doi:10.1111/evo.12040

ON OPTIMAL LEARNING SCHEDULES AND THE MARGINAL VALUE OF CUMULATIVE CULTURAL EVOLUTION

Laurent Lehmann,[1,2] Joe Yuichiro Wakano,[3] and Kenichi Aoki[4]

[1] *Department of Ecology and Evolution, University of Lausanne, Lausanne 1015, Switzerland*

[2] *E-mail: laurent.lehmann@unil.ch*

[3] *Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University, 1-1-1 Higashi Mita, Tama-ku, Kawasaki, Kanagawa 214-8571, Japan*

[4] *Department of Biological Sciences, University of Tokyo, Hongo 7-3-1, Bunkyoku, Tokyo 113-0033, Japan*

Received July 5, 2012
Accepted December 4, 2012

The age-dependent choice between expressing individual learning (IL) or social learning (SL) affects cumulative cultural evolution. A learning schedule in which SL precedes IL is supportive of cumulative culture because the amount of nongenetically encoded adaptive information acquired by previous generations can be absorbed by an individual and augmented. Devoting time and energy to learning, however, reduces the resources available for other life-history components. Learning schedules and life history thus coevolve. Here, we analyze a model where individuals may have up to three distinct life stages: “infants” using IL or oblique SL, “juveniles” implementing IL or horizontal SL, and adults obtaining material resources with learned information. We study the dynamic allocation of IL and SL within life stages and how this coevolves with the length of the learning stages. Although no learning may be evolutionary stable, we find conditions where cumulative cultural evolution can be selected for. In that case, the evolutionary stable learning schedule causes individuals to use oblique SL during infancy and a mixture between IL and horizontal SL when juvenile. We also find that the selected pattern of oblique SL increases the amount of information in the population, but horizontal SL does not do so.

KEY WORDS: Cumulative cultural evolution, optimal learning, life history, Rogers’ paradox.

Many organisms can change their behaviors and/or select novel actions among alternatives during their life span when exposed to novel environmental conditions (Dugatkin 2004). Individual learning (IL) is a generic term for the cognitive processes underlying such a change in behavior in the absence of interactions with conspecifics (Boyd and Richerson 1985; Rogers 1988; Dugatkin 2004). It comprises processes such as trial-and-error learning, statistical inference, induction and deduction, or insight. But individuals from many species can also acquire from others information about the appropriate phenotype(s) to express in a given environment. Social learning (SL) is a generic term for the psychological processes underlying the acquisition of information from others (Cavalli-Sforza and Feldman 1981; Boyd and Richerson 1985; Rogers 1988; Dugatkin 2004). It involves processes such as imitation, instructed learning, and local enhancement.

Individuals equipped with the ability to perform both IL and SL may obtain information from others and then build on it. This is likely to result in cumulative cultural evolution, where the phenotype(s) of one individual will depend on the implementation of IL and SL by other individuals in past time periods. Cumulative cultural evolution results in the acquisition of behaviors and other phenotypes that would not be possible to obtain by an individual performing learning in isolation. This underlies the ecological success of the human lineage (Klein 2009) and may occur in other animals (Tennie et al. 2009; Pradhan et al. 2012).

A necessary but not sufficient condition for the evolution of cumulative culture is that organisms use a composite learning strategy in which SL precedes IL (Boyd and Richerson 1985; Enquist et al. 2007; Aoki 2010). SL may precede IL when individuals can acquire information obliquely, from others living in the parental generation. But SL may also occur horizontally, between individuals of the same generation. In both cases, if SL precedes IL, the joint expression of these learning modes may result in cumulative information build up. But only vertical and/or oblique SL allows for a gradual increase in cultural complexity across generations.

If a reliance on vertical and/or oblique SL is necessary to absorb the extant culture, natural selection is expected to act differentially on the ontogeny of IL versus SL. This raises the question of what is the optimal resource allocation schedule toward IL and SL during an individual's development and its consequences for the pattern of cumulative cultural evolution within and across generations. Will selection favor SL at a young age to absorb information about the environment generated by individuals living in previous time periods? Primates seem to increase their reliance on IL as they grow (Reader and Laland 2001), and they take several years of SL alternated with practice (IL) to acquire tool-use techniques (Lonsdorf et al. 2004). Young honeybees foragers are more likely to rely on SL than mature individuals (Biesmeijer and Seeley 2005). This suggests that a reliance on SL at a young age may be a learning pattern common in natural populations.

But no species expresses as much SL at early ages as humans do (Konner 2010). The acquisition of nongenetically encoded information by learning to exploit hard to acquire food of high energetic value may lead to substantial fitness gains (Kaplan et al. 2000; Kaplan and Robson 2002). A learning intensive way to forage requires a long apprenticeship (Blurton Jones and Marlowe 2002), which may last up to 20 years (Gurven et al. 2006). If information acquired by the parental generation can be useful in this context, one expects selection to increase the reliance on learning during an individual's life span and to favor transfer of information across generations. This circumvents the need to innovate all solutions to environmental conditions de novo.

Learning is also costly in terms of time and energy (Kaplan et al. 2000; Mery and Kawecki 2004; Snell-Rood et al. 2011). Thus, IL and SL schedules are likely to impact on survival and fecundity trade-offs (Kaplan et al. 2000; Kaplan and Robson 2002; Kaplan et al. 2009). In other words, cumulative cultural evolution is expected to coevolve with the life history of a species.

Although the coevolution of IL and SL has been intensively studied (e.g., Rogers 1988; Stephens 1991; Boyd and Richerson 1995; Feldman et al. 1996; Wakano et al. 2004; Enquist et al. 2007), and cumulative cultural evolution increasingly investigated (e.g., Enquist et al. 2008; Lehmann and Feldman 2009; Strimling et al. 2009; Mesoudi 2011; Pradhan et al. 2012), there are but a few quantitative studies addressing the coevolution of IL and SL in the context of cumulative culture. Some investigations consider the evolution of social-learner-explorer strategies (where SL precedes IL), which can produce a change in nongenetically encoded phenotype(s) across generations (Boyd and Richerson 1995; Borenstein et al. 2008; Aoki 2010). But these studies do not consider explicitly the dynamics of culture within generations, which is taken into account in studies that either neglect intergenerational transfers of information (van Schaik and Pradhan 2003; Lehmann et al. 2010) or offer only the possibility of oblique SL (Aoki et al. 2012).

The trade-offs between no learning and learning, IL and SL, and oblique and horizontal SL, have thus not been simultaneously taken into account, and their consequences for life-history evolution have not been studied. But for understanding the conditions under which cumulative cultural evolution is likely to be an evolutionary stable strategy (ESS, Maynard Smith 1982), SL should coevolve with IL in a situation where alternative learning schedules are possible.

In this article, we use evolutionary game theory that we couple with optimal control theory to identify optimal learning schedules in a model where organisms have three distinct life stages. These are "infants," "juveniles," and "adults." IL and/or oblique SL may occur during infancy so that information generated in past generations can be acquired, while IL or horizontal SL can take place among juveniles. Resources are gathered during adulthood according to the amount of nongenetically encoded information acquired during the learning stages. Because the length of these stages as well as the behavioral options can evolve, our model allows for no learning as well as a schedule of learning supportive of cumulative cultural evolution both within and between generations to be a possible evolutionary stable strategy.

Model

LIFE-HISTORY ASSUMPTIONS

We consider a haploid population of constant but very large size. We assume that the individuals in this population are endowed with physiological mechanisms allowing them to produce nongenetically encoded information by IL, and to acquire such information from others by SL. For ease of presentation of the basic components of our model and to gain an intuitive understanding of the intergenerational explicit cultural dynamics, we present the life history of the individuals by assuming a population where all individuals express the same phenotypes (no genetic variation). This assumption is relaxed when the evolutionary analysis is carried out in Supporting Information Appendices A and B, so that our results do not rely on it.

We measure nongenetically encoded information on a continuous scale and denote by $A(t)$ the amount of adaptive information

Table 1. List of symbols.

Symbol	Definition
K	Length of the infant period.
L	Length of the juvenile period.
G	Length of a generation.
$A(t)$	Information carried by an age $0 \leq t \leq G$ individual.
$P(t)$	Total information in an age $0 \leq t \leq G$ population.
φ	Conversion factor of amount of resources into offspring number.
α	Oblique SL learning rate of infants.
β	Horizontal SL learning rate of juveniles.
ϵ	Decay rate of information.
ρ	Nonoverlapping fraction of information produced by different ILs.
γ	Fraction of information of the parental generation acquired by the infants of the descendant generation at equilibrium of the learning dynamics.
ϕ	Dependence of resource intake on cultural information.

held by a representative individual of age t in the population (see Table 1 for a list of symbols). This variable can also be interpreted as the number of units of adaptive knowledge held by the individual, and may be obtained as the size (minimal) of the set describing this information (Reiter 2001). The units of adaptive knowledge may involve information about the abundance and locations of preys and predators, items such as techniques to build arrows or tailored clothes, lists of poisonous foods, methods to obtain sources of heat. We assume that the environment is constantly and smoothly changing so that each unit of information becomes obsolete at rate ϵ per unit time. Alternatively, ϵ may also be interpreted as the rate of forgetting of individuals.

The life cycle of the organism is as follows. (1) During a period of infancy of length K, every individual in the population may learn information about the environment from the individuals in the parental generation through oblique social transmission. At the end of this period, each individual has acquired $A(K)$ units of information and individuals of the parental generation die. (2) During a juvenile period of length L, each individual can produce novel information by IL or acquire information from others by horizontal SL. At each time $t \in [K, K + L]$, an individual devotes fractional effort $u(t)$ to IL, while allocating the complementary fraction of effort $1 - u(t)$ to acquiring information from others by SL. (3) Juveniles reach adulthood. During a period of time of length $G - K - L$, where G is the generation time, each adult gathers resources according to the amount of information obtained by the end of the juvenile learning period (see Fig. 1 for an overview of learning in the life cycle). (4) Each individual produces a large number of offspring. Density-dependent competition occurs and the population is regulated back to its census size. The cycle starts again, but adults survive after giving birth at age G and serve as exemplars for the infants of the next generation (post-reproductive period of length K).

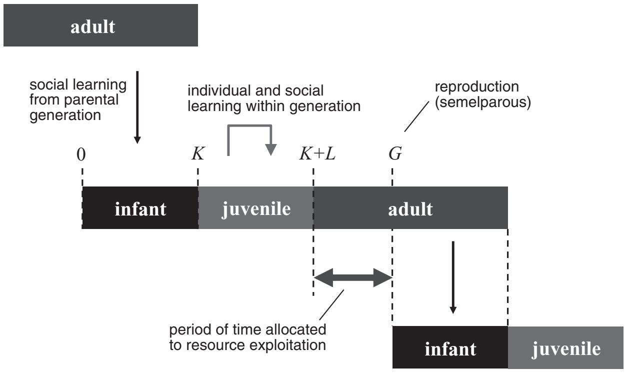

Figure 1. Cultural transmission and partition of an individual's life span into three stages.

Our aim in this article is to determine the coevolutionary stable candidate learning schedule $u^*(t)$, K^*, and L^*. This requires that we specify the dynamics of information $A(t)$ held by an individual of age $t \in [0, G]$. This will be carried out by assuming that the information dynamics has already reached an equilibrium so that $A(t)$ denotes the amount of information held by an individual of age t in the parental generation as well as that in the offspring generation.

INFORMATION DYNAMICS

Adults

Right at the start of adulthood, an individual of age $t = K + L$ has $A(K + L)$ units of information. Because it decays at rate ϵ, an individual of age $t = K + L + \tau$ carries $A(K + L)e^{-\epsilon\tau}$ units of information ($\tau \in [0, G - K - L]$). We assume that the individual gathers energy at rate $1 + \phi A(K + L)e^{-\epsilon\tau}$ at age $t = K + L + \tau$. Here, "1" is the baseline unit of energy obtained without expressing culturally acquired information, while the parameter ϕ denotes the dependence of foraging efficiency on acquired cultural information. The energy is stored and consumed at age G to produce offspring with a conversion factor φ. This gives total energy amount $\int_0^{G-K-L}[1 + \phi A(K + L)e^{-\epsilon\tau}]\,\mathrm{d}\tau$ and offspring number produced (fecundity) as

$$f = \varphi\left[(G - K - L) + \left(\frac{1 - e^{-\epsilon(G-K-L)}}{\epsilon}\right)\phi A(K + L)\right]. \quad (1)$$

Infants

A surviving offspring is assumed to acquire information only by SL and from a single exemplar individual randomly sampled from the parental generation so that it can be thought of as acquiring information from an average individual in the population. The information held by such an infant of age $t = 0$ is $A(0) = 0$, whereas that of its exemplar individual is $A(G)$. When the infant is of age $t = \tau$, the exemplar individual has $A(G)e^{-\epsilon\tau}$ units of remaining information, because initial information decays at rate ϵ. We assume that the infant acquires information according to the difference $A(G)e^{-\epsilon t} - A(t)$ between the information it has, $A(t)$, and that of the exemplar individual, $A(G)e^{-\epsilon t}$, at rate α per unit time. The parameter α may capture various cultural acquisition processes such as imitation, instructed learning, or collaborative learning (Tomasello et al. 1993). This leads for infancy ($t \in [0, K]$) to the information dynamics

$$\dot{A}(t) = -\epsilon A(t) + \alpha[A(G)e^{-\epsilon t} - A(t)] \tag{2}$$

whose solution is

$$A(K) = [1 - e^{-\alpha K}]A(G)e^{-\epsilon K}. \tag{3}$$

Juveniles

When an individual produces adaptive information by IL, each unit of effort $u(t)$ allocated to producing leads to μ novel units of information. Different individuals may produce different pieces of information and the total amount of information in the population at time t is denoted $P(t)$. We have $P(t) - A(t) \geq 0$, which is the amount of information in the population a representative individual has not yet acquired. When an individual performs SL, we assume that each unit of effort $1 - u(t)$ allocated to that task leads to the acquisition of $\beta(P(t) - A(t))$ novel units of information, where β is the rate of social transmission of information. To track the dynamics of $A(t)$, we thus need to specify the dynamics of $P(t)$. The information dynamics is assumed to be given for $t \in [K, K + L]$ by the system

$$\begin{aligned} \dot{A}(t) &= -\epsilon A(t) + u(t)\mu + (1 - u(t))\,\beta\,(P(t) - A(t)) \\ \dot{P}(t) &= -\epsilon P(t) + u(t)\mu\,(1 + \rho)\,. \end{aligned} \tag{4}$$

where ρ describes the overlap among individuals of production of information in the population. When $\rho = 0$, every individual produces exactly the same type of information. However, when ρ is approximately as large as population size, every individual would produce different pieces of information at rate μ.

The initial condition $A(K)$ for the juvenile learning dynamics (eq. 4) depends on the amount of information $A(G)$ that can be obtained from the past generation (eq. 3), which itself depends on the amount of information at the end of the juvenile period: $A(G) = A(K + L)e^{-\epsilon(G-K-L)}$. This gives $A(K) = \gamma A(K + L)$, where

$$\gamma = [1 - e^{-\alpha K}]e^{-\epsilon(G-L)} \tag{5}$$

and allows us to close the cycle of information dynamics across generations at the individual level.

To obtain the initial condition $P(K)$ for the juvenile information dynamics (eq. 4), we assume that there is no loss of information in the population due to stochastic sampling when infants copy their exemplar individual. In other words, we assume that the ratio of the amount of information carried by an individual to that carried by the whole population, A/P, is kept constant during the infancy and adult periods and can only change during the juvenile period according to equation 4. This gives $P(K) = \gamma P(K + L)$, which closes the cycle of information dynamics across generations at the population level. These dynamics of information describe cumulative cultural evolution in a quantitative sense, where innovations accumulate by spreading throughout the population by oblique and horizontal transmission.

Results

Having specified the information dynamics within and across generations, we now have all the elements to evaluate the fecundity of an individual expressing mutant trait values in $u(t)$, K, or L, and thereby determine the selection pressure on these evolving phenotypes. The characterization of the optimal learning strategies is carried out in the Supporting Information by using evolutionary game theory (e.g., Maynard Smith 1982; Eshel 1983; Lessard 1990; Dercole and Rinaldi 2008) coupled with optimal control theory (e.g., Iwasa and Roughgarden 1984; Perrin 1992; Bulmer 1994; Day and Taylor 1998; Sydsaeter et al. 2008).

OPTIMAL JUVENILE LEARNING SCHEDULE

We start by evaluating the candidate ES level $u^*(K + \tau)$ of effort allocated to IL at time $\tau \in [0, L]$ of the juvenile period. An individual increasing its level of IL at time τ by one unit causes a change in its reproductive output that is proportional to

$$\mu - \beta\,[P(\tau) - A(\tau)]\,. \tag{6}$$

This provides the selective pressure on IL (eq. A-10 in Supporting Information), where μ is the information obtained by increasing IL at time τ, which thus gives the marginal benefit of IL. The marginal cost of IL is given by the second term, $\beta[P(\tau) - A(\tau)]$, which is the amount of information obtained if SL was increased at time τ by one unit. When the benefit exceeds the cost, IL is favored and $u(K + \tau)$ should go to one. Conversely, when the selection pressure is negative, the cost of IL exceeds the benefit and more information would be gained by expressing SL. Now, $u(K + \tau)$ should go to zero. It may also be the case that the benefit exactly balances the cost, in which case a mixed strategy can be selected for, the value of which depends on how the control schedule affects trait dynamics at the individuals and population level ($P(\tau)$ and $A(\tau)$). Whether a pure strategy ($u(t) = 0$ or $u(t) = 1$) or a mixed strategy ($0 < u(t) < 1$) is favored, equation 6 shows that the selection pressure on IL depends on a direct trade-off between acquiring information by IL and SL, with IL being favored when β tends to be small because only scant information can then be obtained by SL.

Pure strategy

We find that when the horizontal SL learning rate β is below the threshold value

$$\beta < \frac{\epsilon(1 - \gamma e^{-\epsilon L})}{\rho(1 - e^{-\epsilon L})}, \tag{7}$$

which tends to decrease with ρ, γ, and L, and to increase with ϵ, IL should be applied throughout the whole juvenile period: $u^*(K + \tau) = 1$ for all $\tau \in [0, L]$ (eqs. A-10–A-18 in Supporting Information for a proof). This inequality will be satisfied when the juvenile learning length is small and ϵ high so that no useful information has the time to accumulate at the population level to be successfully obtained by SL. For this pure IL schedule, the ES amount of information carried by an individual at the beginning of adulthood is

$$A^*(K + L) = \frac{\mu(e^{\epsilon L} - 1)e^{\alpha K + \epsilon(G-L)}}{\epsilon[1 + (e^{\epsilon G} - 1)e^{\alpha K}]}, \tag{8}$$

which is a decreasing function of G if $\epsilon > 0$ and of ϵ, but is an increasing function of μ, α, K, and L (eqs. A-23–A-24 in Supporting Information).

Mixed strategy

When inequality (7) is reversed, which is likely to occur when the rate β of horizontal SL is high, the optimal control for a juvenile of age $t = K + \tau$ is given by

$$u^*(K + \tau) = \begin{cases} 1 & 0 \le \tau \le t_s \\ \dfrac{\epsilon + \beta}{(1 + \rho)\beta} & t_s < \tau \le L. \end{cases} \tag{9}$$

Here, an individual first allocates all its effort into IL until a switching time t_s is reached. After that, the individual expresses a mixed strategy, where only a fraction of each unit of time is spent on IL. This occurs if enough information has been generated by IL so that it pays to perform SL (eq. 6). The switching time is given by

$$t_s = \frac{1}{\epsilon} \log\left[\frac{\beta\rho - \gamma\epsilon}{\beta\rho - \epsilon}\right], \tag{10}$$

which is a decreasing function of ρ (eqs. A-25–A-33 in Supporting Information for a proof). Thus, when different individuals generate different pieces of information ($\rho > 0$), the switching time to SL is reduced because an individual performing SL can usefully obtain more adaptive information.

For this mixed control schedule, the ES amount of information carried by an individual of age $t = K + \tau$ is the same as under the pure IL schedule, in particular $A^*(K + L)$ is still given by equation 8 (eq. A-34 in Supporting Information). Hence, even if SL is selected for, the amount of information held by an individual at any point in time at an evolutionary equilibrium is not increased relative to the case where pure IL occurs. This is an example of the so-called Rogers' paradox for the evolution of SL (Boyd and Richerson 1995; Enquist et al. 2007; Rendell et al. 2010).

Rogers' paradox occurs when the parameter values of a model are such that (1) a stable equilibrium exists with SL and IL, (2) a monomorphic equilibrium of pure IL also exists, and (3) the two equilibria result in the same mean fecundity and/or survival in the population. In this case, cultural transmission does not increase the vital rates of the individuals in the population. Hence the "paradox," which was first observed in a model where individuals had to learn an optimal solution in a one shot learning period (Rogers 1988). In our model Rogers' paradox occurs because the total time devoted to learning during the juvenile period is fixed, and SL is selected to increase (which concomitantly decreases IL in the population) up to the point where an individual gains the same amount of information is if it would implement pure IL.

ES LEARNING LENGTHS

We now present the results on the optimal length of the infancy and juvenile learning periods (K^* and L^*) by holding the learning schedule $u^*(t)$ at the candidate ES derived in the previous section. Importantly, the results presented below hold for both the pure and mixed control schedules discussed above, and the ES values K^* and L^* depend on $u^*(t)$, which is itself solved as a function of K^* and L^* (see Supporting Information Appendix B). Hence, K, L, and $u(t)$ coevolve because the trait values adjust to each other through the action of natural selection. The results presented below thus characterize the ES infancy and juvenile learning periods at the candidate coevolutionary stable learning pattern ($u^*(t)$, K^*, L^*).

Juvenile learning

The infancy learning period (K) cannot evolve if there is initially no juvenile stage ($L = 0$), because in this case there is no information to be acquired from the parental generation. Hence, we first assume that there is no infancy period ($K = 0$) and consider the evolution of the juvenile period. To highlight the main marginal costs and benefits of varying L, we assume that $\epsilon \to 0$. In this case, an individual increasing its juvenile learning length by one unit in a population fixed for L causes a change in its reproductive output that is proportional to

$$\mu\phi(G-L)-\left(1+\phi\underbrace{(\mu L)}_{A(K+L)}\right) \tag{11}$$

(eq. B-8 in Supporting Information). The first term of this selection gradient on juvenile learning is the marginal benefit of increasing the juvenile learning period. This leads to an increase of information μ that can be used throughout the resources gathering period (length $G-L$) according to the extent ϕ to which it increases fecundity. The second term in the selection gradient is the marginal cost, which stems from the reduction of the resource gathering period when more time is devoted to learning, and where resources are harvested at rate $1+\phi A(L)$ with $A(L)=\mu L$ being the amount of information obtained during juvenile learning.

When $K=0$, the amount of information gained during infancy is zero and equation B-8 in Supporting Information shows that L will be selected for provided life span is not too short ($G \gg 0$) in which case it is likely that $\phi G\mu > 1$. More generally, when there are environmental fluctuations, selection favors learning at the juveniles stage when the innovation rate multiplied by the reliance on nongenetically encoded information exceeds the decay rate ($\mu\phi > \epsilon$) and life span is not too short ($G \gg 0$). If this is the case, the juvenile stage evolves to a period of positive duration

$$L^* = \frac{G}{2} - \frac{2+G\epsilon}{4\mu\phi} + O(\epsilon^2), \tag{12}$$

where $O(\epsilon^2)$ is a term of order ϵ^2, which can be neglected when there is a low rate of environmental change, that is, adaptive information decays slowly (eqs. B11–B-13 in Supporting Information). From equation 8, we then find that the associated ES amount of information held by an individual at the end of the juvenile period is

$$A^*(L^*) = \frac{G\mu}{2} - \frac{1}{2\phi} - \frac{\epsilon(1+G^2\mu^2\phi^2)}{8\mu\phi^2} + O(\epsilon^2). \tag{13}$$

This is an decreasing function of ϵ and an increasing function of the other parameters when positive L is selected for.

In the absence of environmental fluctuations ($\epsilon \to 0$), one obtains $L^* = G/2 - 1/(2\mu\phi)$ and $A^*(L^*) = G\mu/2 - 1/(2\phi)$. Selection thus supports the evolution of learning even in the absence of environmental fluctuations ($\epsilon = 0$). This may seem to contradict the standard result that learning should be selected against in the absence of environmental fluctuations (e.g., Stephens 1991; Wakano et al. 2004). But this is likely to be the case only in an "absolutely fixed environment" (Stephens 1991), where every feature of the environment, whether biotic (e.g., the abundance and spatial position of every single prey and predator) or abiotic (e.g., the velocity of the wind in every food patch) remains forever constant and the same. In this case, the cost of learning may outweigh its benefit if the appropriate phenotype(s) to express can be genetically encoded and the learning machinery is expensive. This not only implies $\epsilon = 0$ in our model, but that investment into learning is not beneficial, which can be captured by setting $\phi = 0$. Now, learning is selected against and our model is indeed consistent with previous formalizations.

Infancy learning

When selection favors reliance on learning in juveniles, there is information to be acquired by oblique SL during infancy because mature individuals carry adaptive information. To analyze the selection pressure on the infancy period (K), we again start by assuming negligible environmental fluctuations. Then, an individual increasing its infancy learning length by one unit in a population fixed for K causes a change in its reproductive output that is proportional to

$$\alpha A(G)e^{-\alpha K}\phi(G-K-L) - [1+\phi A(K+L)] \tag{14}$$

(eq. B-15 in Supporting Information). The first term is the marginal benefit of increasing the infancy learning period because this leads to an increase of $\alpha A(G)e^{-\alpha K}$ units of information used throughout the resources gathering period (length $G-K-L$). The second term is the concomitant cost stemming from the reduction of the resource gathering period.

When the infancy learning rate (α) is large enough, the information taken up at the end of the infancy learning period may balance the cost of reducing the total time of gathering resources. This occurs when

$$\alpha > \frac{2\mu\phi}{G\mu\phi - 1}, \tag{15}$$

which decreases with G, μ, and ϕ, and entails that the infancy learning period is selected for and also implies positive selection on L. Then, the candidates ES learning periods are obtained as

$$\begin{aligned} K^* &= G - \frac{1}{\mu\phi} - \frac{1}{\alpha} - \frac{W\left(e^{\alpha G - 1 - \frac{\alpha}{\mu\phi}}\right)}{\alpha} \\ L^* &= \frac{1}{\alpha}, \end{aligned} \tag{16}$$

where $W(\cdot)$ is the principal solution of the Lambert function (eqs. B-18–B-21 in Supporting Information). Here, L^* depends only on the infancy learning rate and varies inversely with it, whereas K^* depends on all the model's parameters, except β. The ES amount of information associated to these learning periods is

$$A^*(K^*+L^*) = \mu L^* e^{\alpha K^*}, \tag{17}$$

which depends on two components. First, the number of units of information μL^* generated by IL during a single juvenile period of length L^*. Second, the amount of such information that can be amalgamated across generations and is measured by the factor

$e^{\alpha K^*}$. This is equal to 1 when $\alpha = 0$. Hence, αK^* can be thought of as the growth rate of information, which is repeatedly generated by IL and then transmitted by SL across generations.

Because the infancy learning period goes to zero ($K^* \to 0$) as oblique SL becomes perfect ($\alpha \to \infty$), that is, an individual instantly absorbs all the extant culture of its exemplar individual, the total amount of information $A^*(K^* + L^*)$ will remain bounded even if the efficiency of oblique SL becomes maximal. In this case, we obtain from equation 17 that when learning becomes perfect ($\alpha \to \infty$), the amount of information held by an individual is

$$A^*(K^* + L^*) = G\mu - \frac{1}{\phi}. \tag{18}$$

This is an upper bound for the amount of information that can be acquired by cumulative cultural evolution under our model's assumption. Comparing equation 18 with equation 13 shows that the gain in information due to oblique SL (during the infancy period of length K^*) only results in at most a doubling of that acquired by IL.

The general dependence of the learning periods and amount of information on parameter values is illustrated in Figure 2, which shows that $A^*(K^* + L^*)$ is an increasing function of μ, α, and G. Hence, the selected pattern of oblique SL increases the amount of information held by individuals at steady state relative to the case where only the juvenile period is selected for, and which entails that only horizontal SL occurs. This was not the case for the selected pattern of horizontal SL. Hence, SL now leads to cumulative cultural evolution across generations and Rogers' paradox is not observed under oblique SL.

When environmental change is negligible ($\epsilon \to 0$), information is lost only through imperfect transmission across generations ($\alpha < \infty$). Roughly speaking, the dynamics of cultural evolution as α varies can then be classified into three regimes (Fig. 2). (1) When oblique transmission is not efficient, the infant learning period is not selected for ($K^* = 0$). But the juvenile learning period can be selected for, which results in individuals acquiring some amount of knowledge by SL. (2) When oblique transmission becomes more efficient, K is selected for at the expense of the optimal juvenile period L^*, which is reduced in length. The accumulation of knowledge is now enhanced due to oblique SL. (3) As the efficiency of oblique SL becomes even higher, not only is L^* reduced in length but also K^*. This occurs because a shorter learning period allows the accumulation of an otherwise equal amount of information that is already close to its maximum, and which can be used in an extended adulthood period to gather resources.

Overall, increasing α causes a decrease of the total learning period $K^* + L^*$ and an increase of the length of adulthood (Fig. 2). A larger innovation rate μ tends to increase both learning periods. In regime (1), a high innovation rate contributes relatively little to cultural evolution, while it has a larger effect when cumulative cultural evolution has evolved (regimes 2, 3). Generation time, or life expectancy, has the same tendency as innovation rate (Fig. 2).

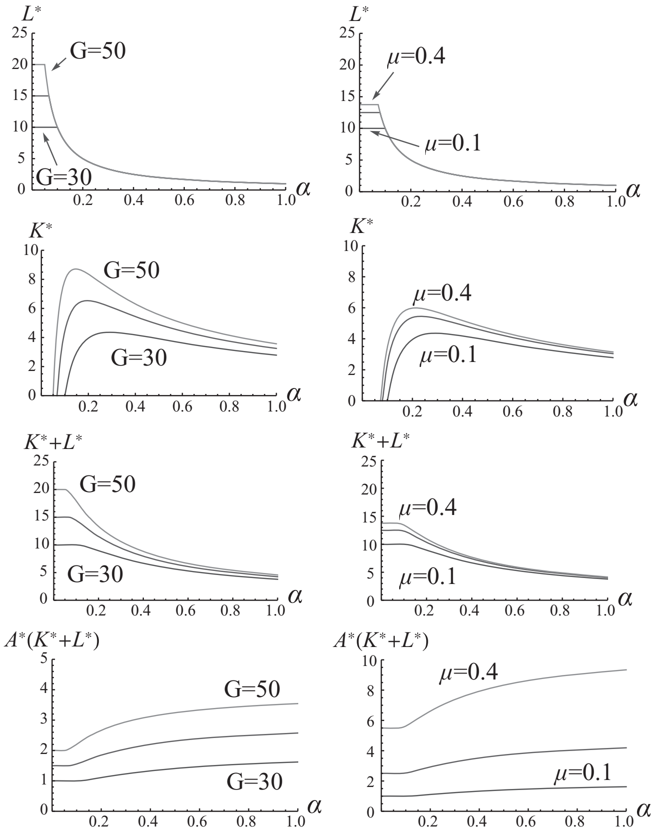

Figure 2. Candidate ES learning length predicted by equation 16 as a function of α for $\phi = 1$. In the first column of panels, one has $\mu = 0.1$, and from the top to the bottom line $G = 30, 40, 50$. In the second column of panels, one has $G = 30$, and from the bottom to the top line $\mu = 0.1, 0.2, 0.4$. In all panels, the horizontal strait lines give the values for L^* and $A^*(K^* + L^*)$ when α does not satisfy inequality 15. Hence, the infancy learning period is not selected for and only horizontal SL may occur during the juvenile period. This can also be seen in the panels for K^*, which takes the value of zero in such cases.

More generally, there will be environmental fluctuations ($\epsilon > 0$). For this case, we were unable to derive analytically the evolutionary stable learning periods (K^* and L^*) and thus relied on numerical work. In Figure 3, we show how ϵ affects K^*, L^*, and $A^*(K^* + L^*)$. The dynamics of cultural evolution as ϵ varies can approximately be classified into two regimes (Fig. 3). (1) When ϵ first increases, it causes a reduction in K^* and an increase in L^* because less information can be learned from the parental generation and more effort needs to be devoted to IL to sustain the total amount of information. This results in an overall increase in the total period devoted to learning (Fig. 3). (2) Above a threshold value of ϵ, the overall time devoted to learning is selected against

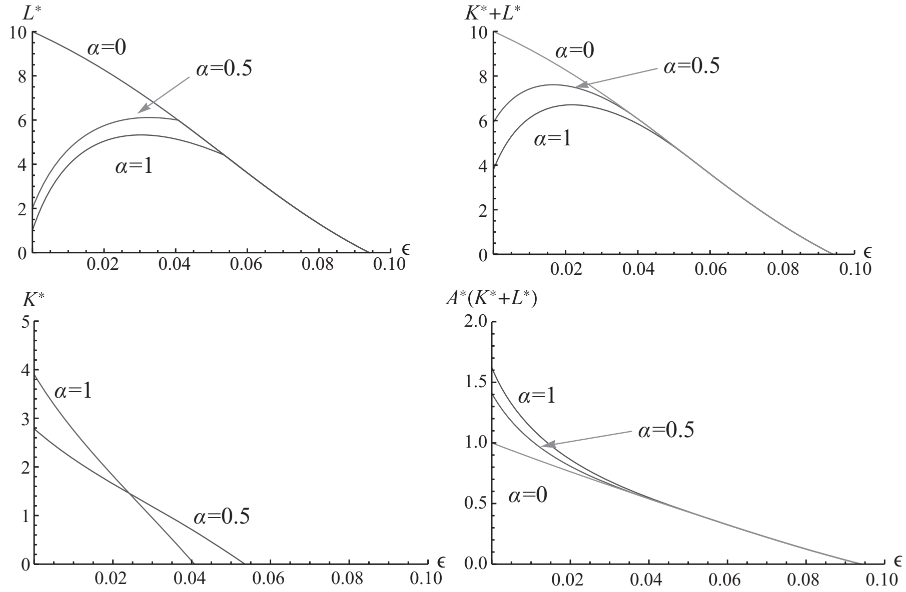

Figure 3. Candidate ES learning length obtained by solving equation B-7 in Supporting Information Appendix B and graphed as a function of ϵ. In all panels, $G = 30$, $\phi = 1$, $\mu = 0.1$, while we have $\alpha = 1, 0.5, 0$. The monotonic decreasing line in the panels for L^* and $K^* + L^*$ is for $\alpha = 0$, in which case K^* is zero.

and both K^* and L^* decrease, eventually reaching a value of zero in which case learning is selected against altogether. This occurs because the rate of decay of information is too large for individuals to acquire enough information to counterbalance the cost of learning, and accords with the view that when the environments becomes too unpredictable, learning should be selected against as no regularity can be exploited by it (Stephens 1991).

Our analysis shows that a learning schedule supportive of cumulative cultural evolution can be selected for. But it may be felt that we should have taken into account the possibility for IL during infancy for the analysis to be really convincing. Taking this into account will not change qualitatively our results because allowing for IL during infancy results in a situation where SL is selected against only if an infant at birth can learn more information on its own than it can acquire from the parental generation (see Supporting Information Appendix C, eqs. C-1–C-4, for a formal argument). It is natural, however, to assume that the rate of IL during infancy is lower than that during the juvenile period (Striano et al. 2001). In this case, if IL is selected for during infancy, the length of the infancy period is likely to be selected against as more traits are learned during the juvenile period. Hence, assuming only SL during infancy is unlikely to affect our qualitative results that oblique SL comes first if learning from the parental generation can evolve.

Discussion

We have investigated the coevolution of an infant and a juvenile learning period, where individuals can allocate effort to both IL and/or SL. Our model, where the environment is constantly but smoothly changing within and between generations allows for the possibility that no learning is selected for or that only IL is evolutionary stable. It also allows for a situation where oblique SL evolves during infancy or that a mixture of IL and horizontal SL occurs in juveniles. A possible outcome of our model is thus a learning schedule that is favored by natural selection to support cumulative cultural evolution both within and across generations, an outcome that has not been taken into account in previous formalizations.

We find that when the innovation rate weighted by the importance of using nongenetically encoded information for gathering resources is larger than the rate of environmental change, juvenile IL can be selected for provided the generation length (G) is not too small. At equilibrium, an individual will thus allocate energy to IL and express a positive ES amount of adaptive information gathered during its life span (eq. 13).

Once juvenile IL has been selected for, individuals carry information that can be learned by others. Hence, it may become beneficial to perform SL horizontally or obliquely. When the infancy learning rate (α) is large enough to overcome the opportunity cost of spending time on learning, individuals are selected to acquire information from the parental generation (eq. 15). Both the optimal infant and juvenile learning periods then coevolve and the coevolutionary optimal learning schedule should unfold according to the following sequence. (1) Pure oblique SL during infancy from age zero to K^*. (2) Pure IL during the juvenile period from age K^* to L^*. This is a learning schedule supportive of cumulative cultural evolution across generations.

Alternatively, when the horizontal transmission rate (β) is large individuals may gain from horizontal SL, in which case the optimal learning schedule unfolds according to the following sequence. (1) Pure oblique SL during infancy until age K^*. (2) Pure IL from age K^* to age $K^* + t_s$, where t_s is a switching time to a mixed strategy. (3) Mixed allocation to individual and SL from age $K^* + t_s$ to L^*. This is again a learning schedule supportive of cumulative culture across generations. Here, however, SL occurs twice, once obliquely from individuals of the parental to the offspring generation and once horizontally.

Our analysis suggests that the ES amount of information acquired during the juvenile period by horizontal SL does not increase the amount of information held by an individual, relative to the case where only IL occurs. Thus, the amount of information held by an individual does not depend on the efficiency β of horizontal transmission (eq. 8, Rogers' paradox obtains). By contrast, when oblique SL evolves through selection of an extended infant learning period, the outcome of evolution is a net increase in the amount of information (Fig. 2). Here, the transfer of information across generations causes individuals to accumulate more information than if they were to perform only IL. The ES amount of information thus depends on the efficiency α of oblique transmission (Fig. 2, Rogers' paradox does not obtain).

That horizontal SL during the juvenile period does not increase the ES amount of information can be understood by noting that the selection pressure on horizontal SL involves a direct trade-off between SL and IL (eq. 6). SL is selected to increase as long as an individual can gain more information by SL than by IL. As the total time spent on IL decreases, the total amount of information generated by IL in the population decreases and thereby the selection pressure on SL. Eventually, a point is reached where selection for increased SL vanishes. This corresponds to a situation where an individual would not gain more information by expressing SL than pure IL and Rogers' paradox obtains.

By contrast, the selection pressure on oblique SL does not directly involve a trade-off between SL and IL, but a trade-off between acquiring information by learning and gathering resources (eq. 14). Increasing the time spent on SL may thus only partially impart on a reduction of time devoted to IL, because it can also reduce the time spent gathering resources. Here, oblique SL is selected for as long as oblique SL results in an increase of the amount of information obtained by an individual, relative to the case where it can only acquire information during the juvenile period. Because the increase in the length of oblique SL does not result in an equivalent reduction of IL length (as the total life span is fixed), Rogers' paradox is no longer expected to obtain.

Suppose, on the other hand, that adulthood length is fixed and that the life span can vary freely as K and L evolve. Although this is not biologically realistic, one expects in this case that oblique SL would be increased to the point where the total amount of information held by an individual is the same as if it performed only IL. This is so because the length of adulthood is fixed, the selection pressure on oblique SL now involves a direct trade-off between SL and IL. This is a situation qualitatively similar to that for horizontal SL and Rogers' paradox should now obtain. This intuition can be confirmed by calculation because if the length of life span (G) is allowed to vary freely, the selected oblique and horizontal SL schedule does not increase the amount of information in the population, relative to pure IL (see Supporting Information Appendix D, eqs. D-3–D-5).

If SL does not involve a direct trade-off with IL, cultural transmission can result in an increase of the amount of information in the population. But SL can only lead to a doubling of this amount relative to the case of pure IL (compare eq. 13 to eq. 18). It can also be shown that SL does not optimize the amount of information from the perspective of the population. Moreover, allowing infants to learn from the total pool of information in the population (by using $P(G)$ instead of $A(G)$ in eq. 2) does not lead to more information accumulation across generations.

The moderate effect of SL on information accumulation relative to pure IL may stem from the fact that SL does not in itself increase the likelihood to find optimal solutions. It mainly decreases the cost of IL because more information can be acquired by SL than by IL (eqs. 2 and 4), which stems from the fact that information is a public good that is nonrival and nonexcludable, and is accumulated across time by oblique and horizontal SL. This results in cumulative culture in a quantitative sense, which is a situation that is more prone to Rogers' paradox than if SL is necessary to find optimal phenotypic solutions (Boyd and Richerson 1995; Enquist and Ghirlanda 2007). In such a case, which could be described as a situation where culture is cumulative in a qualitative sense, we found that the ES learning schedule optimizes the phenotype of individuals from the perspective of the population in a presence of only two discrete learning stages (Aoki et al. 2012). This suggests that the type of culturally transmitted information can have an effect on the extent to which cultural transmission increases the vital rates of individuals in a population. To have a better understanding of the conditions when this will be the case, it seems necessary to consider more explicitly the learning mechanisms behind cultural evolution.

Despite its simplifying assumptions, in particular the neglect of vertical transmission that would require a kin selection analysis to be taken into account, our model predicts some broad guided patterns of cultural acquisition. It shows that the evolution of the age-dependent expression of IL and SL results in a schedule where SL precedes IL. This seems indeed to be the case in primates (Reader and Laland 2001; Lonsdorf et al. 2004; van Schaik 2004) and may occur in honey bees (Biesmeijer and Seeley 2005). Further empirical work is necessary to delineate whether this is a general trend and also the extent to which various species use information acquired by individuals living in past generations.

Our model also predicts a mixed expression of IL and SL during the juvenile's period, which makes it overall very consistent with the pattern of learning observed in humans during childhood. This consists of the acquisition of information from the parental generation during infancy and early childhood (to until about age 6), and then a period of horizontal transmission (SL) and exploration (IL) during middle childhood until adolescence (Konner 2010, chapters 24–26, Hewlett et al. 2011). Our "infancy" stage can be interpreted as corresponding to the periods of real infancy and early childhood, whereas our "juvenile" stage correspond to a period of learning until individuals have enough information to produce resources on their own. This may actually involve the stages of adolescence and early adulthood in natural population as it may take up to 10 or 20 years of IL and SL for developing successful hunting skills (Gurven et al. 2006).

Our model is also a general step in the direction of taking into account the age variability of learning patterns, which affects the information flow within a population and has so far been neglected in evolutionary analysis (Demps et al. 2012). But future work is needed to gain a better understanding of the coevolution between learning and other life-history features as the human life cycle may have been critically shaped by this dynamic interaction (Kaplan et al. 2000; Kaplan and Robson 2002).

ACKNOWLEDGMENTS

We thank R. Barkai for providing useful references. We also thank a reviewer and J. Wilkins for comments that improved the article. This work is partly supported by Monbukagakusho grant 22101001 and Swiss National Science Foundation grant PP00P3-123344.

LITERATURE CITED

Aoki, K. 2010. Evolution of the social-learner-explorer strategy in an environmentally heterogeneous two-island model. Evolution 64:573–581.

Aoki, K., J. Y. Wakano, and L. Lehmann. 2012. Evolutionarily stable learning schedules in discrete generation models. Theor. Popul. Biol. 81:300–309.

Biesmeijer, J. C., and T. D. Seeley. 2005. The use of waggle dance information by honey bees throughout their foraging careers. Behav. Ecol. Sociobiol. 59:133–142.

Blurton Jones, N.G., and F. W. Marlowe. 2002. Selection for delayed maturity: does it take 20 years to learn to hunt and gather? Hum. Nat. 13:199–238.

Borenstein, E., K. Aoki, and M. W. Feldman. 2008. Evolution of learning in fluctuating environments: when selection favors both social and exploratory individuals. Evolution 62:586–560.

Boyd, R., and P. J. Richerson. 1985. Culture and the evolutionary process. Univ. of Chicago Press, Chicago.

———. 1995. Why does culture increase human adaptability? Ethol. Sociobiol. 16:125–143.

Bulmer, M. 1994. Theoretical evolutionary ecology. Sinauer Associates, Massachusetts.

Cavalli-Sforza, L., and M. W. Feldman. 1981. Cultural transmission and evolution. Princeton Univ. Press, NJ.

Day, T., and P. Taylor. 1998. The evolution of temporal patterns of selfishness, altruism, and group cohesion. Am. Nat. 152:102–113.

Demps, K., F. Zorondo Rodríguez, C. García, and V. Reyes-García. 2012. Social learning across the life cycle: cultural knowledge acquisition for honey collection among the Jenu Kuruba, India. Evol. Hum. Behav. 33:460–470.

Dercole, F., and S. Rinaldi. 2008. Analysis of evolutionary processes: the adaptive dynamics approach and its applications. Princeton Univ. Press, Princeton, NJ.

Dugatkin, L. A. 2004. Principles of animal behavior. W. W. Norton and Company, London.

Enquist, M., and S. Ghirlanda. 2007. Evolution of social learning does not explain the origin of human cumulative culture. J. Theor. Biol. 246:129–135.

Enquist, M., K. Eriksson, and S. Ghirlanda. 2007. Critical social learning: a solution to Rogers's paradox of nonadaptive culture. Am. Anthropol. 109:727–734.

Enquist, M., S. Ghirlanda, A. Jarrick, and C. A. Wachtmeister. 2008. Why does human culture increase exponentially? Theor. Popul. Biol. 74: 46–55.

Eshel, I. 1983. Evolutionary and continuous stability. J. Theor. Biol. 103:99–111.

Feldman, M. W., K. Aoki, and J. Kumm. 1996. Individual versus social learning: evolutionary analysis in a fluctuating environment. Anthropol. Sci. 104:209–231.

Gurven, M., H. Kaplan, and M. Gutierrez. 2006. How long does it take to become a proficient hunter? Implications for the evolution of extended development and long life span. J. Hum. Evol. 51:454–470.

Hewlett, B. S., H. N. Fouts, A. H. Boyette, and B. L. Hewlett. 2011. Social learning among Congo Basin hunter-gatherers. Philos. Trans. R. Soc. Lond. Ser. B 366:1168–1178.

Iwasa, Y., and J. Roughgarden. 1984. Shoot/root balance of plants: optimal growth of a system with many vegetative organs. Theor. Popul. Biol. 25:78–105.

Kaplan, H., and A. Robson. 2002. The emergence of humans: the coevolution of intelligence and longevity with intergenerational transfers. Proc. Natl. Acad. Sci. USA 99:10221–10226.

Kaplan, H. K. Hill., J. Lancaster, and A. M. Hurtado. 2000. A theory of human life history evolution: diet, intelligence, and longevity. Evol. Anthropol. 9:156–185.

Kaplan, H., P. L. Hooper, and M. Gurven. 2009. The evolutionary and ecological roots of human social organization. Philos. Trans. R. Soc. Lond. B 364:3289–3299.

Klein, R. G. 2009. The human career. 3rd ed. The Univ. of Chicago Press, Chicago.

Konner, M. 2010. The evolution of childhood. Belknap, Harvard.

Lehmann, L., and M. W. Feldman. 2009. Coevolution of adaptive technology, maladaptive culture, and population size in a producer-scrounger game. Proc. R. Soc. Lond. Ser. B 276:3853–3862.

Lehmann, L., M. W. Feldman, and R. Kaeuffer. 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.

Lessard, S. 1990. Evolutionary stability: one concept, several meanings. Theor. Popul. Biol. 37:159–170.

Lonsdorf, E. V., L. E. Eberly, and A. E. Pusey. 2004. Sex differences in learning in chimpanzees. Nature 428:715–716.

Maynard Smith, J. 1982. Evolution and the theory of games. Cambridge Univ. Press, Cambridge, U.K.

Mery, F., and T. J. Kawecki. 2004. An operating cost of learning in *Drosophila melanogaster*. Anim. Behav. 68:589–598.

Mesoudi, A. 2011. Variable cultural acquisition costs constrain cumulative cultural evolution. PloS One 6.

Perrin, N. 1992. Optimal resource allocation and the marginal value of organs. Am. Nat. 139:1344–1369.

Pradhan, G. R., C. Tennie, and C. van Schaik. 2012. Social organization and the evolution of cumulative technology in apes and hominins. J. Hum. Evol. 63:180–190.

Reader, S. M., and K. N. Laland. 2001. Primate innovation: sex, age and social rank differences. Int. J. Primatol. 22:787–805.

Reiter, S. 2001. Knowledge, discovery, and growth. *In* G. M. Olson, T. W. Malone, and J. B. Smith, eds. Coordination theory and collaboration technology. Lawrence Erlbaum Associates, Mahwah, NJ.

Rendell, L., L. Fogarty, and K. N. Laland. 2010. Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. Evolution 64:534–548.

Rogers, A. R. 1988. Does biology constrain culture? Am. Anthropol. 90:819–831.

Snell-Rood, E. C., G. Davidowitz, and D. R. Papaj. 2011. Reproductive tradeoffs of learning in a butterfly. Anim. Behav. 22:291–302.

Stephens, D. W. 1991. Change, regularity, and value in the evolution of animal learning. Behav. Ecol. 2:77–89.

Striano, T., M. Tomasello, and P. Rochat. 2001. Social and object support for early symbolic play. Dev. Sci 4:442–455.

Strimling, P., J. Sjöstrand, M. Enquist, and K. Eriksson. 2009. Accumulation of independent cultural traits. Theor. Popul. Biol. 76: 77–83.

Sydsaeter, K., P. Hammond, A. Seierstad, and A. Strøm. 2008. Further mathematics for economic analysis. 2nd ed. Prentice Hall, Essex.

Tennie, C., J. Call, and M. Tomasello. 2009. Ratcheting up the ratchet: on the evolution of cumulative culture. Philos. Trans. R. Soc. Lond. B 364:2405–2415.

Tomasello, M., A. Kruger, and H. Ratner. 1993. Cultural learning. Behav. Brain Sci. 16:495–511.

van Schaik, C., and G. Pradhan. 2003. A model for tool-use traditions in primates: implications for the coevolution of culture and cognition. J. Hum. Evol. 44:645–664.

van Schaik, C. P. 2004. Among orangutans: red apes and the rise of human culture. Harvard Univ. Press, Cambridge, MA.

Wakano, J. Y., K. Aoki, and M. W. Feldman. 2004. Evolution of social learning: a mathematical analysis. Theor. Popul. Biol. 66:249–258.

Associate Editor: J. Wilkins

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web site:

Appendix A. Evolutionarily stable schedule *u(t)*.
Appendix B. ES learning periods, *K* and *L*.
Appendix C. IL and SL during infancy.
Appendix D. Optimal learning length with *G* varying.

Online Supplementary material for "On optimal learning schedules and the marginal value of cumulative cultural evolution"

Appendix A: evolutionarily stable schedule $u(t)$

Here, we derive the candidate ES control schedule $u^*(t)$ by using evolutionary game theory (e.g., Maynard Smith, 1982; Eshel, 1983; Lessard, 1990; Dercole and Rinaldi, 2008) that we couple with optimal control theory (e.g., Iwasa and Roughgarden, 1984; Perrin, 1992; Bulmer, 1994; Day and Taylor, 1998; Sydsaeter et al., 2008).

We are only interested in evaluating the candidate ES control $u^*(t)$ for an individual of age $t = K + \tau$ with $\tau \in [0, L]$. We thus simplify the notations so that we don't need to include K in all equations. To that aim, we write any function $h(t) = h(K+\tau)$ for $\tau \in [0, L]$ as $h(\tau)$. Hence, we replace $u(K + \tau)$ and $A(K + \tau)$ by $u(\tau)$ and $A(\tau)$. This means that in this Appendix A, the boundary conditions of the form $u(0)$ and $A(0)$ refer to $u(K)$ and $A(K)$ in the original process.

Mutant control

In order to evaluate $u^*(t)$, we introduce into a resident population a mutant individual expressing a deviant control $u_{\mathrm{m}}(t)$. We then ask under what conditions is the resident immune to invasion by the mutant.

When the population is small, the mutant may affect the dynamic of $A(t)$ of different individuals in the population, but we assume that the population is large enough so that we can neglect the effect of a single individual on $P(t)$. Under this assumption, the amount of information held by a mutant, denoted by $A_{\mathrm{m}}(t)$ for $t \in [0, L]$, satisfies

$$\dot{A}_{\mathrm{m}}(t) = -\epsilon A_{\mathrm{m}}(t) + u_{\mathrm{m}}(t)\mu + (1 - u_{\mathrm{m}}(t))\,\beta\,(P(t) - A_{\mathrm{m}}(t)) \tag{A-1}$$

with boundary condition

$$A_{\mathrm{m}}(0) = \gamma A(L), \tag{A-2}$$

where the dynamics of $A(t)$ and $P(t)$ are assumed to be given by the resident system (eq. 4 of the main text). With the time notation of this Appendix, the resident system is for $t \in [0, L]$

$$\begin{aligned} \dot{A}(t) &= -\epsilon A(t) + u(t)\mu + (1 - u(t))\,\beta\,(P(t) - A(t)) \\ \dot{P}(t) &= -\epsilon P(t) + u(t)\mu\,(1 + \rho)\,, \end{aligned} \tag{A-3}$$

where

$$\begin{aligned} A(0) &= \gamma A(L) \\ P(0) &= \gamma P(L). \end{aligned} \tag{A-4}$$

Importantly, eq. A-2 not only implies that the population is large but also that a mutant can acquire information from the parental generations only from resident individuals. This rules out the possibility that it learns information from its parents. We endorse this assumption because taking into account pure vertical transmission of information would require to track intergenerational effects of mutant-mutant interactions, which is much more complicated.

Before proceeding to the invasion analysis, we note that according to eq. A-3, an individual may gain from SL only insofar there is information $D(t) = P(t) - A(t)$ in the population it lacks. From eq. A-3, this quantity satisfies

$$\dot{D}(t) = -\left[\epsilon + (1 - u(t))\,\beta\right] D(t) + u(t)\mu\rho. \tag{A-5}$$

We will use the variable $D(t)$ in our analyses below (instead of $P(t)$) as it leads to more compact formulations.

Maximum principle

Standard end constraints optimal control problem

In order to find the ES learning schedule, we use optimal control theory under the form of Pontryagin's maximum principle (see Sydsaeter et al., 2008 for a detailed overview). From eq. 1 of the main text, the objective function of the mutant is

$$f(A_{\mathrm{m}}) = \varphi\left[(G - K - L) + \phi A_{\mathrm{m}}(L)\left(\frac{1 - e^{-\epsilon(G-K-L)}}{\epsilon}\right)\right] \tag{A-6}$$

subject to the dynamic constraint eqs. A-1–A-2, which themselves depend on the resident system eq. A-3. This is an optimal control problem with only a terminal payoff (or scrap value), where $u_{\mathrm{m}}(t) \in [0, 1]$ is the control, and $A_{\mathrm{m}}(t)$ and $P(t)$ are the state variables with $A_{\mathrm{m}}(L)$ and $P(L)$ free (Sydsaeter et al., 2008, p. 364–365).

A candidate ES control $u^*(t)$ of a control triple $(u^*(t),\ A^*(t),\ P^*(t))$ is a mutant phenotype $u_{\mathrm{m}}(t)$ that for all $t \in [0, L]$ maximizes the Hamiltonian

$$\begin{aligned} H(u_{\mathrm{m}}(t), u(t), A^*(t), P^*(t)) = {} & \lambda_{\mathrm{A}}(t)\left[-\epsilon A^*(t) + u_{\mathrm{m}}(t)\mu + (1 - u_{\mathrm{m}}(t))\,\beta\,(P^*(t) - A^*(t))\right] \\ & + \lambda_{\mathrm{P}}(t)\left[-\epsilon P^*(t) + u(t)\mu\,(1 + \rho)\right], \end{aligned} \tag{A-7}$$

whose maximum is evaluated at $u_{\mathrm{m}}(t) = u(t) = u^*(t)$. The state variables satisfy eq. A-3 evaluated at $(u^*(t),\ A^*(t),\ P^*(t))$ and the co-state variables satisfy

$$\begin{aligned} \dot{\lambda}_{\mathrm{A}} &= -\frac{\partial H(u^*, u^*, A^*, P^*)}{\partial A^*} \\ \dot{\lambda}_{\mathrm{P}} &= -\frac{\partial H(u^*, u^*, A^*, P^*)}{\partial P^*} \end{aligned} \tag{A-8}$$

with terminal conditions

$$\begin{aligned}\lambda_{\mathrm{A}}(L) &= \frac{\partial f(A^*)}{\partial A^*} \geq 0 \\ \lambda_{\mathrm{P}}(L) &= \frac{\partial f(A^*)}{\partial P^*} = 0.\end{aligned} \tag{A-9}$$

Applying the maximum principle

From eq. A-7, the change in the Hamiltonian due to an individual expressing the mutant control and evaluated at a candidate ESS is

$$\left.\frac{\partial H(t)}{\partial u_{\mathrm{m}}}\right|_{u_{\mathrm{m}}(t)=u^*(t)} = \lambda_{\mathrm{A}}(t)\left[\mu - \beta D^*(t)\right], \tag{A-10}$$

which gives the selection pressure on increasing the level of individual learning at time t. From eq. A-5, $D^*(t)$ satisfies

$$\dot{D}^*(t) = -\left[\epsilon + (1-u^*(t))\,\beta\right] D^*(t) + u^*(t)\mu\rho \tag{A-11}$$

with boundary condition

$$D^*(0) = \gamma D^*(L) < D^*(L). \tag{A-12}$$

Hence, from eqs. A-10–A-11, an evolutionarily stable (ES) control must satisfy:

$$u^*(t) = \begin{cases} 0 & \text{when } D^*(t) > \mu/\beta \\ \frac{\epsilon+\beta}{\beta(1+\rho)} & \text{when } D^*(t) = \mu/\beta \\ 1 & \text{when } D^*(t) < \mu/\beta, \end{cases} \tag{A-13}$$

which shows that $u^*(t)$ can be a step function.

Candidate ES schedules

Here, we find candidate ES schedules among different alternatives. To that aim, we will use the explicit solution of eq. A-11 with $u^*(t) = u$ being constant and boundary condition $D(0)$ at $t = 0$, which produces

$$D^*(t) = \frac{\rho\mu u}{\epsilon + (1-u)\beta} + \left[D^*(0) - \frac{\rho\mu u}{\epsilon + (1-u)\beta}\right] e^{-[\epsilon+(1-u)\beta]t}. \tag{A-14}$$

Let us now consider some specific learning schedules.

(1) $u^*(t) = 0$

From eq. A-14, $D^*(t)$ is monotone decreasing, which violates inequality (A-12). Thus, this is not a candidate for an ES control.

(2) $u^*(t) = 1$

From eq. A-12 and eq. A-14,

$$D^*(L) = \frac{\rho\mu}{\epsilon} + \left[D^*(0) - \frac{\rho\mu}{\epsilon}\right] e^{-\epsilon L} = \frac{1}{\gamma} D^*(0). \tag{A-15}$$

Hence

$$D^*(0) = \frac{\rho\mu}{\epsilon} \frac{1 - e^{-\epsilon L}}{\gamma^{-1} - e^{-\epsilon L}} < \frac{\rho\mu}{\epsilon}, \tag{A-16}$$

Hence, $D^*(t)$ in eq. A-14 is monotone increasing. Thus, consistent with the third line of eq. A-13 if

$$D^*(L) = \frac{\rho\mu}{\epsilon} \frac{1 - e^{-\epsilon L}}{1 - \gamma e^{-\epsilon L}} < \frac{\mu}{\beta}, \tag{A-17}$$

in other words

$$\frac{\rho\beta}{\epsilon} < \frac{1 - \gamma e^{-\epsilon L}}{1 - e^{-\epsilon L}}. \tag{A-18}$$

Thus, $u^*(t) = 1$ for $0 \leq t \leq L$ is a candidate for an ES control.

Let us now calculate $A^*(t)$ associated with this learning schedule. From eq. A-3, this satisfies

$$\dot{A}^*(t) = -\epsilon A^*(t) + \mu, \tag{A-19}$$

which gives

$$A^*(t) = \frac{\mu}{\epsilon}\left[1 - e^{-\epsilon t}\left(1 - \frac{A^*(0)\epsilon}{\mu}\right)\right]. \tag{A-20}$$

Using this equation and the boundary condition $A^*(0) = \gamma A^*(L)$, we have

$$A^*(0) = \gamma\frac{\mu}{\epsilon}\left[1 - e^{-\epsilon L}\left(1 - \frac{A^*(0)\epsilon}{\mu}\right)\right], \tag{A-21}$$

whose solution is

$$A^*(0) = \frac{\gamma\mu\left(e^{L\epsilon} - 1\right)}{\epsilon\left(e^{L\epsilon} - \gamma\right)}. \tag{A-22}$$

On substitution into eq. A-20 yields

$$A^*(t) = \frac{\mu}{\epsilon}\left[1 - \frac{e^{\epsilon(L-t)}(1-\gamma)}{e^{L\epsilon} - \gamma}\right], \tag{A-23}$$

which, when $t = L$, reduces to

$$A^*(L) = \frac{\mu}{\epsilon}\left[\frac{e^{L\epsilon} - 1}{e^{L\epsilon} - \gamma}\right], \tag{A-24}$$

and on substitution of eq. 5 of the main text produces eq. 8.

(3) $u^*(t) = 0$ for $0 \leq t \leq t_{\mathrm{s}}$ and $u^*(t) = 1$ for $t_{\mathrm{s}} < t \leq L$.

This schedule requires $D^*(0) > \mu/\beta > D^*(L)$, which contradicts inequality (A-12). Thus, this is not a candidate for an ES control schedule.

(4) $u^*(t) = 1$ for $0 \leq t < t_\mathrm{s}$ and $u^*(t) = 0$ for $t_\mathrm{s} \leq t \leq L$ This schedule requires $D^*(0) < \mu/\beta$ for $0 \leq t < t_\mathrm{s}$, $D^*(t_\mathrm{s}) = \mu/\beta$, and $D^*(0) > \mu/\beta$ for $t_\mathrm{s} \leq t \leq L$. But $u^*(t) = 0$ entails that $D^*(t)$ is monotone decreasing in this interval. Hence, contradiction. Thus, this is not a candidate for an ES control.

(5) $u^*(t) = 1$ for $0 \leq t \leq t_\mathrm{s}$ and $u^*(t) = (\epsilon + \beta)/\left[\beta(1+\rho)\right]$ for $t_\mathrm{s} < t \leq L$.

Note that $0 < (\epsilon + \beta)/\left[\beta(1+\rho)\right] < 1$ entails that

$$\frac{\rho\beta}{\epsilon} > 1. \tag{A-25}$$

Since $u^*(t) = 1$ for $0 \leq t \leq t_\mathrm{s}$, eq. A-14 entails for this intervall

$$D^*(t) = \frac{\rho\mu}{\epsilon} + \left[D^*(0) - \frac{\rho\mu}{\epsilon}\right] e^{-\epsilon t}, \tag{A-26}$$

while

$$D^*(t) = \frac{\mu}{\beta} \tag{A-27}$$

for $t_\mathrm{s} < t \leq L$. Continuity of $D^*(t)$ at $t = t_\mathrm{s}$ entails

$$\frac{\rho\mu}{\epsilon} - \frac{\mu}{\beta} = \left[\frac{\rho\mu}{\epsilon} - D^*(0)\right] e^{-\epsilon t_\mathrm{s}}, \tag{A-28}$$

where the left member must be positive from inequality (A-25). Hence, we must have

$$D^*(0) < \frac{\mu}{\beta} < \frac{\rho\mu}{\epsilon}. \tag{A-29}$$

If this inequality holds, then $D^*(t)$ increases towards μ/β when $0 \leq t \leq t_\mathrm{s}$, which is a necessary condition for $u^*(t) = 1$ in this interval. Inequality (A-29) is also consistent with the constraint given by eq. A-12, since $D^*(L) = \mu/\beta$. Thus, this is a candidate for an ES control.

Let us now determine the switching time t_s. Eq. A-12 and eq. A-28 entail $D^*(0) = \gamma\mu/\beta$, which, on substitution in eq. A-28, yields

$$\frac{\rho\beta}{\epsilon} - 1 = \left[\frac{\rho\beta}{\epsilon} - \gamma\right] e^{-\epsilon t_\mathrm{s}}. \tag{A-30}$$

Rearranging,

$$\frac{\rho\beta}{\epsilon} = \frac{1 - \gamma e^{-\epsilon t_\mathrm{s}}}{1 - e^{-\epsilon t_\mathrm{s}}}, \tag{A-31}$$

where the right member is monotone decreasing in t_s. Hence, the valid solution

$$t_\mathrm{s} = \frac{1}{\epsilon} \log\left[\frac{\beta\rho - \gamma\epsilon}{\beta\rho - \epsilon}\right] \tag{A-32}$$

with $0 < t_\mathrm{s} < L$ exists only if inequality (A-18) is reversed; that is, if

$$\frac{\rho\beta}{\epsilon} > \frac{1-\gamma e^{-\epsilon L}}{1-e^{-\epsilon L}}. \tag{A-33}$$

Let us now determine $A^*(t)$ associated with this learning schedule. For, $0 \leq t \leq t_\mathrm{s}$, we have from eq. A-3 and $u^*(t) = 1$ that

$$\dot{A}^*(t) = -\epsilon A^*(t) + \mu. \tag{A-34}$$

For $t > t_\mathrm{s}$, we have $D^*(t) = \mu/\beta$. Hence, from eq. A-3 the trait number $A^*(t)$ still satisfies eq. A-34. Therefore, eq. A-34 is satisfied by $A^*(t)$ for $0 \leq t \leq L$ and the solution of this dynamic was calculated before and is given by eq. A-23

Appendix B: ES learning periods, K and L

Here, we evaluate the candidate ES learning periods, K^* and L^*, by holding $u^*(t)$ fixed and performing an invasion analysis. In the original time index $t \in [0, G]$, the equilibrium amount of information held by an individual of age $K+L$ in a resident population is obtained from eq. A-24 with $\gamma = \left[1 - e^{-\alpha K}\right] e^{-\epsilon(G-L)}$, which gives

$$A(K+L) = \frac{\mu\left(e^{\epsilon L}-1\right) e^{\alpha K + \epsilon(G-L)}}{\epsilon\left[1 + \left(e^{\epsilon G}-1\right) e^{\alpha K}\right]}. \tag{B-1}$$

First order condition

From eq. 1 of the main text, the reproductive output of an individual with mutant trait values K_m and L_m is

$$f(K_\mathrm{m}, L_\mathrm{m}) = \varphi\left[(G - K_\mathrm{m} - L_\mathrm{m}) + \left(\frac{1-e^{-\epsilon(G-K_\mathrm{m}-L_\mathrm{m})}}{\epsilon}\right)\phi A(K_\mathrm{m}+L_\mathrm{m})\right], \tag{B-2}$$

where

$$A(K_\mathrm{m}+L_\mathrm{m}) = A(G)\left(e^{\alpha K_\mathrm{m}}-1\right) e^{-[(\alpha+\epsilon)K_\mathrm{m}+\epsilon L_\mathrm{m}]} + \frac{\mu}{\epsilon}\left(1-e^{-\epsilon L_\mathrm{m}}\right), \tag{B-3}$$

where the first term is the number of traits obtained by an individual during the infancy period (through oblique SL), while the second term is the number of traits acquired during the juvenile period. Here

$$A(G) = \frac{\mu\left(e^{L\epsilon}-1\right) e^{K(\alpha+\epsilon)}}{\epsilon\left[1+\left(e^{G\epsilon}-1\right)e^{\alpha K}\right]}, \tag{B-4}$$

which is the equilibrium amount of information of an individual of age G in the resident population (obtained from eq. B-1 and $A(G) = A(K+L)e^{-\epsilon(G-K-L)}$). In a resident population, where $K_\mathrm{m} = K$, $L_\mathrm{m} = L$, eq. B-3 reduces to eq. B-1.

Eq. B-3 is obtained by assuming that a mutant individual affects its reproductive output only through the effect on the length of its own learning periods (K_{m} and L_{m}). For the infancy learning phase, the amount of information held by a mutant in $t \in [0, K_{\mathrm{m}}]$ is obtained from eq. 2 of the main text by solving

$$\dot{A}(t) = -\epsilon A(t) + \alpha \left[A(G)e^{-\epsilon t} - A(t)\right] \tag{B-5}$$

with initial condition $A(0) = 0$ and where $A(G)$ is given by eq. B-4. For the juvenile learning phase with optimal control schedule $u^*(t) = 1$, one has

$$\dot{A}(t) = -\epsilon A(t) + \mu \tag{B-6}$$

for $t \in [K_{\mathrm{m}}, K_{\mathrm{m}} + L_{\mathrm{m}}]$, whose solution gives eq. B-3 when the initial condition $A(K_{\mathrm{m}})$ is obtained by solving eq. B-5.

Eq. B-6 also holds for a monomorphic population with mixed control $u^*(t) = (\epsilon + \beta)/\left[\beta(1+\rho)\right]$ schedule (eq. A-34). So we assume that introducing the mutant for this case only affects the boundary conditions of eqs. B-6–B-5, exactly in the same way as for the pure strategy. Hence, we can use eq. B-3 in eq. B-2 for both the pure and mixed candidate ES control schedules. That eq. B-3 can also be used for the mixed strategy in proved more rigorously below (eqs. B-22–B-25).

In order to find K^* and L^*, we have to find the maximum of $f(K_{\mathrm{m}}, L_{\mathrm{m}})$ with respect to both K_{m} and L_{m} given the constraint $K_{\mathrm{m}} + L_{\mathrm{m}} < G$. This is a nonlinear programming problem with inequality constraints (Sydsaeter et al., 2008, p. 129). Because of the complementary slackness conditions, the first-order condition for $0 < K^*$, $0 < L^*$, and $K^* + L^* < G$ is given by

$$\begin{aligned} \frac{\partial f(K_{\mathrm{m}}, L)}{\partial K_{\mathrm{m}}} &= 0 \\ \frac{\partial f(K, L_{\mathrm{m}})}{\partial L_{\mathrm{m}}} &= 0, \end{aligned} \tag{B-7}$$

where the derivatives are evaluated at $K_{\mathrm{m}} = K = K^*$, $L_{\mathrm{m}} = L = L^*$. The candidate optimal values K^* and L^* are then obtained by solving eq. B-7.

Note that the first order condition (eq. B-7) follow from the assumption that each of the two trait values, K and L, evolve by holding the other control constant (i.e., no pleiotropy or double mutants). In other words, we assume that only one mutant control, K_{m} or L_{m}, deviates from the resident control at any one time.

ES learning lengths

Evolution of L

From eq. B-2, the selection gradient on L is

$$\frac{f(K_{\mathrm{m}}, L_{\mathrm{m}})}{\partial L_{\mathrm{m}}} = \varphi \left[\frac{\phi}{\epsilon}\left(1 - e^{-\epsilon(G-K-L)}\right)\frac{A(K_{\mathrm{m}} + L_{\mathrm{m}})}{\partial L_{\mathrm{m}}} - \left(1 + \phi A(K+L)e^{-\epsilon(G-K-L)}\right)\right],$$

$$\tag{B-8}$$

where the derivatives are evaluated at $K_{\mathrm{m}} = K$ and $L_{\mathrm{m}} = L$. The first term in the brackets is the fecundity benefit of acquiring more information from learning by increasing L, while the second term is the concomitant cost because less time is spent on foraging. The benefit depends on

$$\frac{A(K_{\mathrm{m}} + L_{\mathrm{m}})}{\partial L_{\mathrm{m}}} = e^{-L\epsilon}\left(\mu - \epsilon A(G)\left(e^{\alpha K} - 1\right)e^{-K(\alpha+\epsilon)}\right). \tag{B-9}$$

When $K = 0$ and the initial learning period is also zero ($L = 0$), we have $A(G) = 0$ and the selection gradient reduces to

$$\frac{f(0, L_{\mathrm{m}})}{\partial L_{\mathrm{m}}} = \frac{\mu\phi}{\epsilon}\left[1 - e^{-\epsilon G}\right] - 1. \tag{B-10}$$

Hence, positive L is selected for if

$$\frac{\mu\phi}{\epsilon}\left[1 - e^{-\epsilon G}\right] - 1 > 0. \tag{B-11}$$

Solving $\partial f(0, L_{\mathrm{m}})/\partial L_{\mathrm{m}} = 0$ evaluated at $L_{\mathrm{m}} = L = L^*$, we find that

$$L^* = \frac{1}{\epsilon}\log\left(\frac{e^{\frac{G\epsilon}{2}}\sqrt{\epsilon^2 e^{G\epsilon} + 4\mu^2\phi^2} - \epsilon e^{G\epsilon}}{2\mu\phi}\right). \tag{B-12}$$

A first order Taylor expansion around $\epsilon = 0$ yields

$$L^* = \frac{G}{2} - \frac{2 + G\epsilon}{4\mu\phi} + O(\epsilon^2). \tag{B-13}$$

On substitution into eq. B-1 and Taylor expanding again around $\epsilon = 0$ gives

$$A^*(L^*) = \frac{G\mu}{2} - \frac{1}{2\phi} - \frac{\epsilon\left(1 + G^2\mu^2\phi^2\right)}{8\mu\phi^2} + O(\epsilon^2). \tag{B-14}$$

Coevolution K and L

From eq. B-2, the selection gradient on K is

$$\frac{f(K_{\mathrm{m}}, L_{\mathrm{m}})}{\partial K_{\mathrm{m}}} = \varphi\left[\frac{\phi}{\epsilon}\left(1 - e^{-\epsilon(G-K-L)}\right)\frac{A(K_{\mathrm{m}} + L_{\mathrm{m}})}{\partial K_{\mathrm{m}}} - \left(1 + \phi A(K+L)e^{-\epsilon(G-K-L)}\right)\right], \tag{B-15}$$

where

$$\frac{A(K_{\mathrm{m}} + L_{\mathrm{m}})}{\partial K_{\mathrm{m}}} = A(G)\left[\alpha - \epsilon\left(e^{\alpha K} - 1\right)\right]e^{-K(\alpha+\epsilon)-L\epsilon}. \tag{B-16}$$

Evaluating eq. B-15 at $K = 0$ and $L = L^*$ by using eq. B-12 and letting $\epsilon \to 0$ produces

$$\frac{f(K_{\mathrm{m}}, L_{\mathrm{m}})}{\partial K_{\mathrm{m}}} = \frac{(1 + G\mu\phi)(\alpha G\mu\phi - \alpha - 2\mu\phi)}{4\mu\phi}. \tag{B-17}$$

Hence selection increases K when

$$\alpha > \frac{2\mu\phi}{G\mu\phi - 1}. \tag{B-18}$$

When this condition is satisfied, L and K coevolve, and the first order condition (eq. B-7) informs us that at a candidate ESS we have $\partial f(K_\mathrm{m}, L)/\partial K_\mathrm{m} - \partial f(K, L_\mathrm{m})/\partial L_\mathrm{m} = 0$. Substituting eq. B-8 and eq. B-15 and taking the limit $\epsilon \to 0$, we find from this equality that the candidate juvenile learning period is

$$L^* = \frac{1}{\alpha}. \tag{B-19}$$

Substituting this solution into the first order condition $\partial f(K_\mathrm{m}, L)/\partial K_\mathrm{m} = 0$ produces

$$K^* = G - \frac{1}{\mu\phi} - \frac{1}{\alpha}\left[1 + W\left(e^{\alpha G - 1 - \frac{\alpha}{\mu\phi}}\right)\right], \tag{B-20}$$

where $W(x)$ is the principal solution of the Lambert function defined as $x = We^W$. Using the candidate solutions in eq. B-1 gives

$$A^*(K^* + L^*) = \mu L^* e^{\alpha K^*}. \tag{B-21}$$

Mixed strategy case

Here, we show that eq. B-3 can also be used to evaluate the optimal K^* and L^* for the mixed control schedule $u^*(t) = (\epsilon + \beta)/\left[\beta(1+\rho)\right]$. For this case, the learning dynamics during infancy still follows eq. B-5. For the juvenile learning phase, the amount of information held by a mutant first follows eq. B-6 for $t \in [K_\mathrm{m}, K_\mathrm{m} + t_\mathrm{s}]$, where t_s is held at the resident control value owing to the assumption of no pleitropy. This means that only K or L varies, while everything else is kept constant. From the switching time onwards; that is, for $t \in [K_\mathrm{m} + t_\mathrm{s}, L_\mathrm{m}]$, the amount of information of a mutant during the juvenile period is obtained from eq. A-3 as

$$\dot{A}(t) = a + bA(t) + cP(t), \tag{B-22}$$

where $a = \frac{\mu(\beta+\epsilon)}{\beta(1+\rho)}$, $b = -\frac{\rho(\beta+\epsilon)}{1+\rho}$, and $c = \frac{\beta\rho-\epsilon}{1+\rho}$. From this, we have

$$A(K_\mathrm{m} + L_\mathrm{m}) = A(K_\mathrm{m} + t_\mathrm{s})e^{b(L_\mathrm{m} - t_\mathrm{s})} + \int_{t_\mathrm{s}}^{L_\mathrm{m}} e^{b(L_\mathrm{m} - \tau)}\left[a + cP(t)\right]\mathrm{d}\tau. \tag{B-23}$$

Because $P(t)$ is evaluated at the resident phenotypic value, its value may change discontinuously in $t \in [t_\mathrm{s}, L_\mathrm{m}]$. However, we do not need to evaluate the integral in eq. B-23 explicitly in order to obtain the first order conditions (eq. B-7). To that aim, it is sufficient to evaluate the derivatives of $A(K_\mathrm{m} + L_\mathrm{m})$ with respect to K_m and L_m at the resident phenotypic values.

In order to evaluate the variation of $A(K_\mathrm{m} + L_\mathrm{m})$ in K_m, it suffices to set $L_\mathrm{m} = L$ in eq. B-23 and substitute for $P(t)$ (obtained from eq. A-3) and $A(K_\mathrm{m} + t_\mathrm{s})$ (obtained from eq. B-5–B-6). Then, the right member of eq. B-23 reduces to the right member of eq. B-3 with $L_\mathrm{m} = L$.

In order to evaluate the variation of $A(K_\mathrm{m}+L_\mathrm{m})$ in L_m, we have from eq. B-23

$$\frac{\partial A(K+L_\mathrm{m})}{\partial L_\mathrm{m}} = \frac{\partial}{\partial L_\mathrm{m}}\left(A(K+t_\mathrm{s})e^{b(L_\mathrm{m}-t_\mathrm{s})}\right)\bigg|_{L_\mathrm{m}=L} + [a+cP(L)] + \int_{t_\mathrm{s}}^{L}\frac{\partial}{\partial L_\mathrm{m}}\left(e^{b(L_\mathrm{m}-\tau)}\right)\bigg|_{L_\mathrm{m}=L}[a+cP(t)]\,\mathrm{d}\tau, \quad \text{(B-24)}$$

which depends only on resident phenotypic values. Substituting the values for $P(t)$ and $A(K+t_\mathrm{s})$, it can be shown that

$$\frac{\partial A(K+L_\mathrm{m})}{\partial L_\mathrm{m}} = e^{-\epsilon L}\left(\mu - \epsilon A(K)\right), \quad \text{(B-25)}$$

which is equivalent as computing $\partial A(K+L_\mathrm{m})/\partial L_\mathrm{m}$ by using the right member of eq. B-3. Hence, eq. B-3 can be used to evaluate the candidate L^* and K^* for both the pure and mixed control schedules.

Appendix C: IL and SL during infancy

In this appendix, we show that allowing for IL during the infancy period does not qualitatively affect our results. Our aim is to show that when the innovation rate μ_I during infancy is small, pure SL will be favored at the beginning of infancy. This was our working assumption in the main text.

In order to allow for both IL and SL during infancy (for $t \in [0, K]$), we extend eq. 2 of the main text along the lines of eq. 4. This gives

$$\dot{A}(t) = -\epsilon A(t) + u(t)\mu_\mathrm{I} + (1-u(t))\alpha\left[A(G)e^{-\epsilon t} - A(t)\right]. \quad \text{(C-1)}$$

Applying the maximum principle as in Appendix A, the change in the Hamiltonian due to an individual expressing a mutant control during infancy and evaluated at a candidate ESS is

$$\frac{\partial H(t)}{\partial u_\mathrm{m}}\bigg|_{u_\mathrm{m}(t)=u^*(t)} \propto \mu_\mathrm{I} - \alpha\left[A^*(G)e^{-\epsilon t} - A^*(t)\right], \quad \text{(C-2)}$$

where

$$\dot{A}^*(t) = -\epsilon A^*(t) + u^*(t)\mu_\mathrm{I} + (1-u^*(t))\alpha\left[A^*(G)e^{-\epsilon t} - A^*(t)\right], \quad \text{(C-3)}$$

with initial condition $A^*(0) = 0$.

When $A^*(G) < e^{t\epsilon}(\mu_\mathrm{I} + \alpha A^*(t))/\alpha$, the right member of eq. C-2 is positive, whereby $u^*(t) = 1$. For $t = 0$, this inequality holds if $A^*(G) < \mu_\mathrm{I}/\alpha$, which implies that the inequality holds for all $t \in [0, K]$ since $e^{t\epsilon} \geq 0$ and $A^*(t) \geq 0$. Hence, if $u^*(0) = 1$, the optimal control is pure IL throughout.

When $A^*(G) > e^{t\epsilon}(\mu_\mathrm{I} + \alpha A^*(t))/\alpha$, the right member of eq. C-2 is negative, which implies $u^*(t) = 0$. For $t = 0$, this inequality holds if $A^*(G) > \mu_\mathrm{I}/\alpha$, in which case we have pure SL to begin with: $u^*(0) = 0$. This is likely to obtains when μ_{I} is small. As a consequence of social learning, $A^*(t)$ increases and may eventually reach the point where $A^*(G) = e^{t\epsilon}(\mu_{\mathrm{I}} + \alpha A^*(t))/\alpha$. Can a mixed strategy be favored for some time interval when this inequality is satisfied, and which entails that $A^*(t) = A^*(G)e^{-\epsilon t} - \mu_{\mathrm{I}}/\alpha$. The mixed strategy must satisfy eq. C-3 with $A^*(t) = A^*(G)e^{-\epsilon t} - \mu_{\mathrm{I}}/\alpha$ and on substitution of the latter equation, we have

$$1 + \epsilon/\alpha = 0, \tag{C-4}$$

which is in contradiction with the fact that $\epsilon \geq 0$ and $\alpha \geq 0$. This suggests that we have a bang-bang optimal control when $A^*(G) > \mu_{\mathrm{I}}/\alpha$: $u^*(t) = 0$ for $t < t_{\mathrm{s}} \leq K$, and $u^*(t) = 1$ when $t_{\mathrm{s}} \leq t \leq K$, where t_{s} is the switching time.

Appendix D: optimal learning length with G varying

In our analysis we assumed that the generation length G was a fixed value. Because L and K evolved, selection determined the length $M = G - L - K$ of adulthood (or mature period). We now assume that M is fixed, and study the co-evolution of L and K under this assumption. This entails that $G = L + K + M$ is no longer fixed but will be an outcome of selection on L and K by holding M fixed. For this case, the derivatives in the first order condition (eq. B-7) will be evaluated at $K_{\mathrm{m}} = K = K^*$, $L_{\mathrm{m}} = L = L^*$, and $G = L^* + K^* + M$.

Evolution of L

Evaluating $\partial f(0, L_{\mathrm{m}})/\partial L_{\mathrm{m}} = 0$ at $L = 0$ and $G = M$, we find that L is selected for if

$$\frac{\mu\phi}{\epsilon}\left[1 - e^{-\epsilon M}\right] - 1 > 0. \tag{D-1}$$

A necessary condition for this inequality to hold is $M > 0$ and that the innovation rate multiplied by the reliance on non-genetically encoded information exceeds the decay rate ($\mu\phi > \epsilon$). This is essentially the same result as in eq. B-11.

Solving $\partial f(0, L_{\mathrm{m}})/\partial L_{\mathrm{m}} = 0$ evaluated at L^*, we find that

$$L^* = \frac{\log\left(\frac{\epsilon}{\mu\phi} + e^{-\epsilon M}\right)}{\log\left(e^{-\epsilon}\right)} \tag{D-2}$$

and substituting into eq. B-1 by setting $K = 0$ gives

$$A^*(L^*) = \frac{\mu}{\epsilon}\left[1 - e^{-\epsilon M}\right] - 1/\phi. \tag{D-3}$$

Coevolution of K and L

Given that eq. B-11 is satisfied, the juvenile learning period may evolve towards a candidate ES learning length. Then, solving eq. B-7 at $K_{\mathrm{m}} = K = K^*$, $L_{\mathrm{m}} = L = L^*$, and $G =$

$L^* + K^* + M$, we obtain

$$\begin{aligned} K^* &= \frac{1}{\epsilon+\alpha}\left[\log\left(\alpha\left[\frac{1-e^{-M\epsilon}}{\epsilon}-\frac{1}{\mu\phi}\right]\right)-M\epsilon\right] \\ L^* &= \frac{1}{\epsilon}\log\left(\frac{\alpha-\epsilon e^{\alpha K^*}+\epsilon}{\alpha-\epsilon e^{K^*(\alpha+\epsilon)+M\epsilon}}\right). \end{aligned} \tag{D-4}$$

Although these functions are relatively complicated they are analytic and allow us to evaluate explicitly the amount of information held by an individual at equilibrium at the end of the learning phases. This is

$$A^*(K^*+L^*) = \frac{\mu}{\epsilon}\left[1-e^{-\epsilon M}\right]-1/\phi, \tag{D-5}$$

where the right member is identical to that of eq. D-3. This is independent of α. Hence, whether oblique SL evolves or not results in exactly the same amount of information held by an individual at steady state. In other words, Rogers' paradox now applies to both oblique and horizontal SL.

References

Bulmer, M., 1994. Theoretical Evolutionary Ecology. Sinauer Associates, Massachusetts.

Day, T. and P. Taylor, 1998. The evolution of temporal patterns of selfishness, altruism, and group cohesion. American Naturalist 152:102–113.

Dercole, F. and S. Rinaldi, 2008. Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications. Princeton University Press, Princeton, NJ.

Eshel, I., 1983. Evolutionary and continuous stability. Journal ot Theoretical Biology 103:99–111.

Iwasa, Y. and J. Roughgarden, 1984. Shoot/root balance of plants: optimal growth of a system with many vegetative organs. Theoretical Population Biology 25:78–105.

Lessard, S., 1990. Evolutionary stability: one concept, several meanings. Theoretical Population Biology 37:159–170.

Maynard Smith, J., 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge.

Perrin, N., 1992. Optimal resource allocation and the marginal value of organs. American Naturalist 139:1344–1369.

Sydsaeter, K., P. Hammond, A. Seierstad, and A. Strøm, 2008. Further Mathematics for Economic Analysis. 2th ed. Prentice Hall, Essex.