0, 1\}$. Following Proposition 1(i), F must be improving when $\delta_t = 0$, so $s_0 - s_1 = \sigma$. We adapt (4) and obtain

$$\begin{aligned} x_{t+1,B} - x_{t,B} &= x_{t,A}x_{t,B}(s_0 - s_1)(b - a) + \delta_t x_{t,A}(s_0 - s_1)(b - a) \\ &\quad + \delta_t((1 - b)as_0 + b(1 - a)s_1 + (1 - a)(1 - b)s_0 + abs_1) \\ &= (x_{t,B} + \delta_t)x_{t,A}\sigma(b - a) + \delta_t(s_0 - bs_1). \end{aligned}$$

W.l.o.g. we can assume that B is better than A . We then see that increase in payoffs of novices is largest when $s_0 = 1$. This yields

$$x_{t+1,B} - x_{t,B} = (x_{t,B} + \delta_t)x_{t,A}(1 - s_1)(b - a) + \delta_t(1 - b(1 - s_1)).$$

We now select s_1 by analyzing loss in performance. If best performance is given when $s_1 = 0$ we obtain with a few additional steps that

$$((x_{t,B} + \delta_t)x_{t,A}(b - a) + \delta_t(1 - b))(b - a) - (x_{t+1,B} - x_{t,B})(b - a) \leq \frac{1}{4}s_1.$$

In case best performance is achieved when $s_1 = 1$ we obtain

$$\delta_t(b - a) - (x_{t+1,B} - x_{t,B})(b - a) \leq \frac{1}{4}(1 - s_1).$$

Hence, loss in performance is given by $\frac{1}{4} \max\{s_1, (1 - s_1)\}$ and maximal loss is minimal by choosing $s_1 = \frac{1}{2}$.

Now consider the case of general payoffs in $[0, 1]$. The argument made in the proof of Proposition 2 also shows here that the linear rule that coincides with the selected rule above when payoffs are in $\{0, 1\}$ minimizes maximum loss.

Proposition 3. Consider novices observing choice but not payoff of experienced. The rule F that minimizes maximal loss in performance among all improving rules is imitating and satisfies $F(C, u, D) = 1 - \frac{1}{2}u$ for $u \in [0, 1]$ and $C \neq D$. It guarantees loss in performance below $1/8$.

Note that the rule selected above can be interpreted as mixing equally likely between AI and PRR. AI and PRR are the rules that are best when $\delta_t = 1$ and when $\delta_t = 0$, respectively. Each of them on itself yields a maximal loss in performance equal to $1/4$ while mixing equally likely between them reduces it to $1/8$.

We find that the rule that guarantees the lowest maximal loss in performance is an improving rule that has a value of σ equal to $1/2$. It gives up learning when the difference in experience is marginal (as otherwise $\sigma = 1$ is best when $\delta_t = 0$) in order to benefit from higher switching probability when the difference in experience is larger.

Note the substantial value of observing the payoff of the other, in that case, the maximal loss was $1/31 \approx 0.033$ while without observability we obtain $1/8 = 0.125$.

Similarly, one can consider the case in which own payoffs are not observable. This case arises naturally when one has not made a choice yet. For instance, this is the case in the models of the fish experiments of Pike et al. [10]. Here $F(C, D, v)$ is the probability of not choosing C after observing another who chose D and obtained payoff v . The results are very similar to those obtained when only own payoff is observable. Calculations are straightforward and left to the reader. Once again, experienced will not change behavior when observing a novice. The analysis of novice behavior is analogous to the above. The rule that minimizes maximal loss in performance is the imitating rule that satisfies

$F(C, D, v) = \frac{1}{2}v$ for $C \neq D$. Its maximal loss is $\frac{1}{8}$. This rule arises when choosing equally likely between Never Switch and the Proportional Observation rule where $F(C, D, v) = v$ for $C \neq D$, as defined in [6].

6. Evolution of Experience

We now return to the setting where self and others' payoffs are observable and investigate how differences in experience change over time. The current paper relies on the common knowledge that novices are less experienced than the experienced. This can occur quite naturally, namely when the experienced have been facing the environment several times while the novices have only faced it once. Here one has to assume they both started with the same distribution of actions. The difference in experience is also present when both start with the same distribution, the novices have faced the environment less often and neither group has learned from the other. The present paper applies to the first time there is learning between these two groups. A natural question is to investigate what happens when these two groups are learning from each other the second time or later. To apply the present paper it has to be the case that the difference in experience is still present after the first time they have learned from each other. As they do not know which environment they are facing, this has to hold true for any environment. So we now investigate how the difference in experience evolves over time.

Population dynamics are defined by four rules, depending on which group is learning and which group is being observed. We assume that experienced never switch when observing a novice and that the other three rules belong to the class of improving rules with $\sigma = 1$. Let λ be the probability that novices observe an experienced, so any novice observes another random novice with probability $1 - \lambda$, $\lambda \in [0, 1]$. We distinguish two different cases, according to whom experienced observe.

6.1. Experienced Observe only Experienced

Consider first the situation where experienced only observe other experienced, which defacto means that there are only three different rules are being used. This scenario is motivated by the fact that experienced do not wish to learn from novices. Using (3) we obtain that population dynamics are given by

$$\begin{aligned} x_{t+1,B} - x_{t,B} &= (1 - \lambda)x_{t,A}x_{t,B}(b - a) + \lambda(x_{t,A}x_{t,B}(b - a) + \delta_t x_{t,A}(b - a) + \delta_t f_{BA}) \\ &= x_{t,A}x_{t,B}(b - a) + \lambda\delta_t(x_{t,A}(b - a) + f_{BA}), \\ y_{t+1,B} - y_{t,B} &= y_{t,A}y_{t,B}(b - a). \end{aligned} \tag{8}$$

We now investigate how the difference in experience changes over time and find

$$\begin{aligned} \delta_{t+1} &= y_{t+1,B} - x_{t+1,B} \\ &= y_{t,B} + y_{t,A}y_{t,B}(b - a) - (x_{t,B} + x_{t,A}x_{t,B}(b - a) + \lambda\delta_t(x_{t,A}(b - a) + f_{BA})) \\ &= \delta_t + \left(x_{t,A}x_{t,B} + x_{t,A}\delta_t - \delta_t x_{t,B} - (\delta_t)^2\right)(b - a) \\ &\quad - (x_{t,A}x_{t,B}(b - a) + \lambda\delta_t(x_{t,A}(b - a) + f_{BA})) \\ &= \delta_t[1 + (x_{t,A} - x_{t,B} - \delta_t)(b - a) - \lambda f_{BA} - \lambda x_{t,A}(b - a)] \\ &\geq \delta_t(1 - \lambda)(1 - x_{t,B}(b - a)) \geq 0 \end{aligned} \tag{9}$$

as $\delta_t \leq x_{t,A}$ and $f_{AB} - f_{BA} = \sigma(b - a)$ with $\sigma = 1$ implies $f_{BA} \leq 1 - (b - a)$. Note that $\delta_t < x_{t,A}$ holds if and only if $y_{t,B} < 1$. Hence, even if novices only observe experienced, so if $\lambda = 1$, then $\delta_{t+1} > 0$ if $\delta_t > 0$ and $y_{t,B} < 1$. Moreover, note that if $y_{t,B} < 1$ then $y_{t',B} < 1$ for all $t' > t$. This proves the following.

Proposition 4. Assume that experienced observe only experienced. Then the property of being more experienced in round t persists over time, i.e., $\delta_t(b - a) \geq 0$ implies $\delta_{t'}(b - a) \geq 0$ for all $t' > t$. Novices never become as good as the experienced when not all experienced choose the same action in round t , i.e., $\delta_t(b - a) > 0$ and $y_{t,B} \in (0, 1)$ implies $\delta_{t'}(b - a) > 0$ for all $t' > t$.

As a consequence, following (8), we obtain in the long run (as $t \rightarrow \infty$) that all choose the better action provided each of the two actions are initially chosen by some of the experienced.

6.2. Experienced and Novices Learn from Each Other

Now consider the scenario where subjects observe each other. As a novice observes an experienced with probability λ , this is also the probability that an experienced observes a novice. The dynamics of the group of novices remain unchanged (see (8)). As experienced choose not to change behavior when observing a novice, we obtain

$$y_{t+1,B} - y_{t,B} = (1 - \lambda)y_{t,A}y_{t,B}(b - a).$$

Learning among the experienced is slower than in the previous scenario. Change in experience is now given by

$$\begin{aligned} \delta_{t+1} &= y_{t+1,B} - x_{t+1,B} \\ &= y_{t,B} + (1 - \lambda)y_{t,A}y_{t,B}(b - a) - (x_{t,B} + x_{t,A}x_{t,B}(b - a) + \lambda\delta_t(x_{t,A}(b - a) + f_{BA})) \\ &= \delta_t + (1 - \lambda)(x_{t,A}x_{t,B} + x_{t,A}\delta_t - \delta_t x_{t,B} - (\delta_t)^2)(b - a) \\ &\quad - (x_{t,A}x_{t,B}(b - a) + \lambda\delta_t(x_{t,A}(b - a) + f_{BA})) \\ &= \delta_t[1 + (1 - \lambda)(x_{t,A} - x_{t,B} - \delta_t)(b - a) - \lambda f_{BA} - \lambda x_{t,A}(b - a)] - \lambda x_{t,A}x_{t,B}(b - a). \end{aligned}$$

In particular, subjects who are a novice in round t need not be less experienced in round $t + 1$. For instance, given $x_{t,B} \in (0, 1)$, $\lambda > 0$ and $b > a$, if $y_{t,B}$ is chosen such that δ_t is sufficiently small, then $\delta_{t+1} \approx -\lambda x_{t,A}x_{t,B}(b - a) < 0$. This comes as no surprise as the experienced are choosing not to learn from the novices even though it turns out that the experience of novices is very similar.

Proposition 5. Assume that experienced and novices observe each other. The property of being more experienced need not carry over from round t to round $t + 1$. More specifically, given $x_{t,B} \in (0, 1)$, $\lambda > 0$ and $b \neq a$ there exists $y_{t,B}$ such that $\delta_t(b - a) > 0 > \delta_{t+1}(b - a)$.

So, assume that there are two groups in the population and that in round t it is known that those in one group are on average more experienced than those in the other group. Then our analysis and selection of rules in this paper only applies to round t only. The subjects in the population cannot infer that those more experienced in round t are also more experienced in round $t + 1$.

6.3. Individuals Are Free to Choose Whom to Learn From

Finally, consider the following scenario. There are two groups, and it is known in round t that the subjects belonging to the one group are on average more experienced than those belonging to the other. In each round, with probability μ a subject can choose which group of subjects to learn from and with probability, $1 - \mu$ can only observe someone from the same group. Then clearly each subject will choose in round t , when possible, to observe an experienced. This leads to the dynamics in Section 6.1 with $\lambda = \mu$. The same group remains more experienced in round $t + 1$, and again all subjects wish to observe these subjects. Differences in experience carry over from one round to the next round and persist forever.

6.4. Discussion

We summarize our findings. The difference in experience does not necessarily persist when experienced learn from novices whenever novices learn from the experienced. Initially one might think that novices cannot overtake the experienced. However, on the second inspection, it is clear how they can turn more experienced. This namely happens when the difference in experience is very small. Novices are learning and hence increase the choice of the better action. Experienced on the other hand do not change their action. Hence more of the novices may choose the better action in the next round than the experienced. So, when novices learn from experienced as often as vice versa then the current paper is only applicable the first time experienced and novices learn from each other.

On the other hand, we find that difference in experience persists if everyone learns from those with better experience. In this case, the present paper can be applied in each round. Novices never completely catch up with the experienced.

In this paper, there is no assumption about the size of the difference in experience. However, one would imagine that the difference in experience is initially bounded away from zero. In that case, one expects that the difference in experience persists for several rounds. An analysis of such aspects is however outside the scope of the present paper.

7. Conclusions

Cultural transmission can come in many forms. It can be one generation passing on traits to the next generation, it can be individuals learning from others who have different experiences. In this paper, we caution not to forget that newcomers have their own experiences and old-timers need not all have found the best solutions. The suggestion is that, when passing on info to the newcomers, the success of both sides should be incorporated. For instance, what if the strategy used by the old-timer was not successful, and what if the initial efforts of the newcomer using a different method had been successful? The objective of this paper is to find out whether such additional information should be taken into account, and how this should be done.

Schlag [6] introduced a new methodology and thus opened insights into the value of imitation for social learning. Yet many of the papers that have built on these insights have applied them to a setting where it does not apply. The original paper [6] considers a setting where individuals within the same group are learning from each other how to make good choices. Yet many of the applications consider one group learning from another more experienced group. Those who are making the choice are disjoint from those they are learning from. This paper closes this gap in the literature and investigates exactly this situation, where those choosing belong to a different group than those who they are observing.

The results of this paper are consoling in two senses. First of all, the findings show that it is possible to have a social learning rule that enables learning from other groups in any environment. Second of all, the learning rules (improving rules) that enable learning the better action are the same as they were in the original paper. So those papers applying the within-group learning insights to learning between groups are partially validated. Specifically, rules are improving the within-a-group setting if, and only if, they are for learning from the more experienced.

On the other hand, new aspects arise in how to best learn. Recall for learning within groups that the amount of switching should be kept minimal (within the constraints imposed by the class of improving rules) to reduce variance in finite populations. Yet when learning from more experienced we find that switching should be maximized in order to take advantage of the superior experience of the observed. We learn that all the insights of the original paper by Schlag [6] cannot be automatically applied to other contexts. In particular, the best rule for within-group learning (PIR) does not perform very well when learning from more experienced.

In terms of cultural transmission of traits from the older generation (experienced) to the younger (novices), we obtain the following insights. The older generation should

not change their own behavior. At the same time, they should convince the younger to copy their behavior unless by chance the younger are more successful than they are in which case the younger should continue with their own behavior with a probability that is proportional to how much better they are. In particular, it is not best for them to both adopt the same choice. This is not a contradiction but a consequence of their different objectives, each of them wishes to improve payoffs within their group. This makes the experienced very conservative, worried that the novices are doing the wrong thing so that the experienced do not change their behavior. On the other hand, the novices have the opportunity to learn from more experienced and hence can change.

Given the straightforward analysis in this paper, we observe that the methodology from [6] is easily adapted to investigate new contexts. In this paper, the new context is learning from other groups. Minor adjustments are made to the original methodology as we now include minimax loss as a measurement of performance (as in [11]) and use settings in which extreme payoffs can only be generated to gain insights on the general rules.

We briefly mention other contexts to which the original methodology has been applied. Alós-Ferrer and Schlag [12] show how to learn when some subjects are more likely to be observed than others. They also investigate the setting where individuals do not face independent choice problems. Schlag [13] and Hofbauer and Schlag [14] consider the availability of multiple observations influences learning behavior and success. Schlag [11] allows for learning rules to be time-dependent. All of these investigations concern learning within a group setting.

Numerous gaps in our understanding of social learning remain. How will learning look when subjects are more sophisticated, such as when they have memory and can learn about the situation they face? How should learning between groups be done when there are more than two actions? How can individuals find the right cues to condition their learning on? How can this model be extended when not all face the same environment? Last but not least, how does this theory compare to how real human subjects learn under minimal information? On this matter, note that [3] only provides partial insights as there is not enough variability in the payoffs to be able to get an understanding of the details of the learning rule.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

Note

- ¹ When payoffs are known to belong to $[\alpha, \omega]$, for some $\alpha < \omega$, replace each $x \in [\alpha, \omega]$ by $(x - \alpha)/(\omega - \alpha)$ and then proceed as if payoffs belong to $[0, 1]$.

References

1. Henrich, J.; Gil-White, F. The Evolution of Prestige: Freely Conferred Status as a Mechanism for Enhancing the Benefits of Cultural Transmission. *Evol. Hum. Behav.* **2001**, *22*, 1–32. [[CrossRef](#)]
2. Henrich, J.; Broesch, J. On the Nature of Cultural Transmission Networks: Evidence from Fijian Villages for Adaptive Learning Biases. *Philos. R. Soc.* **2011**, *366*, 1139–1148. [[CrossRef](#)] [[PubMed](#)]
3. Apesteguia, J.; Huck, S.; Oechssler, J. Imitation—Theory and Experimental Evidence. *J. Econ. Theory* **2007**, *136*, 217–235. [[CrossRef](#)]
4. Galef, B.G.; Laland, K.N. Social Learning in Animals: Empirical Studies and Theoretical Models. *BioScience* **2005**, *55*, 489–499. [[CrossRef](#)]
5. Bisin, A.; Verdier, T. The Economics of Cultural Transmission and the Dynamics of Preferences. *J. Econ. Theory* **2001**, *97*, 298–319. [[CrossRef](#)]
6. Schlag, K.H. Why Imitate, and if so, How? A Boundedly Rational Approach to Multi-Armed Bandits. *J. Econ. Theory* **1998**, *78*, 130–156. [[CrossRef](#)]
7. Dean, L.G.; Vale, G.L.; Laland, K.N.; Flynn, E.; Kendal, R.L. Human Cumulative Culture: A Comparative Perspective. *Biol. Rev.* **2014**, *89*, 284–301. [[CrossRef](#)] [[PubMed](#)]
8. Savage, L.J. The Theory of Statistical Decision. *J. Amer. Stat. Assoc.* **1951**, *46*, 55–67. [[CrossRef](#)]

9. Milnor, J. Games Against Nature. In *Decision Processes*; Thrall, R.M., Coombs, C.H., Davis, R.L., Eds.; John Wiley & Sons: New York, NY, USA, 1954.
10. Pike, T.W.; Kendal, J.R.; Rendell, L.E.; Laland, K.N. Learning by Proportional Observation in a Species of Fish. *Behav. Ecol.* **2010**, *21*, 570–575. [[CrossRef](#)]
11. Schlag, K.H. Imitation and Minimax Regret, Mimeo. Available online: <http://homepage.univie.ac.at/karl.schlag/research/imitation/regretim47.pdf> (accessed on 3 December 2004).
12. Alós-Ferrer, C.; Schlag, K.H. Imitation and Social Learning. In *The Handbook of Rational and Social Choice*; Anand, P., Pattanaik P., Puppe, C., Eds.; Oxford University Press: Oxford, UK, 2009.
13. Schlag, K.H. Which one should I imitate? *J. Math. Econ.* **1999**, *31*, 493–522. [[CrossRef](#)]
14. Hofbauer, J.; Schlag, K.H. Sophisticated Imitation in Cyclic Games. *J. Evol. Econ.* **2000**, *10*, 523–543. [[CrossRef](#)]