

Trade-off between learning and exploitation: The Pareto-optimal versus evolutionarily stable learning schedule in cumulative cultural evolution

Joe Yuichiro Wakano*, Chiaki Miura

Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University, 164-8525 Tokyo, Japan

ARTICLE INFO

Article history:

Available online 14 September 2013

Keywords:

Cumulative culture
Pareto-optimal strategy
Evolutionarily stable strategy
Two-stage life history
Public goods game

ABSTRACT

Inheritance of culture is achieved by social learning and improvement is achieved by individual learning. To realize cumulative cultural evolution, social and individual learning should be performed in this order in one's life. However, it is not clear whether such a learning schedule can evolve by the maximization of individual fitness. Here we study optimal allocation of lifetime to learning and exploitation in a two-stage life history model under a constant environment. We show that the learning schedule by which high cultural level is achieved through cumulative cultural evolution is unlikely to evolve as a result of the maximization of individual fitness, if there exists a trade-off between the time spent in learning and the time spent in exploiting the knowledge that has been learned in earlier stages of one's life. Collapse of a fully developed culture is predicted by a game-theoretical analysis where individuals behave selfishly, e.g., less learning and more exploiting. The present study suggests that such factors as group selection, the ability of learning-while-working ("on the job training"), or environmental fluctuation might be important in the realization of rapid and cumulative cultural evolution that is observed in humans.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

In the evolution of modern humans, innovative stone tools might have played an important role. Advanced and complex lithic industries could never have been invented by a single individual without building on knowledge acquired from others. They are the results of cumulative cultural evolution in which knowledge is inherited from parental generations and passed on to descendant generations. However, if all individuals simply copied what was already known, there would be no advancement in technology. Contribution to culture by improving on the preceding knowledge, such as a discovery of a new adaptive use or form of stone tools, is also crucial for cultural evolution. Thus, it is an important problem to determine what learning strategy can support the cumulative culture that develops the fastest: when and how should an individual perform learning during one's life?

The balance between the time allocated to learning and the time allocated to the exploitation of the learned knowledge is then the next question. To maximize the accumulation rate of culture, one should spend sufficient time to absorb the existing knowledge in the parental generation and then spend all the rest of the lifetime to improve the knowledge by individual learning. However, such a

learning schedule does not necessarily maximize the fitness of an individual, i.e., the expected number of biological offspring.

Although many modeling studies have been performed (e.g., Feldman and Cavalli-Sforza, 1984, Boyd and Richerson, 1985, Rogers, 1988, Wakano et al., 2004, Aoki et al., 2005, Rendell et al., 2010), researchers have started surveying the evolution of learning as life-history strategies only recently (Aoki, 2010; Aoki et al., 2012; Lehmann et al., 2013). A learning schedule considered in this paper is a broad developmental pattern over a single individual's lifetime in the differential use of IL and SL with regard to behaviors that may take a long time to be acquired. It differs from the ordered application of IL and SL in skill acquisition at any one time, which may occur repeatedly during one's lifetime (e.g., Boyd and Richerson, 1985, Borenstein et al., 2008). One of the well-known studies on a learning schedule is the Social Learning Tournament (Rendell et al., 2010). However, the Tournament included so many factors and the analytic treatment is impossible. Enquist et al. (2007) compared the performances of pure SL strategy, pure IL strategy, and "critical social learner" who performs individual learning (IL) only if social learning (SL) had failed to achieve an OK solution. They showed that critical social learner outperforms the pure strategies and thus evolves. Aoki et al. (2012) performed a more exhaustive study of a two-stage model in which any mixture of SL and IL is allowed in each stage and the environment may fluctuate. They showed that in a constant environment pure SL followed by pure IL is an evolutionarily stable strategy when the efficiency of SL is not too low.

* Correspondence to: 4-21-1 Nakano, Nakano-ku, 164-8525 Tokyo, Japan.
E-mail address: joe@meiji.ac.jp (J.Y. Wakano).

Table 1
Parameters and variables.

Variable	Classification	Meaning	Range
z	Dynamical variable	Cultural level	$0 \leq z$
α	Parameter	Gain of cultural level per unit effort of IL	$0 < \alpha < 1$
β	Parameter	Efficiency of SL	$0 < \beta < 1$
u_0	Life history strategy	Allocation of learning effort to IL in 1st stage	$0 \leq u_0 \leq 1$
u_1	Life history strategy	Allocation of learning effort to IL in 2nd stage	$0 \leq u_1 \leq 1$
v	Life history strategy	Allocation of time to learning efforts (IL+SL) in 2nd stage	$0 \leq v \leq 1$

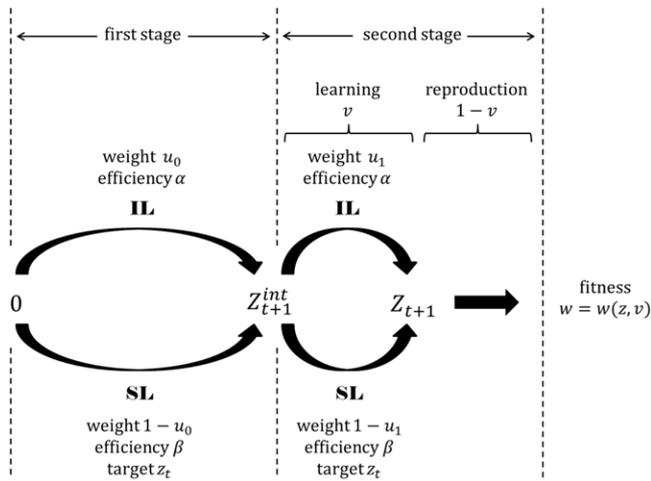


Fig. 1. Schematic illustration of staged life learning schedule in this model.

In a series of studies on the evolutionarily stable learning schedule, Lehmann et al. (2013) studied a case when trade-off between learning and exploiting exists. This model is far more complex as it includes many factors (i.e., continuous time axis, horizontal transmission, and environmental changes). One of their results is that, when environmental change is negligibly rare, the equilibrium of the cultural level realized by the evolutionarily stable learning schedule is not very high compared to the level realized by the pure IL strategy. This result is counter-intuitive because improvement achieved by IL should accumulate over generations when environmental change is rare (a sufficiently small decay rate of information), and thus we naturally expect the cultural level to increase to reach the maximum level that is transmittable to the next generation (limited only by the efficiency of SL). However, the predicted cultural level was far below this level.

The goal of the present study is to clarify the factors that facilitate or suppress the cumulative cultural evolution over generations. When and how does the maximization of individual fitness (natural selection) also maximize the cultural level of the society after a sufficient number of generations? For this purpose, we first propose a relatively simple model with a two-stage life history. Second, we perform a full mathematical analysis of the Pareto-optimal strategy and the evolutionarily stable strategy (ESS) when there is no trade-off between learning and exploiting. Third, an analysis of the case with the trade-off follows. We also show an analysis of a case when the fitness function is an exponential function of the cultural level. Finally we discuss the similarity of our model to models of public goods game.

2. Model and results

2.1. Model

Our model is a modification and extension of Aoki et al. (2012). Here we assume a constant environment. Let z denote the cultural

level, e.g., complexity of stone tool or the industry. We assume that the cultural level is represented by a one-dimensional variable. For example, the Acheulean lithic culture might be represented by $z = 10$, while the Aurignacian lithic culture might be represented by $z = 20$. We use the terminologies ‘culture’, ‘cultural level’, and ‘information (level)’, interchangeably. For simplicity, we assume that all individuals with the same cultural level carry the same information. Thus, social learning from another person with the same z value brings no new information.

Life history assumptions are summarized in Fig. 1. Every individual is born with no information ($z = 0$). For simplicity, we assume two learning stages in one’s life. In each stage, an individual can perform IL, SL, or any mixture of IL and SL. What is learned in the first stage and in the second stage are added and become the mature phenotype, that is the cultural level of an adult. Adults reproduce offspring, act as exemplars for the next generation, and die.

When an individual performs IL, she gains information α per unit effort. For SL, we only consider oblique transmission across generations. When performing SL, a young individual tries to copy the mature phenotype in the parental generation. The increase of cultural level per unit effort of SL is proportional to the difference in the cultural level of the social learner and that of a randomly chosen exemplar in the previous generation. The coefficient is β . See Table 1 for a summary of parameters and variables.

Allocation of effort is the evolving strategy in this study. In the first stage, an individual performs IL with effort u_0 and SL with effort $1 - u_0$. In the second stage, the individual allocates the learning effort to IL and SL with a ratio $u_1 : 1 - u_1$. During this second stage, the individual can reproduce offspring. Reproductive success is determined by two factors; cultural level and the effort put into reproduction. Most animals including human can learn something before the body sexually matures. It also seems adaptive to allocate more learning effort in the earlier stage of life than allocating learning effort equally in all life stages. Thus, we assume that an individual does not reproduce in the first stage. In the second stage, the fraction v of time is dedicated to learning, and the fraction $1 - v$ to exploiting the knowledge for reproduction. Thus in the second stage, efforts $u_1 v$, $(1 - u_1)v$, and $1 - v$ are allocated to IL, SL, and exploitation, respectively. The life history strategy is represented by the triplet (u_0, u_1, v) .

The model assumptions above are denoted by the following recursion that describes the dynamics of cultural level z_t in generation t ;

$$\begin{aligned} z_{t+1}^{\text{INT}} &= u_0 \alpha + (1 - u_0) \beta z_t \\ z_{t+1} &= z_{t+1}^{\text{INT}} + v [u_1 \alpha + (1 - u_1) \beta (z_t - z_{t+1}^{\text{INT}})], \end{aligned} \quad (1)$$

where all individuals adopt the same life-history strategy (Aoki et al., 2012). Our model slightly differs from the previous study (Aoki et al., 2012) as we assume a constant rate of improvement by IL while the previous study assumes that the improvement by IL is proportional to the difference between the current z value and the target value.

2.2. Maximization of cultural level

According to (1), z_t approaches an equilibrium value, z^* , after a sufficient number of generations. In this section, before we consider reproductive success (fitness), we study which learning schedule maximizes z^* .

2.2.1. Pareto-optimal learning schedule (for society)

For the pure IL strategy, i.e. $(u_0, u_1) = (1, 1)$, the cultural equilibrium is

$$z^* = \alpha(1 + v) \quad (2)$$

and for the SL–IL strategy, $(u_0, u_1) = (0, 1)$, it is

$$z^* = \frac{v\alpha}{1 - \beta}. \quad (3)$$

In general, the cultural equilibrium can be written as $z^*(u_0, u_1, v)$. It is a fixed point of the recursion (1) and satisfies

$$z^* = (u_0 + vu_1)\alpha + [(1 - u_0)\beta z^* + (1 - u_1)v\beta \times \{z^* - u_0\alpha - (1 - u_0)\beta z^*\}] \quad (4)$$

whose solution is

$$z^*(u_0, u_1, v) = \alpha \frac{vu_1 + u_0(1 - v\beta(1 - u_1))}{(1 - \beta(1 - u_0))(1 - v\beta(1 - u_1))}. \quad (5)$$

The learning schedule (u_0, u_1, v) that maximizes $z^*(u_0, u_1, v)$ is the Pareto-optimal learning schedule (the optimal strategy for society). Since

$$\frac{\partial z^*}{\partial u_1} = \frac{v\alpha(1 - v\beta)}{(1 - \beta(1 - u_0))(1 - v\beta(1 - u_1))^2} > 0 \quad (6)$$

and

$$\frac{\partial z^*}{\partial v} = \frac{\alpha u_1}{(1 - \beta(1 - u_0))(1 - v\beta(1 - u_1))^2} > 0, \quad (7)$$

we find that $u_1 = 1$ and $v = 1$ maximize z^* . The partial derivative

$$\frac{\partial z^*}{\partial u_0}(u_0, 1, 1) = \frac{\alpha(1 - 2\beta)}{(1 - \beta(1 - u_0))^2} \quad (8)$$

changes sign at $\beta = 1/2$.

Thus, pure IL with the maximum learning effort in the second stage ($u_1 = 1, v = 1$)—this entails no reproduction—always maximizes the cultural level. Moreover, the IL–IL strategy is Pareto-optimal when the efficiency of SL satisfies $\beta < 1/2$. Otherwise, the SL–IL strategy is Pareto-optimal.

2.2.2. Optimal learning schedule for an individual (ES learning schedule)

An evolutionarily stable (ES) learning schedule is a different concept from the Pareto-optimal learning schedule. It is defined as such a strategy that no other mutant can achieve a higher payoff than the residents if the residents are adopting the ES schedule. For a constant resident cultural level z^R , the recursion for the cultural level z_t^M of rare mutants adopting the strategy (u_0^M, u_1^M, v^M) is given by

$$z_{t+1}^M = u_0^M \alpha + (1 - u_0^M) \beta z^R \quad (9)$$

$$z_{t+1}^M = z_{t+1}^M + v^M [u_1^M \alpha + (1 - u_1^M) \beta (z^R - z_{t+1}^M)],$$

and thus the equilibrium of the cultural level of the mutant is

$$z^M = \{\alpha u_0^M + \beta(1 - u_0^M)z^R\} + v^M \left[\alpha u_1^M + \beta(1 - u_1^M)(z^R - \{\alpha u_0^M + \beta(1 - u_0^M)z^R\}) \right]. \quad (10)$$

Aoki et al. (2012) have provided a detailed justification of the usage of (10) to derive an ESS (quasi-stationary approximation of cultural dynamics). Using the method in Aoki et al. (2012) and also in this paper in Section 2.3.2, it is a straightforward calculation to show that z^M is maximized by the IL–IL (resp. SL–IL) strategy when residents adopt the IL–IL (resp. SL–IL) strategy for $\beta < 1/2$ (resp. $\beta > 1/2$). Thus, the Pareto-optimal strategy is also the ES learning strategy.

The results above show that if the objective is to gain a higher cultural level, maximization at a group level and at an individual level yield the same optimal strategy. No social dilemma exists in this case. Cumulative cultural evolution, supported by the SL–IL strategy, is the expected evolutionary outcome when the SL efficiency is high. This result is qualitatively the same as Aoki et al. (2012).

2.3. Maximization of fitness

Here we consider fitness as the objective function to be maximized. Reproductive success, or fitness, should be an increasing function of the cultural level (z) that a focal individual has learnt. It should also be a decreasing function of the effort allocated to learning (v), as we assume a trade-off between learning and exploitation. We first consider the following simple fitness function

$$w(z, v) = (1 - v)z \quad (11)$$

where the reproductive success is given by a product of the cultural level and the time spent in exploiting the information.

2.3.1. Pareto-optimal life history (for society)

To maximize the fitness for society, we simply look for the triplet (u_0, u_1, v) that maximizes the fitness function (11) evaluated at the cultural equilibrium given by (5). Recall that the cultural level at equilibrium is also a function of v , i.e., $z^* = z^*(u_0, u_1, v)$. Formally,

$$(u_0^{\text{OPT}}, u_1^{\text{OPT}}, v^{\text{OPT}}) \equiv \arg \max_{u_0, u_1, v} w(z^*(u_0, u_1, v), v). \quad (12)$$

Since

$$\frac{\partial w}{\partial u_1} = (1 - v) \frac{\partial z^*}{\partial u_1} > 0, \quad (13)$$

we find that $u_1 = 1$ always maximizes fitness. Next, using

$$\frac{\partial w}{\partial v}(u_0, 1, v) = \frac{\alpha(1 - u_0 - 2v)}{1 - \beta(1 - u_0)}, \quad (14)$$

we find that $v = (1 - u_0)/2$ always maximizes fitness, which is

$$w\left(u_0, 1, \frac{1 - u_0}{2}\right) = \frac{\alpha(1 + u_0)^2}{4(1 - \beta + \beta u_0)}. \quad (15)$$

The maximum of this function in the interval $u_0 \in [0, 1]$ depends on β . Either $u_0 = 0$ or $u_0 = 1$ gives the maximum, where $w(1, 1, 0) = \alpha$ and $w(0, 1, 1/2) = \alpha/(4 - 4\beta)$ respectively.

Case (1) $\beta < \frac{3}{4}$.

In this case,

$$(u_0^{\text{OPT}}, u_1^{\text{OPT}}, v^{\text{OPT}}) = (1, 1, 0) \quad (16)$$

is the solution, i.e. the pure IL strategy with no learning in the second stage. The corresponding cultural level and fitness are

$$z^{\text{OPT}} = \alpha, \quad w^{\text{OPT}} = \alpha. \quad (17)$$

When the efficiency of social learning is relatively low, it is better for society to abandon cumulative cultural evolution.

Case (2) $\beta > \frac{3}{4}$.
In this case,

$$(u_0^{\text{OPT}}, u_1^{\text{OPT}}, v^{\text{OPT}}) = \left(0, 1, \frac{1}{2}\right) \quad (18)$$

is the solution, i.e. pure SL in the first stage and then pure IL in the second stage. Half of the time in the second stage is allocated to reproduction. The corresponding cultural level and fitness are

$$z^{\text{OPT}} = \frac{\alpha}{2-2\beta}, \quad w^{\text{OPT}} = \frac{\alpha}{4-4\beta}. \quad (19)$$

When the efficiency of social learning is high, society can enjoy the benefit of cumulative cultural evolution. Note that

$$\lim_{\beta \uparrow 1} z^{\text{OPT}} = \infty, \quad \lim_{\beta \uparrow 1} w^{\text{OPT}} = \infty \quad (20)$$

which means that the cultural level and also fitness can be infinitely large if the efficiency of social learning is almost perfect. This corresponds to the situation where the accumulation of information over generations continues endlessly.

2.3.2. Optimal life history for an individual (evolutionarily stable life history)

Here we show that the ESS allows neither social learning nor cumulative cultural evolution. Substituting (10) into (11), the mutant fitness function is given by

$$\begin{aligned} w(z^M, v) &= w^M(u_0, u_1, v) \\ &= (1-v) \left[\{\alpha u_0 + \beta(1-u_0)z^R\} \right. \\ &\quad \left. + v[\alpha u_1 + \beta(1-u_1)(z^R - \{\alpha u_0 + \beta(1-u_0)z^R\})] \right] \end{aligned} \quad (21)$$

where superscript M for the arguments is omitted for simplicity. First, let us assume $\beta > 3/4$ and check whether the Pareto-optimal schedule $(0, 1, 1/2)$ is an ESS or not. Using $z^R = \frac{\alpha}{2-2\beta}$ (see (19)), we have

$$\frac{\partial w^M}{\partial v} \left(0, 1, \frac{1}{2}\right) = -\frac{\alpha\beta}{2(1-\beta)} < 0 \quad (22)$$

which shows that a mutant can gain higher fitness by decreasing her learning effort v . Thus, the Pareto-optimal schedule is *not* an ESS.

Which learning schedule is an ESS? To derive an ESS, we first assume that an ESS exists and realizes the cultural equilibrium $z^R = z^{\text{ESS}}$. When residents adopt this ESS, the derivative of the mutant fitness function with respect to u_0 evaluated at the assumed ESS is

$$\frac{\partial w^M}{\partial u_0}(u_0, u_1, v) = (1-v)(\alpha - \beta z^{\text{ESS}})(1 - v\beta(1-u_1)). \quad (23)$$

Note that the value of z^{ESS} depends on the assumed ESS but it remains constant when we calculate the fitness of rare mutants. It is obvious that $v = 1$ cannot be an ESS because this strategy yields $w = 0$, and hence the mutant strategy $(1, 1, 0)$ can invade. Thus, we can restrict our search for an ESS to the region $v < 1$. Eq. (23) shows that the best mutant strategy depends on the sign of $\alpha - \beta z^{\text{ESS}}$.

Case (1) $z^{\text{ESS}} > \alpha/\beta$.

We prove that this case is impossible. By assumption, $\frac{\partial w^M}{\partial u_0} < 0$ always holds and $u_0 = 0$ maximizes mutant fitness. Then we have

$$\frac{\partial w^M}{\partial u_1}(0, u_1, v) = v(1-v)(\alpha - \beta(1-\beta)z^{\text{ESS}}). \quad (24)$$

When residents adopt $v = 0$, (5) assures that $z^{\text{ESS}} \leq \alpha$ (the equality holds only when $u_0 = 0$). Thus, in Case 1, it is enough

to consider the region $0 < v < 1$. If $z^{\text{ESS}} > \alpha/\beta(1-\beta)$ then $u_0 = u_1 = 0$ maximizes mutant fitness. However, $(0, 0, v)$ is the SL–SL strategy and it yields the cultural equilibrium $z^* = 0$ when they are residents. Thus $z^* > \alpha/\beta(1-\beta)$ is impossible as an ESS result. For $z^{\text{ESS}} < \alpha/\beta(1-\beta)$, $u_1 = 1$ maximizes mutant fitness. If a mutant adopts the strategy $(u_0, u_1) = (0, 1)$,

$$\frac{\partial w^M}{\partial v}(0, 1, v) = \alpha - \beta z^{\text{ESS}} - 2\alpha v < 0 \quad (25)$$

holds by assumption, and $v = 0$ maximizes mutant fitness. However, the strategy $(u_0, u_1, v) = (0, 1, 0)$ performs only SL and thus yields $z^* = 0$ at the cultural equilibrium. Thus, it cannot be an ESS with $z^{\text{ESS}} > \alpha/\beta$.

Case (2) $z^{\text{ESS}} < \alpha/\beta$ and $v > 0$.

We prove that this case is also impossible. In this case, $u_0 = 1$ maximizes mutant fitness and we have

$$\frac{\partial w^M}{\partial u_1}(1, u_1, v) = v(1-v)(\alpha + \alpha\beta - \beta z^{\text{ESS}}) \quad (26)$$

which is positive by assumption. Thus, $(u_0, u_1) = (1, 1)$ maximizes mutant fitness. We have

$$\frac{\partial w^M}{\partial v}(1, 1, v) = -2v\alpha \quad (27)$$

which is negative. Thus, a mutant with a smaller v value can always invade. Since we assume $v > 0$, there always exists a mutant that has a smaller v value than residents. Hence there is no ESS in Case 2.

Case (3) $z^{\text{ESS}} < \alpha/\beta$ and $v = 0$.

From (23), $u_0 = 1$ maximizes mutant fitness. As $v = 0$, (21) entails that w^M does not depend on u_1 , and hence u_1 can take an arbitrary value. We choose $u_1 = 1$. Thus, $(u_0, u_1, v) = (1, 1, 0)$ maximizes mutant fitness. When residents adopt this strategy, from (27), a mutant with $v > 0$ cannot invade. The cultural equilibrium is $z^* = \alpha$, which is consistent with $z^{\text{ESS}} < \alpha/\beta$. Thus, we finally obtain

$$(u_0^{\text{ESS}}, u_1^{\text{ESS}}, v^{\text{ESS}}) = (1, 1, 0). \quad (28)$$

The above derivation shows that the only ES learning schedule in our model is $(u_0, u_1, v) = (1, 1, 0)$, irrespective of parameter values (as long as they satisfy the range shown in Table 1). The ESS performs pure IL only in the first stage. The whole duration of the second stage is devoted to exploitation, and no SL is performed by the ESS. We have also shown that no other ESS exists in our model. The corresponding cultural level and fitness are

$$z^{\text{ESS}} = \alpha, \quad w^{\text{ESS}} = \alpha. \quad (29)$$

See Fig. 2 for the comparison between the fitness realized by the Pareto-optimal life history strategy and that by the ES life history strategy. Clearly, a considerable difference exists when the efficiency of social learning is high. For the ESS, fitness is only determined by the efficiency of individual learning, while the fitness can be much larger by utilizing high efficiency of social learning in the maximization of cultural level at a group level.

2.3.3. A model with a more general fitness function

Our result is not an artifact of our choice of the simple fitness function (11). For the fitness function $w = (1-v)f(z)$ with $f(z) > 0$ and $f'(z) > 0$, the sign of $\frac{\partial w^M}{\partial u_0}$ is still determined by $\alpha - \beta z^{\text{ESS}}$. (This is rather trivial because this quantity determines whether IL or SL is more efficient to obtain information in the first stage). The sign of $\frac{\partial w^M}{\partial v}$ is negative and our argument still holds if

$$\left. \frac{d}{dz} \log(f(z)) \right|_{z=v\alpha+\beta z^{\text{ESS}}} < \frac{1}{(1-v)\alpha}. \quad (30)$$

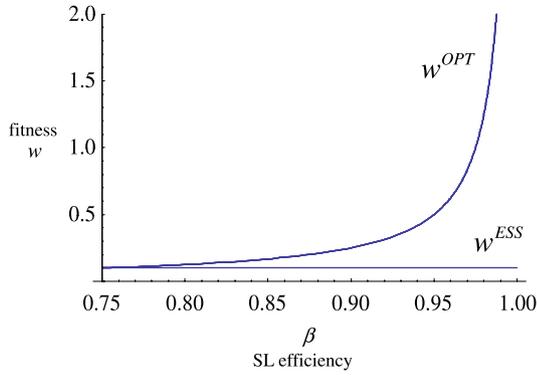


Fig. 2. The possible maximum value of fitness (w^{OPT}) and the value of fitness realized by the ES learning schedule (w^{ESS}) as functions of the SL efficiency β . See (19) and (29) in the main text. Parameters: $\alpha = 0.1$.

The pure-IL strategy with no learning in the second stage is the unique ESS if (30) holds. In particular, when $f(z) = \exp(Cz)$ we cannot expect cultural accumulation if $C < \frac{1}{(1-v)\alpha}$. As $0 \leq v \leq 1$, C must be at least larger than $\frac{1}{\alpha}$ for a strategy with SL to evolve.

It is difficult to derive the Pareto-optimal and ES learning schedules for a general fitness function. If we assume $f(z) = \exp(Cz)$ and if we restrict ourselves to search only a limited set of strategies of the form $(0, 1, v)$, then we can show

$$v^{OPT} = 1 - \frac{1 - \beta}{C\alpha}, \quad w^{OPT} = \frac{1 - \beta}{C\alpha} \exp\left[\frac{C\alpha + \beta - 1}{1 - \beta}\right], \quad (31)$$

and

$$v^{ESS} = 1 - \frac{1}{C\alpha}, \quad w^{ESS} = \frac{1}{C\alpha} \exp\left[\frac{C\alpha - 1}{1 - \beta}\right]. \quad (32)$$

This suggests that cumulative cultural evolution might be maintained as an evolutionary outcome when C is large. However, the ratio

$$\frac{w^{OPT}}{w^{ESS}} = (1 - \beta) \exp\left[\frac{\beta}{1 - \beta}\right] \quad (33)$$

goes to infinity as $\beta \rightarrow 1$. Thus, although SL is part of the ES schedule, the drop in fitness value from the Pareto-optimal strategy to the ES strategy is significantly large. The result suggests that the cumulative cultural evolution is difficult to achieve even when the fitness depends exponentially on the cultural level.

2.4. Numerical illustration of evolutionary paths

As we have the derivatives of fitness functions, we can numerically illustrate the vector field showing the direction of selection in (u_0, u_1, v) -space. We numerically found that u_1 was always positively selected, and a numerical example of the vector field in the (u_0, v) -plane at $u_1 = 1$ is shown in Fig. 3. We can also gain some insight on the three-dimensional evolutionary pathway by numerically integrating the following canonical equation (e.g., Wakano and Aoki, 2006)

$$\begin{aligned} \frac{du_0}{d\tau} &= \frac{\partial w^M}{\partial u_0}, \\ \frac{du_1}{d\tau} &= \frac{\partial w^M}{\partial u_1}, \\ \frac{dv}{d\tau} &= \frac{\partial w^M}{\partial v}, \end{aligned} \quad (34)$$

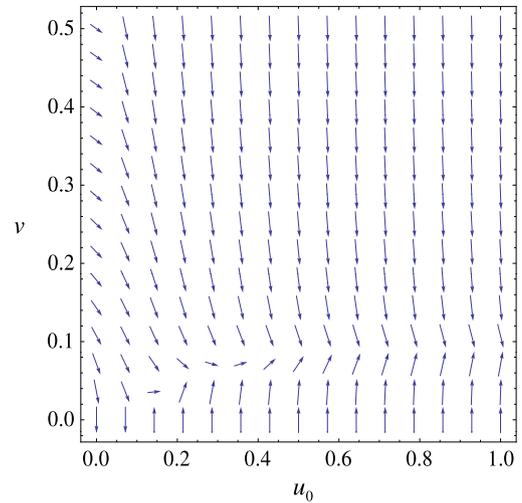


Fig. 3. The vector field showing the direction of selection in (u_0, v) -plane at $u_1 = 1$. The arrow lengths are normalized to the constant length. Parameters: $\alpha = 0.1, \beta = 0.9$.

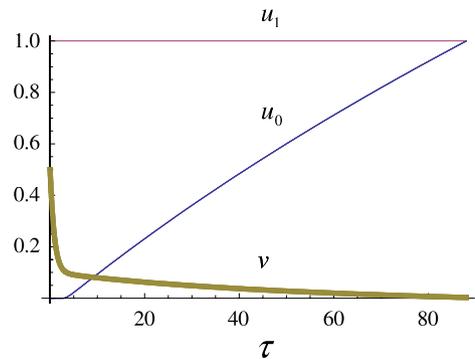


Fig. 4. A numerical example of the evolutionary dynamics according to (34) in the main text. Fitness function is $w = (1 - v)z$. Parameters: $\alpha = 0.1, \beta = 0.9$.

where τ represents a slow timescale of genetic evolution. An example is shown in Fig. 4 illustrating the “collapse”, which is defined as a drastic decrease in mean fitness (given by (11)) from the Pareto-optimal level to the ES level as a result of competition among individuals. Initially, residents adopt the strategy $(0, 1, 1/2)$ which maximizes the cultural level. Then, a mutant with smaller effort in learning (v) invades. As v decreases, the cultural level also decreases because less effort is allocated to IL. In other words, individuals no longer provide a sufficient amount of new information that balances the loss of information in social transmission due to non-perfect SL efficiency. As v drops to a certain level, the cultural level approaches the threshold (α/β) below which IL instead of SL is the best learning strategy in the first stage. Then the collapse continues as the information held by the parental generation becomes more and more useless for the next generation. Numerical results (Fig. 4) suggest that the first collapse of learning effort (v) occurs very quickly, while it takes some time until the learning effort is reduced to zero and SL is completely abandoned. Cumulative cultural dynamics collapses at the initial rapid drop of learning effort. Since individuals cannot reproduce in the first stage (by model assumption), they perform only IL in the first stage and all effort in the second stage is allocated to exploitation.

The dynamics for the case $f(z) = \exp(Cz)$ is shown in Fig. 5. The evolutionary scenario is similar except that the learning effort v does not evolve to reach zero. However, the cultural level z drastically drops as v decreases. Since fitness is an exponential function of z , the collapse is even more prominent in fitness values.

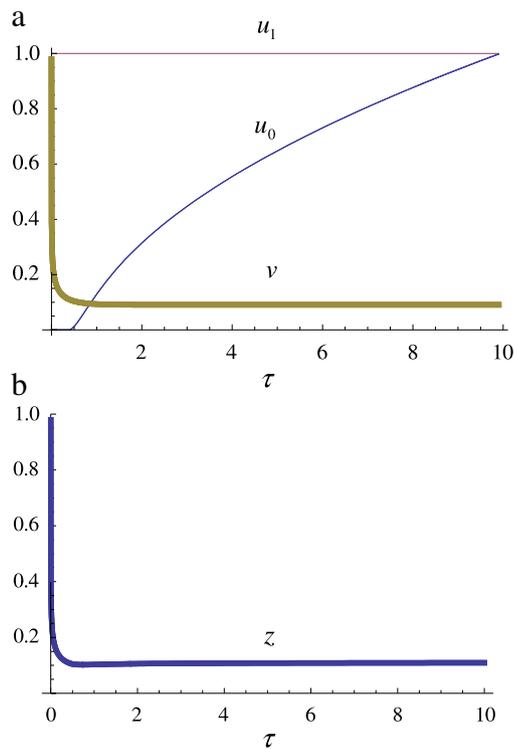


Fig. 5. A numerical example of the evolutionary dynamics according to (34) in the main text. Fitness function is $w = (1 - v) \exp(Cz)$. (a) Evolutionary dynamics of the strategies. (b) The dynamics of the quasi-equilibrium value of cultural level during the slow genetic timescale τ . Parameters: $\alpha = 0.1$, $\beta = 0.9$, $C = 11$.

3. Discussion

3.1. Summary

We have analyzed a two-stage life history model of learning schedule. Allocation of lifetime to IL, SL and exploitation (reproduction) at each stage defines a strategy. We studied cases with and without trade-off between learning and exploitation. In both cases, the optimal strategy for society is to perform SL then IL, when the efficiency of SL is high. If the SL efficiency is almost perfect, the equilibrium of the cultural level can increase to almost infinity. However, when a trade-off between learning and exploitation exists, we show that social learning cannot evolve as part of an ES learning schedule in a simple model with linear dependence of the fitness on the cultural level. Thus, cumulative cultural evolution is impossible, although individuals would enjoy a very large fitness if they all cooperated by performing IL after SL to maintain culture. If fitness increases exponentially with cultural level, it is suggested that the result qualitatively remains the same although a certain magnitude of SL is performed by the ES schedule.

The reason is intuitively explained as follows. When individuals are allowed to increase their exploiting effort (for reproduction) at the expense of decreased individual learning effort, there exists a temptation to allocate more effort to exploitation. The temptation is stronger if society's cultural level is high, because an individual can add only little by her own individual learning. It is best to socially learn in the first half of life, and spend all the rest of her life in exploitation. But if everybody does this, nobody contributes new knowledge and the collapse occurs. The increase of fitness by increased exploiting effort is a private component of fitness. On the other hand, new knowledge contributed by IL is shared by everyone and carried over generations, so it can be seen as the public component of fitness. The situation is well-explained by

considering IL as “information-producer” and SL as “information-scrounger” (Kameda and Nakanishi, 2002). It is very similar to a public goods game where public goods is information that society carries. Without another exogenous mechanism such as kin selection or spatial structure, social dilemma always leads to total defection, which is parallel to the situation in our model where all individuals perform maximum exploitation and no learning in the second stage.

3.2. Comparison with previous studies

The present result without trade-off is qualitatively the same as in Aoki et al. (2012), although the models are slightly different. (We assume additive effect of IL and no environmental fluctuation.) Lehmann et al. (2013) have studied an evolutionary model of learning schedule and have shown that the cultural explosion cannot be an evolutionary outcome. They have studied a more realistic model with continuous time axis and possibility of horizontal social learning and achieved some analytic results. The present study is a minimal extension of Aoki et al. (2012)'s two-stage model, which is sufficient to show that trade-off between learning and exploiting can drastically change the theoretical prediction. Such a trade-off is also assumed in Lehmann et al. (2013)'s model, and the results are parallel.

In the field of computer simulation, the Social Learning Tournament (Rendell et al., 2010) has allowed three actions corresponding to IL, SL, and exploitation in a discretely staged life. Thus, the motivation is fairly similar to the present study. However, they have observed relatively heavy reliance on social learning. The discrepancy might be related to many factors that the Tournament considers. In the Tournament, a fixed payoff is assigned to each behavior. Thus, although life history is modeled, it is a game in which players seek the best behavior by many stages of learning. Players cannot improve payoff values themselves by IL. This might be the primary reason why SL evolved in the Tournament. The existence of environmental fluctuation might be another reason.

In our model, there exist a diversity of strategies, but the cultural level is assumed to be a one-dimensional trait. Lehmann and Wakano (2013) have shown that learning dynamics for a multidimensional trait could have a very different theoretical result. We have also assumed that individuals learn nothing during exploitation. Real humans learn while working, perhaps inevitably. If learning can be performed at no cost when individuals are exploiting their knowledge, then cumulative cultural evolution seems more likely to occur.

3.3. Relation to public goods games

In order to avoid the collapse and to maintain a high cultural level by cumulative cultural evolution, some mechanisms might be necessary that decrease or eliminate the private component of fitness. One of the most plausible mechanisms in real human populations might be the social interaction among relatives where they share their reproductive success. Experimental studies show that contemporary hunter-gatherers have band structures (Hill et al., 2011; Apicella et al., 2012), and most exchange of information (particularly transmission across generations that we focus on in this paper) might be performed within a band. The present model, and also many models of evolution of learning strategies, uses an evolutionary game theory that basically assumes competition among players. Among fully competitive or hostile players, social learning is similar to “information stealing”. Naively, we do not expect flourishing culture in such a situation. Many previous studies have assumed no private component of fitness (cultural level is identical to individual fitness) so cumulative cultural evolution has been

predicted. The present study suggests the importance of cooperation among members in the same society for evolution of a learning strategy that can support cumulative culture (see also Henrich, 2012). Evolution of cooperation has been extensively studied in genetic evolution (e.g., Hamilton, 1964. Taylor and Irwin, 2000, Ohtsuki et al., 2006, Wakano et al., 2009), but combining evolution of cooperation with evolution of learning creates a very complicated problem in a gene–culture co-evolution framework. Recent theoretical developments in inclusive fitness theory, including the calculation of relatedness across generations (Lehmann, 2010) and relatedness between gene and meme (Kobayashi and Ohtsuki, personal communication), might usefully inform future work along these lines (see also Lehmann et al., 2010, Lehmann and Rousset, 2010).

Acknowledgments

We thank K. Aoki for many fruitful comments and suggestions. This research was supported in part by Monbukagakusho grant 22101004.

References

- Aoki, K., 2010. Evolution of the social-learner–explorer strategy in an environmentally heterogeneous two-island model. *Evolution* 64, 2575–2586.
- Aoki, K., Wakano, J.Y., Feldman, M.W., 2005. The emergence of social learning in a temporally changing environment: a theoretical model. *Curr. Anthropol.* 46, 334–340.
- Aoki, K., Wakano, J.Y., Lehmann, L., 2012. Evolutionarily stable learning schedules and cumulative culture in discrete generation models. *Theor. Popul. Biol.* 81, 300–309.
- Apicella, C.L., Marlowe, F.W., Fowler, J.H., Christakis, N.A., 2012. Social networks and cooperation in hunter–gatherers. *Nature* 481, 497–501.
- Borenstein, E., Feldman, M.W., Aoki, K., 2008. Evolution of learning in fluctuating environments: when selection favors both social and exploratory individual learning. *Evolution* 62, 586–602.
- Boyd, R., Richerson, P.J., 1985. *Culture and the Evolutionary Process*. Univ. of Chicago Press, Chicago, IL.
- Enquist, M., Eriksson, K., Ghirlanda, S., 2007. Critical social learning: a solution to Roger's paradox of nonadaptive culture. *Am. Anthropol.* 109, 727–734.
- Feldman, M.W., Cavalli-Sforza, L.L., 1984. Cultural and biological evolutionary processes: gene–culture disequilibrium. *Proc. Natl. Acad. Sci. USA* 81, 1604–1607.
- Hamilton, W.D., 1964. The genetical evolution of social behaviour. I. *J. Theoret. Biol.* 7, 1–16.
- Henrich, J., 2012. Hunter–gatherer cooperation. *Nature* 481, 449–450.
- Hill, K.R., Walker, R.S., Bozicevic, M., Eder, J., Headland, T., Hewlett, B., Hurtado, A.M., Marlowe, F., Wiessner, P., Wood, B., 2011. Co-residence patterns in hunter–gatherer societies show unique human social structure. *Science* 331, 1286–1289.
- Kameda, T., Nakanishi, D., 2002. Cost–benefit analysis of social/cultural learning in a nonstationary uncertain environment: an evolutionary simulation and an experiment with human subjects. *Evol. Hum. Behav.* 23, 373–393.
- Lehmann, L., 2010. Space–time relatedness and Hamilton's rule for long-lasting behaviors in viscous populations. *Am. Nat.* 175, 136–143.
- Lehmann, L., Feldman, M.W., Kaeuffer, R., 2010. Cumulative cultural dynamics and the coevolution of cultural innovation and transmission: an ESS model for panmictic and structured populations. *J. Evol. Biol.* 23, 2356–2369.
- Lehmann, L., Rousset, F., 2010. How life history and demography promote or inhibit the evolution of helping behaviours. *Philos. Trans. R. Soc. B* 365, 2599–2617.
- Lehmann, L., Wakano, J.Y., 2013. The handaxe and the microscope: individual and social learning in a multidimensional model of adaptation. *Evol. Hum. Behav.* 34, 109–117.
- Lehmann, L., Wakano, J.Y., Aoki, K., 2013. On optimal learning schedules and the marginal value of cumulative cultural evolution. *Evolution* 67, 1435–1445.
- Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M.A., 2006. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 502–505.
- Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M.W., Fogarty, L., Ghirlanda, S., Lillicrap, T., Laland, K.N., 2010. Why copy others? Insights from the social learning strategies tournament. *Science* 328, 208–213.
- Rogers, A.R., 1988. Does biology constrain culture? *Am. Anthropol.* 90, 819–831.
- Taylor, P.D., Irwin, A.J., 2000. Overlapping generations can promote altruistic behavior. *Evolution* 54, 1135–1141.
- Wakano, J.Y., Aoki, K., 2006. A mixed strategy model for the emergence and intensification of social learning in a periodically changing natural environment. *Theor. Popul. Biol.* 70, 486–497.
- Wakano, J.Y., Aoki, K., Feldman, M.W., 2004. Evolution of social learning: a mathematical analysis. *Theor. Popul. Biol.* 66, 249–258.
- Wakano, J.Y., Nowak, M.A., Hauert, C., 2009. Spatial dynamics of ecological public goods. *Proc. Natl. Acad. Sci. USA* 106, 7910–7914.