

# Evolution of learning strategies in temporally and spatially variable environments: A review of theory



Kenichi Aoki<sup>a,\*</sup>, Marcus W. Feldman<sup>b</sup>

<sup>a</sup> Organization for the Strategic Coordination of Research and Intellectual Properties, Meiji University, Nakano 4-21-1, Nakano-ku, Tokyo 164-8525, Japan

<sup>b</sup> Department of Biology, Stanford University, Stanford CA 94305, USA

## ARTICLE INFO

### Article history:

Available online 7 November 2013

### Keywords:

Dynamical models  
Evolutionarily stable strategies  
Monte Carlo/agent-based simulations  
Rogers' paradox

## ABSTRACT

The theoretical literature from 1985 to the present on the evolution of learning strategies in variable environments is reviewed, with the focus on deterministic dynamical models that are amenable to local stability analysis, and on deterministic models yielding evolutionarily stable strategies. Individual learning, unbiased and biased social learning, mixed learning, and learning schedules are considered. A rapidly changing environment or frequent migration in a spatially heterogeneous environment favors individual learning over unbiased social learning. However, results are not so straightforward in the context of learning schedules or when biases in social learning are introduced. The three major methods of modeling temporal environmental change – coevolutionary, two-timescale, and information decay – are compared and shown to sometimes yield contradictory results. The so-called Rogers' paradox is inherent in the two-timescale method as originally applied to the evolution of pure strategies, but is often eliminated when the other methods are used. Moreover, Rogers' paradox is not observed for the mixed learning strategies and learning schedules that we review. We believe that further theoretical work is necessary on learning schedules and biased social learning, based on models that are logically consistent and empirically pertinent.

© 2013 Elsevier Inc. All rights reserved.

## 1. Introduction

Learning is a means of acquiring information about the environment and of expressing a phenotype (behavior) appropriate to that environment. Two forms of learning may be distinguished by the source of the information acquired. Individual learning (IL) occurs when an organism depends on its personal experience to gather the information directly from the environment, e.g., by trial-and-error. The second form of learning is social learning (SL), which occurs when an organism obtains the information indirectly by copying other organisms, e.g., by imitation.

A learning strategy is the way in which an organism combines IL and SL, either simultaneously or sequentially, and its relative dependence on each. Biases associated with SL in the choice of whom to copy are also an integral part of a learning strategy. The simplest strategies involve the use of IL or SL but not both. Each learning strategy can be regarded as a genetic adaptation to a specific kind of environmental variability. A learning strategy supports culture, to the extent that an innovation produced by IL is propagated through the population by SL. The learning strategy available to a

species will – in conjunction with other factors such as its demography – determine the nature and properties of its culture.

Evolutionary models of learning, the subject of this review, are to be distinguished from classical learning models in psychology, which were constructed as mathematical formulations for how to assess the probabilities of alternative behaviors upon presentation of stimuli to a subject. These probabilities changed dynamically so that the subject's behavior over time would also change. The focus was on modifications of individual behavior over the course of such trials (Bush and Mosteller, 1955; Hanania, 1959). Extensions of such models have been made to competitive situations where the members of a set of players adopt behaviors at each time step that depend on the history of decisions made by all the players (e.g., Izquierdo and Izquierdo, 2008). Common applications allow players to choose one of two behaviors, and the time-dependent and asymptotic probabilities of adopting each behavior are computed.

Our focus is on the evolution of learning strategies in a population. Each learning strategy is assumed to be genetically determined and – in the models that we consider in this review – not modifiable by learning. The fitness of a learning strategy in a given environment depends on whether the behavior(s) it dictates is (are) adaptive or maladaptive in that environment. The environment may change in time or vary spatially, and a behavior that

\* Corresponding author.

E-mail address: [kenaoki@meiji.ac.jp](mailto:kenaoki@meiji.ac.jp) (K. Aoki).

may have been the best, in terms of natural selection, in one environment may not be the best in another. The fitness of a learning strategy also depends in a frequency-dependent manner on what the competing strategies are doing. Earlier studies (e.g., Boyd and Richerson, 1985; Rogers, 1988; Feldman et al., 1996) emphasized SL, as this form of learning is essential for culture. More recently, learning strategies combining IL and SL that support cumulative culture are receiving attention (e.g., Enquist et al., 2007; Borenstein et al., 2008; Aoki, 2010; Lehmann et al., 2010; Aoki et al., 2012).

In ecology, the evolution of learning has been widely studied in the context of foraging (e.g., Barnard and Sibly, 1981; Stephens, 1991; Rodriguez-Gironés and Vásquez, 1997; Giraldeau and Caraco, 2000; Eliassen et al., 2009; Dubois et al., 2010; Katsnelson et al., 2011; Arbilly et al., 2011). The models in this area often address complex situations and posit specific targets of learning, such as where to forage or whether to produce or to scrounge. As such, these models are usually not amenable to a formal mathematical treatment. The evolutionary models of learning that we consider in this review are more “abstract”, in the sense that the behavioral alternatives are distinguished only by whether they are adaptive or maladaptive, or by the degree of adaptedness. Some models are phrased in terms of the number of adaptive cultural traits carried by an organism (Lehmann and Feldman, 2009; Nakahashi, 2010), but they will not be addressed in this review. In the simplest situations, we can write down the dynamical equations describing the changes in the frequencies of the competing learning strategies in terms of their variable fitnesses in the different environments to which they are exposed. More complicated situations involving strategies that differ in the probabilities of using IL or SL can sometimes be modeled by the evolutionarily stable strategy (ESS) approach (Maynard Smith, 1982).

The models reviewed in detail in this paper are numbered sequentially from 1 to 11. We seek the stable equilibria of the dynamical equations or alternatively the ES learning strategy. In addition, we briefly discuss several interesting but complex models, some of which have been investigated using Monte Carlo/agent-based simulations. It will be seen that the results obtained from the simpler models can usefully be applied to interpreting the observations on the more complex models. Finally, we ask whether the presence of SL will improve the (geometric) mean fitness of a population relative to when it is absent—i.e., we address the so-called Rogers’ paradox (Rogers, 1988; Boyd and Richerson, 1995). Table 1 summarizes the provenance of models 1–11 and indicates for each model whether or not Rogers’ paradox occurs.

## 2. Dynamical models in temporally variable environments

The basic models of this section assume the simplest learning strategies, namely those that involve the use of IL or SL but not both. They also assume dichotomous variation in the phenotype (behavior) that can be acquired by learning. It is then possible to write down the difference equations governing the frequency dynamics of the learning strategies and phenotypes, which is done here for three of the four models.

### 2.1. Model 1: infinite-states $l$ -cycle coevolutionary model

This model, which was first described by Feldman et al. (1996) in a slightly less general form, is coevolutionary in the sense that the learning strategies and behaviors can coevolve. Consider an infinite population of haploid organisms in which a genetic locus with two alleles determines whether an organism is an obligate individual learner or an obligate social learner. Among the adults of each generation, we distinguish two behaviors, correct or wrong, which are adaptive or maladaptive, respectively, in the environment faced by that generation. Behaviors are defined

**Table 1**  
Provenance of Models 1–11 and possibility of Rogers’ paradox.

| Model | Provenance                      | Comments <sup>a</sup>    | Rogers’ paradox    |
|-------|---------------------------------|--------------------------|--------------------|
| 1     | Feldman et al. (1996)           | Generalization           | Always observed    |
| 2     | Feldman et al. (1996)           | Parameter range extended | Sometimes resolved |
| 3     | Rogers (1988)                   | Modified formulation     | Always observed    |
| 4     | Kendal et al. (2009)            | Simplification           | Sometimes resolved |
| 5     | Feldman et al. (1996)           | Detailed analysis        | Resolved           |
| 6     | Boyd and Richerson (1988, 1995) | Modified formulation     | Not considered     |
| 7     | Aoki and Nakahashi (2008)       | Unmodified               | Sometimes resolved |
| 8     | Enquist et al. (2007)           | Reworded                 | Sometimes resolved |
| 9     | Aoki et al. (2012)              | Unmodified               | Resolved           |
| 10    | Nakahashi et al. (2012)         | Unmodified               | Not addressed      |
| 11    | Wakano and Aoki (2006)          | Unmodified               | Not addressed      |

<sup>a</sup> Comments refer to the present analysis and discussion of the models in the corresponding references.

relative to the environment, so that when the environment changes, so do the behaviors that are correct or wrong. These adults reproduce asexually without fertility differences.

A newborn individual learner gathers information directly from the environment and achieves the correct behavior on its own before becoming an adult. However, it suffers a cost,  $c$ , which can be interpreted as the probability of making a fatal mistake. Hence, a fraction  $1 - c$  of individual learners survive to adulthood, and they all show the correct behavior.

A newborn social learner, on the other hand, acquires its behavior by faithfully copying (i.e., imitating) a random member of the parental generation. Its behavior will be correct only if the behavior that it copies from its exemplar (i.e., cultural parent) is correct in the environment into which it is born. We assume that the environment changes every  $l$  generations, with that change occurring just prior to birth. Moreover, an environmental change results in a previously unknown state, which entails that neither of the two preexisting behaviors (correct or wrong) can be correct after the environmental change. Hence, only the individual learners can acquire the correct behavior immediately after an environmental change; this is known as the infinite environmental states assumption. A social learner with correct behavior has fitness (relative viability) 1, whereas the fitness associated with wrong behavior is  $1 - s$ . We assume  $0 < c < s < 1$ ; otherwise, the individual learners will be selected out unconditionally.

Hence, among the surviving adults of any generation, there can be three phenogenotypes (i.e., genotype–phenotype combinations, Feldman and Cavalli-Sforza, 1984): individual learner, social learner with correct behavior (SLC), and social learner with wrong behavior (SLW). Let us denote their respective frequencies in the parental generation by  $z$ ,  $x$ , and  $y$ , and the corresponding frequencies in the offspring generation by  $z'$ ,  $x'$ , and  $y'$ . Then, the difference equations governing the dynamics of these variables can be written as follows. When there is an intervening environmental change, which occurs once every  $l$  generations, we have

$$Vx' = 0, \quad (1.1a)$$

$$Vy' = (1 - s)(x + y), \quad (1.1b)$$

$$Vz' = (1 - c)z, \quad (1.1c)$$

where

$$V = (1 - c)z + (1 - s)(1 - z). \quad (1.1d)$$

On the other hand, when the environment does not change between the generations,

$$Wx' = (x + y)(x + z), \quad (1.2a)$$

$$Wy' = (1 - s)(x + y)y, \quad (1.2b)$$

$$Wz' = (1 - c)z, \quad (1.2c)$$

where

$$W = (1 - c)z + (1 - s)y(1 - z). \quad (1.2d)$$

Eqs. (1.2) are to be applied consecutively  $l - 1$  times, since there are  $l - 1$  generations of environmental stasis after a change. The variables  $V$  and  $W$  in Eqs. (1.1d) and (1.2d) give the population mean fitnesses when the environment changes and when it does not change, respectively, and serve to normalize the equations.

To illustrate how these recursions follow from the assumptions, we briefly explain the derivation of Eqs. (1.1a), (1.1b), (1.2a) and (1.2b). Note first that SLC and SLW have the same genotype although their phenotypes differ, so that the frequency of social learners among the newborns of the offspring generation will be the sum of their frequencies,  $x + y$ . When the environment changes, neither the correct behavior nor the wrong behavior in the parental generation is correct as viewed from – i.e., in the environment faced by – the offspring generation. Hence, all newborn social learners will be SLW, and since their viability is  $1 - s$ , we obtain Eq. (1.1b). In this case the frequency of SLC will of course be 0, resulting in Eq. (1.1a). With environmental stasis, on the other hand, fractions  $x + z$  and  $y$  of the parental generation – recall  $x$ ,  $y$ , and  $z$  are the frequencies of SLC, SLW, and individual learners, respectively – have correct and wrong behaviors, respectively, as viewed from the offspring generation. Hence, when a newborn social learner copies a random member of the parental generation, it acquires the correct behavior with probability  $x + z$  and the wrong behavior with complementary probability  $y$ . On including the effects of viability selection, we obtain Eqs. (1.2a) and (1.2b).

There are two monomorphic equilibria of the model, which always (i.e., for all valid parameter combinations) exist. At the first such equilibrium, the individual learners are fixed, and we have  $\hat{z}^{(i)} = 1$  for  $0 \leq i \leq l - 1$ , where the hat indicates equilibrium, and the generations are counted with  $i = 0$  indicating the generation immediately before an environmental change. This equilibrium is (locally) stable if

$$1 - s < (1 - c)^l. \quad (1.3)$$

At the second monomorphic equilibrium, SLW are fixed, such that  $\hat{y}^{(i)} = 1$  for  $0 \leq i \leq l - 1$ . This equilibrium is always unstable. In addition, we conjecture the existence of a periodic fully polymorphic (all three phenotypes are present) equilibrium, which is stable when inequality (1.3) is reversed. Clearly, inequality (1.3) is more likely to be satisfied when  $l$  is small, where  $l$  is the period of environmental change. Since a shorter period corresponds to greater environmental instability, the model predicts that individual learners are more likely to be fixed in a more changeable environment. See Appendix A for details.

## 2.2. Rogers' paradox

Rogers (1988) reasoned that the introduction of social learners into a population of individual learners would improve the mean fitness of that population, because social learners make culture possible, and culture is presumably an adaptation the capacity for which evolved by natural selection. Contrary to expectation, his simple model (see Model 3) suggested that this was not true. Rogers' paradox, as it has come to be called, is said to occur when the parameter values are such that (a) a stable equilibrium exists with social learning (learners) present, (b) a monomorphic

equilibrium of individual learning (learners) also exists, and (c) the two equilibria have the same mean fitness (Boyd and Richerson, 1995; Enquist et al., 2007). In such a case, culture does not appear to impart a fitness advantage: hence the paradox.

For Model 1 described above, numerical work suggests that a periodic fully polymorphic equilibrium exists and is stable when inequality (1.3) is reversed. Because  $\hat{z}^{(i)} > 0$  for  $0 \leq i \leq l - 1$  at this equilibrium, Eqs. (1.1c) and (1.2c) entail that the equilibrium mean fitness in any generation is equal to  $1 - c$ , and hence that the geometric mean of the mean fitnesses over the  $l$ -cycle is  $1 - c$ . The same argument shows that the geometric mean when individual learners are fixed is also  $1 - c$ . Thus, Rogers' paradox is observed in this model.

## 2.3. Model 2: two-state $l$ -cycle coevolutionary model

We modify Model 1 by assuming that the environment fluctuates between two states, e.g., hot and cold, again with period  $l$ . There are two behaviors each of which is correct in one environmental state and wrong in the other. Hence, when there is an intervening environmental change, the wrong behavior in the parental generation becomes the correct behavior as viewed from the offspring generation. For this model, the difference equations (1.2) still apply, but we must replace Eq. (1.1) by

$$Vx' = (x + y)y, \quad (2.1a)$$

$$Vy' = (1 - s)(x + y)(x + z), \quad (2.1b)$$

$$Vz' = (1 - c)z, \quad (2.1c)$$

where

$$V = (1 - c)z + [1 - s(1 - y)](1 - z). \quad (2.1d)$$

Note that  $x + z$  and  $y$  in Eqs. (2.1a) and (2.1b) are interchanged, reflecting the reversal of correct and wrong after an environmental change.

In relation to Rogers' paradox, an interesting aspect of this model is that a population that is monomorphic for social learners – but in which SLC and SLW coexist – cannot be invaded by individual learners if

$$(1 - c)^2 < 1 - s. \quad (2.2)$$

There are an infinite number of periodic equilibria where social learners are fixed, which are neutrally stable with regard to perturbations in the frequencies of SLC and SLW. Nevertheless, each such equilibrium is stable to invasion by individual learners provided inequality (2.2) is satisfied. Moreover, the geometric mean of the mean fitnesses is  $\sqrt{1 - s}$ , whereas it is  $1 - c$  when individual learners are fixed. Hence, a population that is monomorphic for social learners can be stable and have a higher mean fitness than one that is monomorphic for individual learners—Rogers' paradox is resolved. See Appendix B for details. (In Feldman et al. (1996), we stated that the fixation of social learners was always unstable. This result followed from the parameterization used in that paper which entailed that  $c < s/2$ .)

## 2.4. Model 3: Rogers' two-timescale model

The two-timescale model assumes that the genetic evolution is much slower than the cultural evolution and hence that genetic variables can be regarded as constants of cultural evolution (Rogers, 1988; Enquist et al., 2007). Although this is not a dynamical model for phenotypes, we include it here because it is closely related to the other models in this section. We explain Rogers' result using an argument that differs slightly from the original. There are two environmental states as in Model 2 described above. The possible phenotypes and their fitnesses are as in

Models 1 and 2, which is a more general parameterization than Rogers (1988).

Let  $p$  be the genetically-determined frequency of social learners among adults. Consider a focal newborn social learner and a social transmission chain extending backward in time with random copying from the previous generation. This chain may comprise social learners of ascending generations, but will eventually end in an individual learner. Since  $p$  is assumed constant during cultural evolution, the probability that an individual learner occurs in this chain for the first time exactly  $t$  generations ago is

$$p^{t-1}(1-p). \quad (3.1)$$

Next, let  $\varepsilon$  be the probability of an environmental change between generations. Since the environment fluctuates between two states, fitness is clearly dependent on whether the number of environmental changes is even or odd. Appendix C shows that the probability of no environmental change or an even number of changes in  $t$  generations is

$$\frac{1}{2} [1 + (1 - 2\varepsilon)^t]. \quad (3.2)$$

In this case, the (correct) behavior acquired  $t$  generations ago by an individual learner is correct in the current generation. If, on the other hand, there are an odd number of environmental changes, which occurs with complementary probability  $\frac{1}{2} [1 - (1 - 2\varepsilon)^t]$ , this behavior is now wrong, and the fitness will be  $1 - s$  rather than 1. Hence, the expected fitness of a social learner is

$$\tilde{w}_s^{(2)} = \frac{1}{2} \sum_{t=1}^{\infty} p^{t-1} (1-p) \{ [1 + (1 - 2\varepsilon)^t] + [1 - (1 - 2\varepsilon)^t] (1-s) \}, \quad (3.3)$$

which reduces to

$$\tilde{w}_s^{(2)} = 1 - \frac{s\varepsilon}{1 - p(1 - 2\varepsilon)}. \quad (3.4)$$

The superscript (2) indicates that the fitness of the social learner has been calculated on the two environmental states model. Rogers (1988) sets the fitnesses of SLC and SLW to  $w + b$  and  $w - b$ , respectively. It can be shown with the appropriate substitutions ( $w \rightarrow 1 - s/2$ ,  $b \rightarrow s/2$ ,  $u \rightarrow 2\varepsilon$ ) that Eq. (3.4) is equivalent to the unnumbered equation in Rogers (1988) immediately above Eq. (2) of that paper.

The fitness of an individual learner is  $1 - c$ . Thus, at the genetically polymorphic equilibrium

$$1 - c = 1 - \frac{s\varepsilon}{1 - p(1 - 2\varepsilon)}. \quad (3.5)$$

Solving Eq. (3.5) yields

$$\hat{p} = \frac{c - s\varepsilon}{c(1 - 2\varepsilon)}, \quad (3.6)$$

which is valid for  $s\varepsilon < c$ . If  $s\varepsilon > c$ , on the other hand, the only valid equilibrium is one that is monomorphic for individual learners.

Let us now briefly redo the calculations for an infinite environmental states version of Rogers' model. The probability of no environmental change in  $t$  generations is  $(1 - \varepsilon)^t$ , and the probability of at least one change is  $1 - (1 - \varepsilon)^t$ . Hence, the expected fitness of the newborn social learner in this modified Rogers' model is

$$\tilde{w}_s^{(\infty)} = \sum_{t=1}^{\infty} p^{t-1} (1-p) \{ (1 - \varepsilon)^t + [1 - (1 - \varepsilon)^t] (1-s) \}, \quad (3.7)$$

which reduces to

$$\tilde{w}_s^{(\infty)} = 1 - \frac{s\varepsilon}{1 - p(1 - \varepsilon)}. \quad (3.8)$$

The superscript ( $\infty$ ) distinguishes the infinite from the two environmental states model. Hence, the frequency of social learners at the genetically polymorphic equilibrium for this case is

$$\hat{p} = \frac{c - s\varepsilon}{c(1 - \varepsilon)}. \quad (3.9)$$

We have two comments on the analysis by Rogers (1988). First, the so-called Rogers' paradox is inherent in his derivation, since the genetically polymorphic equilibrium,  $\hat{p}$ , is obtained by equating the fitnesses of the individual learners and the social learners (see Eq. (3.5)). Second, it fails to identify the stable monomorphism of social learners that is predicted by the dynamical model (Model 2).

## 2.5. Model 4: information decay model

Assume, as in the modified (infinite environmental states) Rogers' model, that the correct behavior becomes outdated and hence wrong with probability  $\varepsilon$  per generation. Parameter  $\varepsilon$  would appear to have the same meaning as in the modified Rogers' model. Then, with the possible phenotypes and their fitnesses as in Model 1, the difference equations are

$$Wx' = (x + y)(x + z)(1 - \varepsilon), \quad (4.1a)$$

$$Wy' = (1 - s)(x + y) [(x + z)\varepsilon + y], \quad (4.1b)$$

$$Wz' = (1 - c)z, \quad (4.1c)$$

where

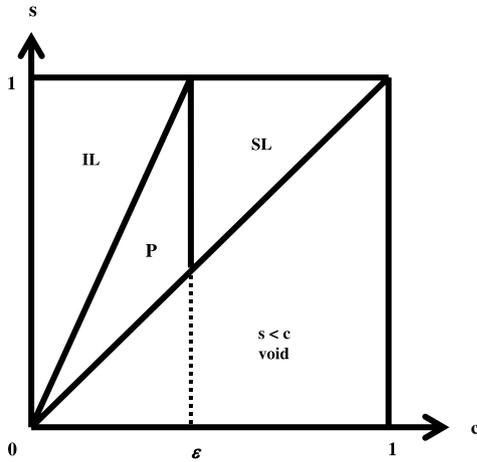
$$W = (1 - c)z + \{1 - s[\varepsilon + y(1 - \varepsilon)]\} (1 - z). \quad (4.1d)$$

In Eqs. (4.1a) and (4.1b), the term  $(x + z)\varepsilon$  represents the decay of adaptive information. With this model, we do not require a separate set of equations for environmental change and stasis. This modeling approach has been used by Kendal et al. (2009), Lehmann and Feldman (2009), Lehmann et al. (2010), and Nakahashi (2010). In fact, the inclusion of Model 4 in the current paper was motivated by the model of Kendal et al. (2009) for the case of unbiased social learning.

The equilibria of Eqs. (4.1) and their stability properties are as follows.

- (A)  $\hat{x} = 0$ ,  $\hat{y} = 0$ ,  $\hat{z} = 1$ ,  $\hat{W} = 1 - c$ ; this is monomorphic for individual learners, always (i.e., for all parameter combinations) exists, and is stable if  $c < s\varepsilon$ .
- (B)  $\hat{x} = 0$ ,  $\hat{y} = 1$ ,  $\hat{z} = 0$ ,  $\hat{W} = 1 - s$ ; this always exists but is never stable.
- (C)  $\hat{x} = \frac{s - \varepsilon}{s(1 - \varepsilon)}$ ,  $\hat{y} = \frac{\varepsilon(1 - s)}{s(1 - \varepsilon)}$ ,  $\hat{z} = 0$ ,  $\hat{W} = 1 - \varepsilon$ ; this is genetically monomorphic for social learners, exists if  $\varepsilon < s$ , and is stable if  $\varepsilon < c$ .
- (D)  $\hat{x} = \frac{(c - s\varepsilon)(s - c)}{s(1 - \varepsilon)c(1 - s)}$ ,  $\hat{y} = \frac{c - s\varepsilon}{s(1 - \varepsilon)}$ ,  $\hat{z} = \frac{(s - c)(\varepsilon - c)}{(1 - \varepsilon)c(1 - s)}$ ,  $\hat{W} = 1 - c$ ; this is a fully polymorphic equilibrium, exists if  $s\varepsilon < c < \varepsilon$ , and is stable whenever it exists. See Appendix D for details.

Fig. 1 illustrates the mutually exclusive regions in the parameter space of  $c$  and  $s$  where equilibria (A), (C), and (D) exist and are stable. Equilibrium (A), which is monomorphic for individual learners, coexists with the other two equilibria throughout the upper triangular region of Fig. 1, and is always associated with a mean fitness of  $1 - c$ . Social learners are present at equilibria (C) and (D). The mean fitness at equilibrium (D) is  $1 - c$ , which is identical to that at equilibrium (A), and these two equilibria can coexist for the same parameter values (i.e.,  $s\varepsilon < c < \varepsilon$ ). Hence, Rogers' paradox occurs in this case.



**Fig. 1.** Regions of local stability of the equilibria of Model 4 (information decay model) in the  $(c, s)$ -parameter space. In the triangular regions labeled IL, SL, and P, the parameter values are such that fixation of individual learners, fixation of social learners, and polymorphism of both are, respectively, the unique locally stable equilibrium. The lower triangular region is void by assumption. Parameter  $\varepsilon$  represents the rate of decay of adaptive information per generation.

When  $c > \varepsilon$  on the other hand, the equilibrium (C) is stable, and the mean fitness at this equilibrium,  $1 - \varepsilon$ , is greater than at the coexisting unstable monomorphism of individual learners. Hence, Rogers' paradox is resolved in this case. Interestingly, an equilibrium corresponding to (C) does not exist in the model of Kendal et al. (2009), which entails that Rogers' paradox is not eliminated (in the case of unbiased social learning). But when notational idiosyncrasies are accounted for, their model is identical to Model 4 considered here, except that selection acts through fertility rather than viability differences. Clearly, whether or not Rogers' paradox occurs is determined by subtle differences in the assumptions.

Equilibrium (D) in this model would seem to correspond to the genetically polymorphic equilibrium in the modified Rogers' model given by Eq. (3.9). Hence, one might expect  $\hat{x} + \hat{y} = \hat{p}$  to hold, but this is not the case. In fact,  $\hat{x} + \hat{y} = \hat{p} \frac{1-c}{1-s}$ .

### 3. ESS models in temporally variable environments

#### 3.1. Model 5: mixed strategy model with infinite-states $l$ -cycle

The four models described above share the perhaps "unrealistic" assumption that an organism can engage in individual learning (IL) or social learning (SL) but not both. Here, we consider a model due to Feldman et al. (1996) in which IL and SL are both available to the same organism. The analysis requires us to posit two genetically determined mixed strategies, the resident and the mutant, that differ in the probability of using IL, which we denote by  $L$  for the resident and  $L + dL$  for the mutant. As before, a behavior is either correct or wrong. Hence, we need to distinguish four phenogenotypes: resident with correct behavior, resident with wrong behavior, mutant with correct behavior, and mutant with wrong behavior. Their frequencies among asexually reproducing adults are  $x, \bar{x}, y,$  and  $\bar{y}$ , respectively. Note that, these variables are defined differently from those in the dynamical models considered in the previous section. An environmental change results in a state that has not been experienced before (infinite states).

When the environment changes, the correct behavior is achieved only by IL. Hence, the difference equations are

$$Vx' = (x + \bar{x})L(1 - c), \tag{5.1a}$$

$$V\bar{x}' = (x + \bar{x})(1 - L)(1 - s), \tag{5.1b}$$

$$Vy' = (y + \bar{y})(L + dL)(1 - c), \tag{5.1c}$$

$$V\bar{y}' = (y + \bar{y})(1 - L - dL)(1 - s), \tag{5.1d}$$

with

$$V = (x + \bar{x}) [L(1 - c) + (1 - L)(1 - s)] + (y + \bar{y}) [(L + dL)(1 - c) + (1 - L - dL)(1 - s)]. \tag{5.1e}$$

When there is no environmental change between the generations, the correct behavior can be achieved either by IL, or by SL from an appropriate exemplar. Hence,

$$Wx' = (x + \bar{x}) [L(1 - c) + (1 - L)(x + y)], \tag{5.2a}$$

$$W\bar{x}' = (x + \bar{x})(1 - L)(1 - s)(\bar{x} + \bar{y}), \tag{5.2b}$$

$$Wy' = (y + \bar{y}) [(L + dL)(1 - c) + (1 - L - dL)(x + y)], \tag{5.2c}$$

$$W\bar{y}' = (y + \bar{y})(1 - L - dL)(1 - s)(\bar{x} + \bar{y}), \tag{5.2d}$$

with

$$W = (x + \bar{x}) \{L(1 - c) + (1 - L)[1 - s(\bar{x} + \bar{y})]\} + (y + \bar{y}) \{(L + dL)(1 - c) + (1 - L - dL)[1 - s(\bar{x} + \bar{y})]\}. \tag{5.2e}$$

As before, we assume  $0 < c < s < 1$ .

We show in Appendix E that the genetically monomorphic periodic equilibrium of the resident can be written as

$$\hat{x}^{(i)} = \frac{1 - \alpha}{1 - \alpha + \beta(1 - \alpha^i)}, \quad \hat{x}^{(i)} = 1 - \hat{x}^{(i)} \tag{5.3a}$$

for  $1 \leq i \leq l$ , where

$$\alpha = \frac{L(1 - c) + 1 - L}{(1 - L)(1 - s)}, \quad \beta = \frac{L(1 - c)}{(1 - L)(1 - s)}. \tag{5.3b}$$

This equilibrium has period  $l$ , since  $\hat{x}^{(1)} = \frac{1}{1+\beta}$  every  $l$  generations.

Hence,  $\hat{x}^{(0)} = \hat{x}^{(l)}$ . In the special case of  $L = 1$ , we clearly have  $\hat{x}^{(i)} = 0$  for  $1 \leq i \leq l$  from Eqs. (5.1b) and (5.2b).

Furthermore, the eigenvalue governing the invasion of the mutant is

$$\lambda = A \prod_{i=1}^{l-1} B^{(i)}, \tag{5.4a}$$

where

$$A = 1 + dL \frac{s - c}{L(1 - c) + (1 - L)(1 - s)}, \tag{5.4b}$$

$$B^{(i)} = 1 + dL \frac{\hat{s}\bar{x}^{(i)} - c}{L(1 - c) + (1 - L)(1 - \hat{s}\bar{x}^{(i)})}.$$

It can readily be shown that the  $L = 0$  (pure SL) strategy cannot be an ESS, and that a necessary condition for the  $L = 1$  (pure IL) strategy to be an ESS is  $l < s/c$ . The latter condition entails that shorter periods of environmental change are necessary for the evolution of IL, which is consistent with our findings based on the previous models. Furthermore, when  $dL$  is small, a necessary condition for an internal ESS is given by

$$f(L) = \frac{s - c}{L(1 - c) + (1 - L)(1 - s)} + \sum_{i=1}^{l-1} \frac{\hat{s}\bar{x}^{(i)} - c}{L(1 - c) + (1 - L)(1 - \hat{s}\bar{x}^{(i)})} = 0. \tag{5.5}$$

The solution(s) for  $L$  can be found numerically for given values of  $s, c,$  and  $l$  on substituting Eq. (5.3a) with (5.3b) into (5.5).

Let us now examine the special case of environmental period 2 ( $l = 2$ ) in some detail. We wish to determine the conditions for an interior ESS – which is analogous to a genetically polymorphic equilibrium in the dynamical models – to exist and whether Rogers' paradox is observed. Eqs. (5.3) and (5.4) reduce to

$$\hat{x}^{(0)} = \frac{1}{1 + \beta(1 + \alpha)}, \quad \hat{x}^{(1)} = \frac{1}{1 + \beta}, \quad (5.6)$$

with  $\alpha$  and  $\beta$  given by Eq. (5.3b), and

$$\begin{aligned} \lambda = 1 + dL & \left[ \frac{s - c}{L(1 - c) + (1 - L)(1 - s)} \right. \\ & \left. + \frac{\hat{s}\hat{x}^{(1)} - c}{L(1 - c) + (1 - L)(1 - \hat{s}\hat{x}^{(1)})} \right] \\ & + (dL)^2 \left[ \frac{s - c}{L(1 - c) + (1 - L)(1 - s)} \right] \\ & \times \left[ \frac{\hat{s}\hat{x}^{(1)} - c}{L(1 - c) + (1 - L)(1 - \hat{s}\hat{x}^{(1)})} \right]. \quad (5.7) \end{aligned}$$

The coefficient of  $dL$  in Eq. (5.7) is clearly  $f(L)$  as defined in Eq. (5.5) for the special case of  $l = 2$ . Using Eqs. (5.3) and (5.6), we find that  $f(0) > 0$  since  $c < s$  by assumption, and that  $f(1) < 0$  provided  $s/c < 2$ . We assume  $s/c < 2$  in what follows. There is an (at least one) interior root of  $f(L) = 0$  in this case, which we denote by  $L^*$ . Of the two terms in  $f(L^*)$ , the first is positive, so the second term must be negative. Hence, the coefficient of  $(dL)^2$  is also negative, ensuring that  $L^*$  is an ESS.

Next, the product of the mean fitnesses in generations 0 and 1 is given by the quadratic

$$g(L) = \hat{W}^{(0)}\hat{W}^{(1)} = [L(1 - c) + (1 - L)(1 - s)]^2 + L(1 - L)s(1 - c). \quad (5.8)$$

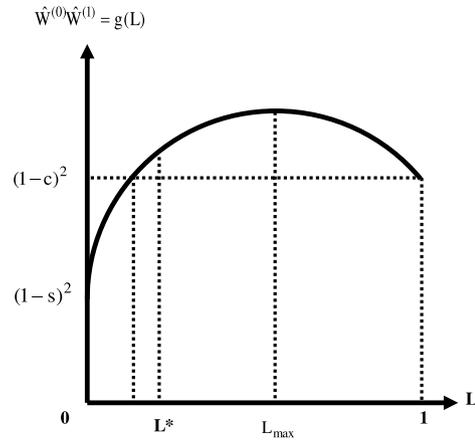
Clearly,  $g(0) = (1 - s)^2$ ,  $g(1) = (1 - c)^2$ , and  $g(L)$  has a maximum between 0 and 1, as shown schematically in Fig. 2. If Rogers' paradox occurs, then we should have  $g(L^*) = (1 - c)^2$ . Unfortunately, although  $f(L)$  and  $g(L)$  are both simple functions of  $L$ , we are unable to show analytically whether this is true or false. On the other hand, we find on numerically solving  $f(L) = 0$  for  $L^*$  that  $g(L^*) > (1 - c)^2$ . Thus, the product (or alternatively the geometric mean) of the mean fitnesses is apparently greater for the interior ESS than in a population using the  $L = 1$  (pure IL) strategy, but is not maximized.

### 3.2. Model 6: sampling the environment

We briefly describe a model due to Boyd and Richerson (1988, 1995), which we include here because it has a mathematical structure that is quite similar to Model 5 (Wakano and Aoki, 2007). The original formulation and analysis by Boyd and Richerson (1988, 1995) were based on a combination of the two-timescale (see Model 3) and the information decay (see Model 4) approaches.

In this model, the environment fluctuates between two states. Using a parameter  $\mu$ , the two states are denoted  $\mu > 0$  and  $-\mu < 0$ . We assume that the temporal fluctuations occur periodically, rather than with a given probability as in the model of Boyd and Richerson (1988, 1995). An organism first samples the environment by IL, and then resorts to oblique SL when the information obtained is indecisive. More specifically, suppose that the environment is in state  $\mu > 0$ . Information,  $z$ , obtained on sampling the environment is distributed normally as

$$\phi(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(z - \mu)^2}{2\sigma^2}\right\}. \quad (6.1)$$



**Fig. 2.** Product of the mean fitnesses in two successive generations ( $g(L)$ ) plotted against the probability of using individual learning ( $L$ ) for the special case of Model 5 (mixed strategy model) with environmental period 2 ( $l = 2$ ).  $L^*$  denotes the evolutionarily stable probability of individual learning. Note that  $g(L^*) > (1 - c)^2$ , whereas equality would hold if Rogers' paradox were to apply.

Large positive values of  $z$  correctly suggest that the environment is in state  $\mu > 0$ , while moderately large negative values incorrectly suggest that the environment is in state  $-\mu < 0$ . The resident strategy achieves the correct behavior by IL with probability

$$\pi_c = \int_d^\infty \phi(z) dz \quad (6.2a)$$

when  $z$  exceeds a genetically determined threshold,  $d$ . Similarly, it acquires the wrong behavior by IL with probability

$$\pi_w = \int_{-\infty}^{-d} \phi(z) dz \quad (6.2b)$$

when  $z$  is small, and it resorts to SL with probability

$$K = \int_{-d}^d \phi(z) dz \quad (6.2c)$$

when  $z$  takes intermediate values. Hence  $K + \pi_c + \pi_w = 1$ , and due to the way in which the three parameters are defined in Eq. (6.2), only one is independent.

Analogous equations apply when the environment is in state  $-\mu < 0$ . For the mutant strategy, we substitute  $K + dK$  for  $K$ ,  $\pi_c + d\pi_c$  for  $\pi_c$ , and  $\pi_w + d\pi_w$  for  $\pi_w$ , where  $dK + d\pi_c + d\pi_w = 0$ .

As before, we denote the frequencies among adults of the residents with correct behavior, the residents with wrong behavior, the mutants with correct behavior, and the mutants with wrong behavior by  $x$ ,  $\bar{x}$ ,  $y$ , and  $\bar{y}$ , respectively. Then, when the environment changes between generations,

$$Vx' = (x + \bar{x}) [K(\bar{x} + \bar{y}) + \pi_c], \quad (6.3a)$$

$$V\bar{x}' = (x + \bar{x}) [K(x + y) + \pi_w] (1 - s), \quad (6.3b)$$

$$Vy' = (y + \bar{y}) [(K + dK)(\bar{x} + \bar{y}) + \pi_c + d\pi_c], \quad (6.3c)$$

$$V\bar{y}' = (y + \bar{y}) [(K + dK)(x + y) + \pi_w + d\pi_w] (1 - s), \quad (6.3d)$$

and when the environment remains constant between generations

$$Wx' = (x + \bar{x}) [K(x + y) + \pi_c], \quad (6.4a)$$

$$W\bar{x}' = (x + \bar{x}) [K(\bar{x} + \bar{y}) + \pi_w] (1 - s), \quad (6.4b)$$

$$Wy' = (y + \bar{y}) [(K + dK)(x + y) + \pi_c + d\pi_c], \quad (6.4c)$$

$$W\bar{y}' = (y + \bar{y}) [(K + dK)(\bar{x} + \bar{y}) + \pi_w + d\pi_w] (1 - s). \quad (6.4d)$$

Let the period of environmental fluctuations be  $l$ . Assuming that  $dK$ ,  $d\pi_c$ , and  $d\pi_w$  are small, Wakano and Aoki (2007) showed that

the  $K = 0$  (i.e.,  $d = 0$ ; pure IL) strategy is unstable if  $l > 2$  and the  $K = 1$  (i.e.,  $d \rightarrow \infty$ ; pure SL) strategy is always unstable. Hence, at least one interior ESS,  $0 < K^* < 1$  ( $0 < d^* < \infty$ ), is predicted for  $l > 2$ , and can be identified numerically.

A comparison of Eqs. (5.1) and (5.2) with Eqs. (6.3) and (6.4) shows that Model 6 indeed has a similar mathematical structure to Model 5, although there are two significant differences. First, Model 6 assumes two environmental states as does Model 2, so that the wrong behavior in the parental generation becomes the correct behavior in the offspring generation after an environmental change. Second, IL can result in the wrong behavior, but in Model 6 does not entail an exogenous cost measured by  $c$  in Model 5. The model of Boyd and Richerson (1988, 1995) on which Model 6 is based has often been interpreted as prescribing a learning schedule in which IL is followed by SL. However, the similarities with Model 5 indicate that the sequential use of IL and SL is not a necessary aspect of this learning strategy.

Boyd and Richerson (1995) find that the interior ESS is associated with a greater mean fitness than the  $K = 0$  (pure IL) strategy (i.e., Rogers' paradox is resolved), but does not maximize the mean fitness, which again is qualitatively similar to the result for Model 5.

#### 4. Population structure, spatially variable environment, and migration

Spatial environmental heterogeneity such that migrating organisms will experience novel environments has been addressed by Boyd and Richerson (1985, 1988, 1995), Henrich and Boyd (1998), Rendell et al. (2010), Aoki (2010), Nakahashi et al. (2012) among others. Here, we briefly review one dynamical model due to Aoki and Nakahashi (2008), which permits a fairly complete analysis.

##### 4.1. Model 7: finite island model with environmentally heterogeneous sites

Organisms are of two genetically distinct types, individual learners and social learners. They may occupy any of  $n$  equally-connected sites in a spatially heterogeneous world. Each site has a different environment. We distinguish  $n$  behaviors, each of which is locally adapted to one particular environment but maladaptive in the  $n - 1$  other environments. Behaviors that are maladaptive in all  $n$  environments are not incorporated into the dynamics. Let  $X_{ij}$  ( $1 \leq i, j \leq n$ ) be the number of social learners at site  $i$  that are adapted to the environment of site  $j$ . Then, at site  $i$  there are  $X_i = \sum_{j=1}^n X_{ij}$  social learners, of which  $X_{ii}$  are behaving adaptively (SLC), and  $X_i - X_{ii}$  are behaving maladaptively (SLW) at fitness cost  $s$ . Similarly, let  $Z_i$  ( $1 \leq i \leq n$ ) be the number of individual learners at site  $i$ . Individual learners always acquire the correct behavior, but suffer an exogenous cost  $c$ . Write  $N_i = X_i + Z_i$  for the total population at site  $i$ . These numbers are enumerated at the adult stage just prior to reproduction. Birth is followed by social learning, migration, individual learning, and then viability selection, in that order.

The life cycle begins with reproduction, where each organism gives birth asexually to  $b(N_i)$  offspring according to the discrete logistic

$$b(N_i) = 1 + r(1 - N_i/K). \tag{7.1}$$

Here, the intrinsic growth rate,  $r > 0$ , and the carrying capacity,  $K > 0$ , are assumed to be the same for all sites. Second, social learners acquire their behavior by copying a random member of the parental generation at their site. As a result, the number of social learners at site  $i$  that are adapted to the environment of site

$j$  becomes

$$X_i b(N_i) \frac{X_{ij} + Z_i \delta_{ij}}{N_i}, \tag{7.2}$$

where  $\delta_{ij} = 1$  for  $i = j$  and 0 otherwise. All members of the parental generation die immediately afterward. The third event is migration, where a fixed fraction of the organisms at each site emigrate (constant forward migration rate). We assume reciprocal migration between all pairs of sites at rate  $m/(n - 1)$  with  $0 < m \leq 1/2$ . Fourth, the individual learners acquire the behavior suitable to their environment. And fifth, viability selection occurs such that all social learners behaving adaptively (SLC) survive, a fraction  $1 - s$  of social learners behaving maladaptively (SLW) survive, and a fraction  $1 - c$  of individual learners survive, where  $0 < c < s < 1$ .

The above assumptions entail that the difference equations for this island model be written as

$$X'_{ii} = (1 - m) X_i b(N_i) \frac{X_{ii} + Z_i}{N_i} + \frac{m}{n - 1} \sum_{k \neq i} X_k b(N_k) \frac{X_{ki}}{N_k}, \tag{7.3a}$$

$$X'_{ij} = (1 - s) \left\{ (1 - m) X_i b(N_i) \frac{X_{ij}}{N_i} + \frac{m}{n - 1} X_j b(N_j) \frac{X_{ij} + Z_j}{N_j} + \frac{m}{n - 1} \sum_{k \neq i, j} X_k b(N_k) \frac{X_{kj}}{N_k} \right\}, \tag{7.3b}$$

$$Z'_i = (1 - c) \left\{ (1 - m) Z_i b(N_i) + \frac{m}{n - 1} \sum_{k \neq i} Z_k b(N_k) \right\}, \tag{7.3c}$$

where  $1 \leq i, j \leq n$  and  $i \neq j$  in Eq. (7.3b). Consider, for example, Eq. (7.3b). Noting Eq. (7.2), the three terms in parentheses on the right hand side represent social learners – all with behavior adapted to site  $j$  – that are natives of site  $i$ , immigrants originating at site  $j$ , and immigrants originating at sites other than  $i$  and  $j$ , respectively.

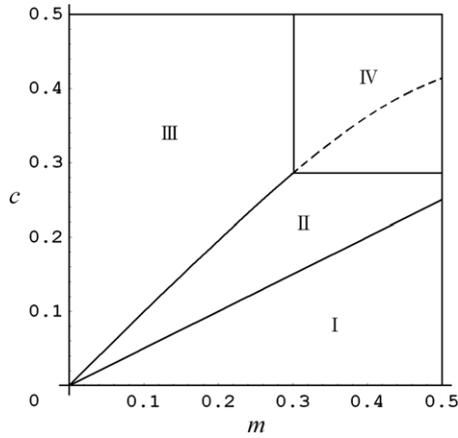
Using local stability analysis and heuristic arguments, Aoki and Nakahashi (2008) identify four classes of stable equilibria of Eq. (7.3), fixation of individual learners ( $\hat{X}_{ii} = 0, \hat{X}_{ij} = 0, \hat{Z}_i > 0$ ; class I), polymorphism of individual learners and social learners ( $\hat{X}_{ii} > 0, \hat{X}_{ij} > 0, \hat{Z}_i > 0$ ; class II), fixation of social learners ( $\hat{X}_{ii} > 0, \hat{X}_{ij} > 0, \hat{Z}_i = 0$ ; class III), and extinction ( $\hat{X}_{ii} = 0, \hat{X}_{ij} = 0, \hat{Z}_i = 0$ ; class IV), which exist for mutually exclusive regions of the parameter space of  $m$  and  $c$  (Fig. 3). A higher migration rate between the environmentally heterogeneous sites of a subdivided population is analogous to a greater instability of the temporally changing environment, as pointed out by Boyd and Richerson (1985, 1988, 1995). Clearly, the evolution of individual learners is more likely at higher migration rates.

However, as we will show later in connection with learning schedules and biased SL, rapid environmental change and high migration rates do not necessarily ensure that (pure) IL will evolve.

Let us now consider Rogers' question of whether the presence of social learners – in addition to, or to the exclusion of, individual learners – enhances the adaptedness of a population at equilibrium. Total population size serves as a proxy measure of adaptedness in this and some other models (Lehmann and Feldman, 2009; Rendell et al., 2010). From symmetry considerations, the total population at the class II equilibria (stable polymorphism of individual learners and social learners) is

$$n\hat{N} = nK [1 - c/r(1 - c)], \tag{7.4}$$

which is the same as when individual learners are fixed. At the class III equilibria (stable monomorphism of social learners), on the other hand, the total population size exceeds the value given by Eq. (7.4). Hence, Rogers' paradox is observed in the former case, but is resolved in the latter.



**Fig. 3.** Heuristic diagram showing the four regions of the  $(m, c)$ -parameter space ( $0 < m \leq 1/2$ ,  $0 < c < s$ ), corresponding to the four classes of stable equilibria of Model 7 (finite island model with environmentally heterogeneous sites). Region I: fixation of individual learners. Region II: polymorphism of individual learners and social learners. Region III: fixation of social learners. Region IV: extinction. Boundary between regions I and II:  $c = ms$ . Boundary between regions II and III:  $c = m(1 - \theta)$ , where  $\theta = \left\{ m - s - \frac{(1-s)m}{n-1} + \sqrt{\left[ m - s - \frac{(1-s)m}{n-1} \right]^2 + 4 \frac{(1-s)m^2}{n-1}} \right\} / (2m)$ . Boundary between regions II and IV: horizontal straight line  $c = r / (1 + r)$ . Boundary between regions III and IV: vertical straight line defined implicitly by  $r = m(1 - \theta) / [1 - m(1 - \theta)]$ . Other parameter values are  $s = 0.5$ ,  $r = 0.4$ ,  $n = 15$ ,  $K = 100$ .

## 5. Learning schedules

With the possible exception of Model 6, all models considered so far assume “one-shot” strategies in which an organism uses IL and/or SL with fixed probabilities throughout its life. In this section, we consider models that allow for two learning stages per generation, in which the probabilities of using IL and/or SL can differ between the two stages. We assume an unstructured population of infinite size with asexual reproduction.

### 5.1. Model 8: critical social learning

Enquist et al. (2007) introduce two novel strategies called “critical social learning” and “conditional social learning”. Critical social learning is a learning schedule in which an organism first uses SL. If the correct behavior is acquired by SL, where the organism is assumed to be capable of judging whether it has succeeded in doing so, no further learning occurs. However, if the wrong behavior is acquired by SL, the organism next tries IL. In conditional social learning, the order in which SL and IL are used is interchanged. For both learning strategies, the exemplar in SL is chosen randomly, and the occurrence of the second learning stage is contingent on the outcome of the first.

In this section, we summarize their proof that critical social learning can be an ESS in the absence of conditional social learners. The competing strategies are then obligate individual learner, obligate social learner, and critical social learner. Enquist et al. (2007) adopt the two-timescale approach, which assumes that cultural evolution can be investigated with the genetic variables held constant (see also Model 3). They obtain the frequency of the correct behavior – they call this “the OK solution” – at the equilibrium of the cultural dynamics,  $q_t^{\text{OK}}$ , which they then use to calculate the selection coefficients – fitnesses, in their terminology – of the competing strategies.

Let us define the following parameters as in their model, using a more compact notation:  $\varepsilon$  is the probability of environmental change per generation (i.e., information decay rate, see Model 4),  $\alpha$  is the probability that an individual learner acquires the

correct behavior,  $\beta$  is the probability that a social learner faithfully copies the exemplar,  $c_i$  is the exogenous cost of IL, and  $c_s$  is the exogenous cost of SL. Then, we can write the selection coefficients of obligate individual learners, obligate social learners, and critical social learners as

$$w_i = \alpha - c_i, \quad (8.1a)$$

$$w_s = q^{\text{OK}}(1 - \varepsilon)\beta - c_s, \quad (8.1b)$$

$$w_{si} = q^{\text{OK}}(1 - \varepsilon)\beta - c_s + [1 - q^{\text{OK}}(1 - \varepsilon)\beta](\alpha - c_i), \quad (8.1c)$$

respectively. Note that  $w_{si} > w_s$  if, as we assume, individual learning is adaptive, i.e.  $\alpha > c_i$ .

To determine  $q^{\text{OK}}$  from the cultural dynamics, we denote the frequencies of obligate individual learners, obligate social learners, and critical social learners – which are by assumption fixed during cultural evolution – by  $q_i$ ,  $q_s$ , and  $q_{si}$ , respectively, and the variable frequency of the OK solution in generation  $t$  by  $q_t^{\text{OK}}$ . Then,

$$q_{t+1}^{\text{OK}} = q_i\alpha + q_s q_t^{\text{OK}}(1 - \varepsilon)\beta + q_{si} \{ q_t^{\text{OK}}(1 - \varepsilon)\beta + [1 - q_t^{\text{OK}}(1 - \varepsilon)\beta]\alpha \}. \quad (8.2)$$

In deriving Eq. (8.2), the exogenous costs,  $c_i$  and  $c_s$ , that appear in Eq. (8.1) have been ignored, using the standard assumption of the two-timescale approach that selection is weak. Hence, at equilibrium (obtained by setting  $q_{t+1}^{\text{OK}} = q_t^{\text{OK}} = q^{\text{OK}}$ ), we have

$$q^{\text{OK}} = \alpha(1 - q_s) / \{ 1 - (1 - \varepsilon)\beta [q_s + q_{si}(1 - \alpha)] \}. \quad (8.3)$$

Substitution of Eq. (8.3) into (8.1) yields the selection coefficients of the three strategies, which depend on their frequencies,  $q_i$ ,  $q_s$ , and  $q_{si}$ .

Since  $w_{si} > w_s$  holds by assumption for all frequencies, critical social learner will be an ESS if  $w_{si} > w_i$  when  $q_{si} = 1$ . Hence, the condition for critical social learner to be an ESS is

$$c_i > \alpha - 1 + c_s [1 - (1 - \alpha)(1 - \varepsilon)\beta] / [\alpha(1 - \varepsilon)\beta]. \quad (8.4)$$

Clearly, Rogers’ paradox is resolved since  $w_{si} > w_i$  at the evolutionarily stable equilibrium where critical social learners are fixed.

### 5.2. Model 9: evolutionarily stable learning schedules with a continuous phenotype

All models presented so far have assumed a dichotomous phenotype that is either correct or wrong. Here, we posit a continuous phenotype with a fitness optimum that depends on the environmental state. The question of learning schedules was first addressed by Boyd and Richerson (1985), who introduced a learning strategy or process (in this latter context) that they called “guided variation”. Guided variation entails that an organism acquires its “initial phenotype” by SL – from the parental generation according to a “blending rule” – which is then adaptively modified by IL to yield the “mature phenotype”. Such a strategy permits a gradual and cumulative improvement of the phenotype over generations.

IL or SL or a mixture of the two can be sequentially combined in an infinite number of ways to form a learning schedule. Aoki et al. (2012) used a method informed by optimal control theory to ask what the evolutionarily stable learning schedules would be under various regimes of environmental change.

Suppose that there are two learning stages per discrete generation, stage-0 and stage-1 learning. Genetically determined learning strategies assign weights between 0 and 1 to the information gathered by IL and by SL during each stage. The phenotype of a naïve newborn before any learning occurs will be called the “initial phenotype” and is set to 0. We refer to the phenotype after stage-0 learning as the “intermediate phenotype” and the phenotype after stage-1 learning as the “mature phenotype”. These terms are

used differently here from Boyd and Richerson (1985). With each learning strategy are associated one intermediate phenotype and one mature phenotype that are sequentially expressed. Only the mature phenotype contributes to fitness. Post-learning organisms reproduce in proportion to their fitness (fertility selection) and survive to serve as exemplars (i.e., cultural parents) for the next generation.

In what follows, we deal only with the situation where the environment undergoes periodic changes twice per generation, i.e., between birth and stage-0 learning, and between stage-0 learning and stage-1 learning. The optimal phenotype is  $-z^*$  during stage-0 learning, whereas it is  $z^*$  during stage-1 learning and reproduction. Since the initial phenotype is set to 0, this requires a special assumption regarding the phenotypic scale. The target for IL in each learning stage is the optimal phenotype in that learning stage. The target for SL is always the mature phenotype of the previous generation if the population is genetically monomorphic, or the population mean of the mature phenotypes if more than one learning schedule is segregating. The efficiencies of IL and SL during stage- $i$  learning are  $\alpha_i$  and  $\beta_i$  ( $i = 0, 1$ ), respectively, where efficiency is defined as the proportional reduction in the deviation from the target and assumed to be the same for all learning strategies. We assume  $0 < \alpha_i, \beta_i < 1$  ( $i = 0, 1$ ), which entails that the mature phenotype approaches, but never converges to, the optimal phenotype.

First, we consider a population that is monomorphic for the resident learning strategy, letting  $z_t$  be the mature phenotype in generation  $t$ . This strategy assigns weight  $u_i$  (called the control) to IL and complementary weight  $1 - u_i$  to SL during stage- $i$  learning, where  $0 \leq u_i \leq 1$  ( $i = 0, 1$ ). In generation  $t + 1$ , pure IL would produce the phenotype  $-\alpha_0 z^*$  during stage-0 learning, whereas pure SL would yield the phenotype  $\beta_0 z_t$ . Hence, taking the weights into consideration, we have the intermediate phenotype

$$z_{t+1}^{int} = -u_0 \alpha_0 z^* + (1 - u_0) \beta_0 z_t. \tag{9.1}$$

On further application of pure IL during stage-1 learning, the incremental change in the phenotype would be  $\alpha_1(z^* - z_{t+1}^{int})$ , because the optimal phenotype and hence the target for IL has changed to  $z^*$ . Pure SL during stage-1 learning would produce the incremental change  $\beta_1(z_t - z_{t+1}^{int})$ . Thus, the mature phenotype is

$$z_{t+1} = z_{t+1}^{int} + u_1 \alpha_1 (z^* - z_{t+1}^{int}) + (1 - u_1) \beta_1 (z_t - z_{t+1}^{int}). \tag{9.2}$$

Substituting Eq. (9.1) into (9.2), the difference equation in the mature phenotype across generations can be written as

$$z_{t+1} = Az_t + Cz^*, \tag{9.3a}$$

where

$$A = (1 - u_0) \beta_0 [1 - u_1 \alpha_1 - (1 - u_1) \beta_1] + (1 - u_1) \beta_1, \tag{9.3b}$$

$$C = u_1 \alpha_1 - u_0 \alpha_0 [1 - u_1 \alpha_1 - (1 - u_1) \beta_1]. \tag{9.3c}$$

At equilibrium of the cultural dynamics described by Eq. (9.3), we have

$$\hat{z} = Cz^* / (1 - A). \tag{9.4}$$

Next, introduce a rare mutant strategy of small effect into the population of resident strategists at this equilibrium. Denote the control for this mutant strategy by  $u_i^m$  ( $i = 0, 1$ ), and let  $A_m$  and  $C_m$  be the functions of  $u_i^m$  ( $i = 0, 1$ ) obtained by substituting  $u_i = u_i^m$  ( $i = 0, 1$ ) in Eqs. (9.3a) and (9.3b), respectively. Then, the mature phenotype of the mutant strategy is given approximately by

$$z^m = A_m \hat{z} + C_m z^*. \tag{9.5}$$

The method for determining an evolutionarily stable control,  $u_i^*$  ( $i = 0, 1$ ), prescribes that we define an objective function,

which gives the fitness of a mutant strategy introduced at low frequency into the population of resident strategist at equilibrium (Maynard Smith, 1982). Specifically, we assume an objective function of the Gaussian form,

$$F(u_0^m, u_1^m, u_0, u_1) = \exp \left\{ -\frac{(z^m - z^*)^2}{w} \right\}, \tag{9.6}$$

where  $z^m$  is given by Eq. (9.5). Eq. (9.6) formalizes the assumption that deviations from the optimal phenotype,  $z^*$ , are penalized, with an intensity of selection inversely proportional to  $w$ .

An evolutionarily stable control,  $u_i^*$  ( $i = 0, 1$ ), is one such that  $F(u_0^m, u_1^m, u_0, u_1)$  takes a (local) maximum when  $u_0^m = u_0 = u_0^*$  and  $u_1^m = u_1 = u_1^*$ . Set

$$s_0 = \left. \frac{\partial F}{\partial u_0^m} \right|_{u_0^m=u_0, u_1^m=u_1}, \tag{9.7a}$$

$$s_1 = \left. \frac{\partial F}{\partial u_1^m} \right|_{u_0^m=u_0, u_1^m=u_1}, \tag{9.7b}$$

which represent the selection gradients on the mutant control for stage-0 and stage-1 learning, respectively. In Appendix F, we show that  $s_0 < 0$  and  $s_1 > 0$  for all values of  $u_0$  and  $u_1$ . Hence, there is only one evolutionarily stable control, which is  $u_0^* = 0, u_1^* = 1$ . In words, when the environment changes twice per generation – between birth and stage-0 learning, and between stage-0 learning and stage-1 learning – we predict an evolutionarily stable learning schedule where pure SL is followed by pure IL, which is equivalent to guided variation. Note that, this schedule incorporates SL in spite of the highly changeable environment.

The fitness of a genetically monomorphic population at the cultural equilibrium,  $\hat{z}$ , is given by

$$W(\hat{z}) = \exp \left\{ -\frac{(\hat{z} - z^*)^2}{w} \right\}, \tag{9.8}$$

which is of the same form as Eq. (9.6). To obtain the fitness of a population using the evolutionarily stable strategy, we substitute  $u_0 = u_0^* = 0$  and  $u_1 = u_1^* = 1$  in Eq. (9.3) and then use Eq. (9.4) to compute  $\hat{z}$ . This yields

$$\hat{z} = \frac{z^* \alpha_1}{1 - (1 - \alpha_1) \beta_0}, \tag{9.9}$$

and it can be shown that this value of  $\hat{z}$  maximizes  $W(\hat{z})$ . Hence, any other control including the pure IL strategy ( $u_0 = 1, u_1 = 1$ ) is associated with a lower fitness at the cultural equilibrium. Thus, Rogers' paradox is not observed in this model.

## 6. Biased social learning

The previous models all assume that SL occurs from a random member of the parental generation (unbiased or random SL). The empirical literature sometimes suggests that organisms choose their exemplar(s) according to certain criteria such as conformity or perceived success (e.g., Mesoudi, 2009; Henrich and Broesch, 2011; Chudek et al., 2012). In this section, we review a model due to Nakahashi et al. (2012) that investigates the conditions under which such biases in SL may evolve.

### 6.1. Model 10: conformist and payoff bias with spatial environmental heterogeneity

Nakahashi et al. (2012) extend Model 7 by the simultaneous introduction of two new SL strategies, conformist bias and payoff bias, which are in competition with each other and with individual

learners and unbiased social learners. As before, there are  $n$  environmentally different sites, each associated with a behavior that is correct there but equally wrong at the  $n - 1$  other sites. Migration at rate  $m$  occurs among the sites after SL but before individual learning. An individual learner always achieves the correct behavior, but suffers an exogenous cost  $c$ . The cost of wrong behavior for an unbiased social learner is  $s$ .

Social learning is said to exhibit conformist bias when the probability that a common phenotype – or the majority phenotype when there are only two options – is adopted exceeds the frequency of that phenotype among the possible exemplars. Specifically, let  $b_{ij}$  be the frequency of organisms at site  $i$  with behavior adapted to site  $j$  in the parental generation, where  $\sum_{j=1}^n b_{ij} = 1$  for each  $i$ . Then, an organism of the offspring generation using the conformist bias strategy at site  $i$  acquires the behavior adapted to site  $j$  with probability

$$\rho_{ij} = (b_{ij})^a / \sum_{k=1}^n (b_{ik})^a, \quad (10.1)$$

where  $a > 1$ . When  $a \rightarrow \infty$ , as is assumed here, the most common behavior at site  $i$  is copied. In particular, if  $b_{ii} > b_{ij}$  for each  $j \neq i$ , then  $\rho_{ii} \rightarrow 1$ , so that the locally correct behavior is copied with probability one in this case.

Payoff bias, on the other hand, is a strategy in which the locally correct behavior is always copied. An organism using the payoff bias strategy will acquire the behavior adapted to its natal site. Such behavior always exists provided individual learners are initially present in the population.

Nakahashi et al. (2012) assume exogenous costs  $d$  to conformist bias and  $g$  to payoff bias, in addition to a cost  $s$  for the wrong behavior. Hence, for example, the fitness of an organism with the payoff bias strategy that has acquired the wrong behavior – i.e., all migrants – will be  $(1 - g)(1 - s)$ . It is assumed that

$$0 < d < g < c < s < 1. \quad (10.2)$$

In particular, inequalities (10.2) entail that the payoff bias is more costly to implement than the conformist bias.

Nakahashi et al. (2012) derive conditions for the local stability of monomorphic and polymorphic equilibria at which the conformist bias exists. They find that a high cost of IL (large  $c$ ) in conjunction with a high migration rate (large  $m$ ) favor the evolution of conformist bias. Thus, high migration rates do not necessarily favor IL when in competition with biased SL. On the other hand, fixation of payoff bias is unstable if  $d < g$ , as is assumed, and polymorphisms involving the payoff bias strategy are either not possible or are unstable when they exist. Hence, payoff bias can never evolve. The intuitive reason for this latter result is that the frequency of the correct behavior at a site eventually exceeds the frequency of any one of the  $n - 1$  wrong behaviors—i.e.,  $b_{ii} > b_{ij}$  for each  $j \neq i$ . Thus, both conformist bias and payoff bias will achieve the correct behavior, but the former strategy suffers a smaller exogenous cost and so out-competes the latter (given  $d < g$ ).

## 7. Emergence of learning

We have reviewed ten models for the evolution of learning strategies in species where IL and SL are already well-established. In doing so, we have ignored the possibility that the phenotype (behavior) may be innately determined. We now consider conditions for the emergence of a partial reliance on learning.

### 7.1. Model 11: mixed strategy model that includes innate behavior

Wakano and Aoki (2006) proposed an extension of Model 5 to include innate behavior. As in Model 5, the environment changes

every  $l$  generations to a previously unknown state. However, this extended mixed strategy model has several differences. First, an organism with the resident strategy uses SL with probability  $K$ , IL with probability  $L$ , and acts innately with probability  $1 - K - L$ . The corresponding probabilities for the mutant strategy are  $K + dK$ ,  $L + dL$ , and  $1 - K - dK - L - dL$ , respectively.

Second, there is a genetic locus (other than the locus that determines the strategy), which carries the phenotypic information that is expressed when the organism behaves innately. We call this genetic locus the “innate information” locus and assume that it has many alleles, which can be classified into those that produce adaptive behavior and those that produce maladaptive behavior in a given environment. A small subset amounting to a fraction  $\rho$  of each class has the special property of producing different behaviors before and after an environmental change. In particular, these rare variants produce adaptive behavior in the post-change generation(s) – they can be regarded as pre-adapted alleles awaiting an environmental change – but the behavior they produce in the pre-change generation(s) is maladaptive as viewed from the post-change generations(s).

Third, an organism behaving innately and expressing the correct behavior is assigned a fitness of 1. Relative to this phenogentotype, SL and IL have exogenous costs  $d$  and  $c$ , respectively. An additional cost,  $s$ , accrues to an organism behaving innately or using SL that has acquired the wrong behavior. By assumption, IL always results in the correct behavior. These costs are additive, such that an organism using SL with the wrong behavior, for example, has fitness  $1 - d - s$ . Moreover,

$$0 < d < c < s < d + s < 1. \quad (11.1)$$

We now consider the conditions for a mutant strategy that relies on a small amount of learning to invade a population that is monomorphic for the pure innate resident strategy. In terms of the probabilities of using SL and/or IL, the resident strategy can be defined as  $K = L = 0$ . Then, the corresponding probabilities for the mutant strategy are  $dK$  and  $dL$ , respectively, where  $dK$  and  $dL$  are small and non-negative, and at least one is positive. For small  $dK$  and  $dL$ , the eigenvalue takes the approximate form

$$\lambda = 1 + C_K dK + C_L dL, \quad (11.2)$$

and the invasion condition is  $\lambda > 1$ .

Wakano and Aoki (2006) show that  $C_K < 0$ . Hence,  $C_L > 0$  and  $dL > 0$  are necessary conditions for invasion. The latter inequality entails that a small amount of SL in isolation ( $dK > 0$ ,  $dL = 0$ ) confers no selective advantage over the pure innate strategy, so that a successful mutant strategy must have an IL component ( $dL > 0$ ). Furthermore,  $C_L > 0$  if and only if

$$\frac{\eta(l)}{\eta(l-1)} < 1 - c, \quad (11.3a)$$

where we define

$$\eta(l) = \rho + (1 - \rho)(1 - s)^l. \quad (11.3b)$$

The ratio  $\frac{\eta(l)}{\eta(l-1)}$  is monotone increasing in  $l$ , with  $\frac{\eta(1)}{\eta(0)} = \rho + (1 - \rho)(1 - s)$  and  $\lim_{l \rightarrow \infty} \frac{\eta(l)}{\eta(l-1)} = 1$ . Hence, provided

$$\rho + (1 - \rho)(1 - s) < 1 - c, \quad (11.4)$$

there exists a unique integer  $l_M \geq 1$  such that (11.3) is satisfied for  $l \leq l_M$ . Restated, the pure innate strategy is evolutionarily stable if inequality (11.4) is reversed or if  $l > l_M$ .

Inequality (11.4) can be rewritten as

$$c < (1 - \rho)s. \quad (11.5)$$

The left hand side of inequality (11.5) is the exogenous cost of IL. The right hand side is the fitness loss to innate behavior due to the

expression of non-adapted alleles at the innate information locus immediately after an environmental change. Thus, the conditions that favor the emergence of learning are (1) a small exogenous cost of IL, (2) a high cost of wrong behavior, (3) a low frequency of pre-adapted alleles at the innate information locus, and (4) a changeable environment. In addition, learning must include some IL.

## 8. Discussion

### 8.1. Comparing the dynamical models in temporally variable environments

Models 1, 2, and 4 are three basic dynamical models for the evolution of obligate individual learners and obligate social learners in temporally variable environments. They produce predictions that differ both qualitatively and quantitatively, the main distinction being between Models 2 and 4 for which fixation of social learners can be stable, and Model 1 for which it cannot. The contradictory predictions can be related to the different assumptions. Models 2 and 4 allow for some of the behaviors occurring in the parental generation to remain adaptive in the offspring generation. On the other hand, Model 1 posits that none of the preexisting behaviors can be adaptive after an environmental change. Model 1 has the merit of being relatively easy to analyze. However, Models 2 and 4 may be better representations of reality, at least in the sense that some continuity in what makes a behavior adaptive across generations is permitted.

As noted above, fixation of social learners can be stable in Models 2 and 4. At such equilibria, social learners with the correct behavior (SLC) and social learners with the wrong behavior (SLW) coexist at what may loosely be described, in analogy with genetics, as a “mutation–selection balance”. Moreover, since there is no input of novel behaviors by IL, the same correct behavior and the same wrong behavior are maintained indefinitely. Specifically, Model 2 produces periodic fluctuations in the frequencies of the two behaviors, whereas Model 4 entails that they are constant. In neither case can there be sustained cultural change. Nevertheless, the mean fitness of the population at such equilibria will be higher than at the other possible equilibria, including fixation of individual learners. Hence, Rogers’ paradox is resolved, but this result arguably introduces another paradox, in that the highest mean fitness is associated with a population that does not evolve culturally.

The three models are deterministic and have in common the assumptions that the environment changes at regular intervals and population size is infinite. Aoki et al. (2005) conducted Monte Carlo simulations of an extended version of Model 1 that includes innate determination of behavior as a third alternative. In these simulations, the environment changes between generations with probability  $p$  – as measured by a uniformly distributed pseudo-random number – so that on average the environment changes every  $1/p$  generations. As expected, they found that the equilibrium frequencies of individual learners, social learners, and innate determination depended on  $1/p$  in the same way as on the fixed period of environmental change,  $l$ , in the deterministic model. Specifically, individual learners, social learners, and innate determination were favored by natural selection when environmental changes occurred at short, intermediate, and long intervals, respectively.

We note in passing that the simple forms of vertical transmission (i.e. SL from parents) are congruent with innate determination. Hence, it comes as no surprise when McElreath and Strimling (2008) find that vertical transmission, where one of the two parents is copied with equal probability, should be selectively favored when the environment is stable.

Rendell et al. (2010) conducted agent-based simulations on the evolution of various SL strategies (pure, critical, and conditional SL)

in a finite population with stochastic temporal changes of the environment. More specifically, their model posits that organisms in this population inhabit cells on a lattice that are arranged as a two-dimensional torus. We will have more to say on this study when we discuss the effects of population structure. Here, we consider their results obtained under the global conditions (SL occurs from a randomly chosen member of the population) and when all cells are in the same environmental state at any one time. In the competition between pure social learners and pure individual learners, Rendell et al. (2010) find that the frequency of social learners and the mean fitness at polymorphic equilibrium, obtained from their simulations, are in good “qualitative” agreement with the analytical predictions of Rogers (1988) and Enquist et al. (2007). However, they also indicate that the frequency of social learners is lower than expected, whereas the mean fitness is higher.

It is not clear to us how the analytical predictions of Rogers (1988) and Enquist et al. (2007) on the frequency of social learners at equilibrium would apply, because the model of Rendell et al. (2010) differs in various respects, such as the number of environmental states and the possible transitions among them. However, the mean fitness of the population when individual learners are fixed is  $1 - c$  in all three models, where  $c$  ( $c_a$  in their notation) is the exogenous cost to individual learners, providing a point of comparison. As noted above, Rendell et al. (2010) find that the mean fitness at polymorphic equilibrium is marginally higher than  $1 - c$ , suggesting that Rogers’ paradox is resolved. These authors do not attribute this discrepancy to environmental stochasticity or finite population size. Rather, they argue that the difference may be caused, for example, by their use of the coevolutionary approach in their simulations, as opposed to the two-timescale argument favored by Rogers (1988) and Enquist et al. (2007). This interpretation is consistent with what we have shown in this paper, that Rogers’ paradox is inherent in the two-timescale method at least as applied to the evolution of pure strategies (Model 3), whereas it can be resolved in a coevolutionary model with two (Model 2) – or a finite number (Model 7) of – environmental states.

### 8.2. ESS mixed strategy models in temporally variable environments

Models in which an organism can engage in both IL and SL are arguably more realistic than those that assume it can only do one or the other. The goal with these mixed strategy models is to investigate the dependence of the evolutionarily stable probability,  $L$ , of using IL (as in Model 5) – or the complementary probability,  $K$ , of using SL (as in Model 6) – on the environmental variability.

Model 5, which assumes an infinite number of environmental states, is the direct analog of the dynamical model, Model 1. Hence, it comes as no surprise that the ESS value of  $L$  predicted by Model 5 depends on the environmental variability in qualitatively the same way as the equilibrium frequency of obligate individual learners does in Model 1. Specifically, both increase as the environmental stability decreases, such that total reliance on IL ( $L = 1$ ) and fixation of individual learners can both be stable if the environment is sufficiently changeable. On the other hand, total reliance on SL ( $L = 0$ ), or equivalently fixation of social learners, cannot be stable in either model. However, Rogers’ paradox was seen to occur in Model 1, but to be eliminated in Model 5—for the latter model, the geometric mean of the mean fitnesses is higher at an interior ESS than in a population with  $L = 1$ . The reason for this discrepancy remains unclear.

Model 6 is the coevolutionary version of a model described by Boyd and Richerson (1988, 1995) for fluctuating environments (Wakano and Aoki, 2007). The original model was formulated and analyzed by a combination of the two-timescale and information decay approaches. The ESS analysis of Model 6 agrees qualitatively

with Boyd and Richerson (1988, 1995) in finding that the total reliance on SL ( $K = 1$ ) cannot be an ESS, whereas the total reliance on IL ( $K = 0$ ) can in a rapidly fluctuating environment. At intermediate levels of environmental variability, an organism evolves to depend on both IL and SL. Hence, the original and reformulated models would appear to produce qualitatively consistent predictions.

In the model of Boyd and Richerson (1988, 1995), the mean fitness associated with an ESS mixing IL and SL is greater than the mean fitness of the pure IL strategy. We suspect Rogers' paradox is also resolved in Model 6, since it has a mathematical structure similar to Model 5, for which we have already seen that Rogers' paradox does not occur. Boyd and Richerson (1995) argue that Rogers' paradox is resolved in their model, because increased reliance on SL has the effect of improving the accuracy of IL (i.e., reducing the probability that the wrong behavior is acquired when IL is used), as suggested by their Fig. 3. We have no reason to question this interpretation. However, it should be noted that Rogers' paradox is also eliminated in Model 5, where IL always results in the correct behavior regardless of the degree of reliance on SL.

### 8.3. Spatially variable environment

Model 7 addresses the effect of spatial environmental heterogeneity on the evolution of obligate individual learners and obligate social learners. The habitable world comprises a finite number of islands, each with a different environment and among which organisms may migrate. All stable equilibria of this deterministic model are symmetric; that is the numbers of individual learners, SLC, and SLW are the same at all sites. Fixation of individual learners, polymorphism of individual learners and social learners, fixation of social learners, and extinction are the possible stable equilibria. Let us take a horizontal transect through Fig. 3 at a value of the exogenous cost to individual learners,  $c$ , that is small relative to the cost of maladaptive behavior,  $s$ . Then we can see that fixation of social learners is stable for small values of the migration rate,  $m$ , polymorphism of individual learners and social learners is stable for intermediate values of  $m$ , and fixation of individual learners is stable for large values of  $m$ . In addition, it can be shown that the total population size is greater at the stable monomorphism of social learners than at the coexisting unstable monomorphism of individual learners, which can be interpreted as indicating that Rogers' paradox is resolved.

Let us compare these observations with the findings of Rendell et al. (2010) from their agent-based simulations. Their treatment of the competition between individual learners and pure social learners under the local conditions (SL occurs from neighbors and dispersal is limited to neighboring sites) with spatial variation in the environment corresponds most closely to the postulates of Model 7. An important parameter in the model of Rendell et al. (2010) is the spatial correlation,  $p_n$ , defined as the probability that neighboring cells have the same environmental state. Small values of  $p_n$  in this model would appear to be analogous to large values of  $m$  in Model 7. This is because smaller values of the spatial correlation entail that a newborn social learner is more likely to be exposed to a neighbor that has experienced a different environmental state from its own. Similarly, higher migration rates entail that a newborn social learner is more likely to acquire its behavior from an immigrant with locally maladaptive behavior.

Rendell et al. (2010) find that the proportion of social learners at equilibrium increases as  $p_n$  increases (see their Fig. 4). Moreover, social learners can be “effectively fixed”—their model includes mutation among the strategies. These results mirror the predictions of Model 7 when we note the analogy between  $p_n$  and  $m$ . Rendell et al. (2010) also claim that Rogers' paradox does not

apply when the spatial correlation is high and social learners are effectively fixed. This result is what we would expect from our analysis of Model 7, but is apparently at odds with their Fig. 4 for the case that a social learner does not initially learn from its parent.

### 8.4. Learning schedules and cumulative culture

Boyd and Richerson (1985) were the first to deal with the evolution of learning schedules. In their model of guided variation, a continuous cultural trait is acquired by a two-step process, whereby an initial phenotype that is acquired by oblique SL is adaptively modified by IL to yield the mature phenotype. These authors rely on the two-timescale approach to predict the “evolutionary equilibrium” contributions of SL and IL to the mature phenotype. Enquist et al. (2007) introduced a strategy called critical social learning, which also entails that SL occurs before IL.

Learning schedules are of particular interest as they relate to the possibility of cumulative culture. By cumulative culture we mean a cumulative improvement in the adaptiveness of a cultural trait, although the term can also describe an increase in the number of adaptive cultural traits. A learning schedule in which each organism accurately absorbs an extant variant of a cultural trait by SL and then builds on it by IL – SL followed by IL – can be supportive of cumulative culture. In this regard, guided variation can result in a gradual and cumulative improvement of the phenotype over generations. However, critical social learning cannot, because this learning strategy entails that the behavior acquired by SL is rejected if it is judged to be maladaptive, in which case IL must occur from scratch.

An explicit study of learning schedules in the context of cumulative culture was made by Aoki et al. (2012), part of which has been reproduced as Model 9. The model assumes two learning stages per discrete generation in each of which an organism can use IL, SL, or a mixture of the two. Analysis of Model 9 showed that if the environment is highly changeable, undergoing periodic fluctuations twice per generation, the evolutionarily stable learning schedule comprised pure SL during the earlier learning stage and pure IL during the latter. Pure SL followed by pure IL can also be the evolutionarily stable learning schedule in a constant environment, depending on the efficiencies of SL and IL (Aoki et al., 2012). Interestingly, a mixture of IL and SL was never found to be evolutionarily stable for either of the learning stages.

An “unrealistic” aspect of the model of Aoki et al. (2012) is the prediction noted above that an organism should use only IL or SL to the exclusion of the other during any one learning stage. This result can, of course, be modified to be more consonant with actual learning behavior by introducing stage-dependent exogenous costs to IL and/or SL. The guided variation model of Boyd and Richerson (1985), on the other hand, allows for intermediate levels of IL and SL at evolutionary equilibrium, even in the absence of exogenous costs. Unfortunately, the two models are not really comparable – neither model can be reduced to a special case of the other – so that it is not possible to ascertain whether the contradictory outcomes can be reconciled.

One reason why the guided variation model sometimes predicts intermediate levels of IL and SL may be that this model can be interpreted as providing only one learning stage per generation, instead of two learning stages as in the model of Aoki et al. (2012). Thus, Eq. (4.9) of Boyd and Richerson (1985) defines a weighted average of the phenotypes that are acquired by SL and IL, as explained in further detail by these authors on page 97 of their book. In fact, their Eq. (4.9) is compatible with the assumption that SL and IL occur concurrently, rather than in some temporal sequence. Hence, the intermediate levels of IL and SL arising from this model may be congruent with the one-shot mixed strategies observed in Models 5 and 6.

### 8.5. Biased social learning

Conformist bias has been studied from the standpoint of evolutionary theory and in psychological experiments. Evolutionary theorists have provided several definitions of conformity in the context of SL. Eq. (10.1) is suitable for dealing with the dynamics of three or more culturally transmitted phenotypes (Lachlan et al., 2004; Nakahashi, 2007). On the other hand, when there are only two options, the probability that the focal variant – which exists at frequency  $p$  – is chosen is usually written as

$$p + Dp(1 - p)(2p - 1),$$

where  $D$  is a parameter assumed to be positive (Boyd and Richerson, 1985, p. 208). That is, the majority (minority) phenotype is adopted with a probability greater (less) than its representation in the population. Different formulations of conformist bias are possible in finite populations, where, for example, each newborn samples a relatively small number of exemplars from the parental generation and adopts the majority phenotype in that sample (Eriksson et al., 2007; Aoki et al., 2011).

Empirical evidence for conformity in humans as defined above – psychologists use a slightly different definition – is poor (Boyd and Richerson, 1985, pp. 223–227; Eriksson and Coultas, 2009; Eriksson et al., 2007; Claidière et al., 2012). On the other hand, theoretical work shows that the conditions for the evolution of conformist bias are not particularly stringent. For example, Model 10 predicts that a high migration rate in a spatially heterogeneous environment, in conjunction with a high cost of IL, favors conformist bias. Nakahashi (2007), Wakano and Aoki (2007), and Kendal et al. (2009) obtained an analogous result that conformity is selected for when the environment changes rapidly in time. Henrich and Boyd (1998) conducted a numerical study on a model incorporating both spatial and temporal variability of the environment, and reached the conclusion that “conformist transmission is favored under a very broad range of conditions”. This may be true, but it is also true that selection on conformist bias is often extremely weak (Wakano and Aoki, 2007). Hence, conformist bias may not necessarily evolve, even under conditions that favor it. Eriksson et al. (2007) give additional reasons for doubting that conformist bias is a general phenomenon.

Payoff bias and direct bias are closely related concepts. Direct bias can be defined as “a naïve individual (choosing) his/her exemplar (cultural parent) based only on the competence of that exemplar in the specific skill that is to be copied” (Kobayashi and Aoki, 2012). This definition is consistent with the one given originally by Boyd and Richerson (1985, p. 137) and with the current usage in the theoretical literature on cultural evolutionary rates (Powell et al., 2009; Mesoudi, 2011). More generally, direct bias means that a particular variant of a cultural trait is preferred and an individual carrying that variant is identified and copied. The term payoff bias makes more explicit the assumption that a phenotypic variant is more likely to be adopted if it is associated with a higher fitness; Boyd and Richerson (1985) were agnostic with regard to this aspect of direct bias. Payoff bias can take many forms. For example, Model 10 due to Nakahashi et al. (2012) assumes that an organism implementing payoff bias always identifies and acquires the (most) adaptive behavior, no matter how low the frequency of that behavior in the population. Alternatively, the fidelity of SL may be set proportional to the fitness benefit from adopting the adaptive behavior (Kendal et al., 2009).

Empirical studies in the laboratory (for review see Mesoudi, 2009; Chudek et al., 2012) and in the field (Henrich and Broesch, 2011) may suggest present-day humans are capable of payoff and/or direct bias. Apparently, the evidence for direct and indirect biases – the latter includes attraction to prestigious and/or successful individuals (Boyd and Richerson, 1985) – is stronger

than for conformist bias. It is then ironic that Model 10 predicts conformist bias will out-compete payoff bias in a spatially variable environment. Nakahashi et al. (2012) also consider the temporal infinite environmental states analog of Model 10, for which they find that payoff bias can be maintained in the population. Unfortunately, this model is not well formulated, since it assumes the presence of individual learners from which the payoff bias strategists can acquire the adaptive behavior, without explicitly incorporating them into the dynamical equations. On the other hand, Kendal et al. (2009), using the information decay approach to model temporal environmental change (see Model 4), show that payoff bias in competition with individual learners can reach polymorphic frequencies. This latter result suggests that Nakahashi et al. (2012) may be right for the wrong reasons.

Clearly, the evolution of biases in SL is an important issue that needs to be pursued further. Such biases can interact with population size to have a large effect on cultural evolutionary rates and have been invoked to explain archeological observations on changes in lithic traditions of various hominid species (Henrich, 2004; Powell et al., 2009; Mesoudi, 2011; Aoki et al., 2011; Kobayashi and Aoki, 2012).

### 8.6. Extensions

Our assumptions may be translated into the language of decision making and behavior. In most of the models (e.g., Models 1–5, 7, 10, 11), individual learners always make the correct decision about what they would gain (in fitness) from the environment, but usually pay a cost to do so. Social learners pay a cost when they make an incorrect decision, in which case they behave inappropriately for the environment they are in, and the latter cost is greater than that paid by individual learners. For human learners, it may be difficult to discern which behavior should be “invented” as an individual learner or “copied” as a social learner, because the optimum behavior in a given environment may be difficult to identify. Kahneman and Tversky (1979) showed that in choosing between sets of payoffs with different probabilities it is not always the highest mean payoff that humans decide upon. In our terminology, it may be very difficult to decide whom to copy in a given environment, or what the payoffs to possible decision choices are. Quantification of uncertainty and the probability of acting upon perceived measures of uncertainty are features of evolution that are crucial to learning, and in principle both should contribute to fitnesses of different learning strategies.

A similar issue has been shown to arise in recent agent-based models of SL and IL in uncertain environments. The context is the producer–scrounger game, where some organisms, producers, discover resources, and others, scroungers, then join them and take advantage of the producers’ discoveries (Barnard and Sibly, 1981; Giraldeau and Caraco, 2000). This game provides a frequency-dependent situation in which IL and SL strategies can compete with innate behavior, and the structure of the competitive environment can be manipulated to give appropriate optimum combinations of these strategies.

Arbilly et al. (2011) used this approach with an environment they designed to have the highest payoff occur with the lowest probability. In their simulations, agents learned in which patch to forage either individually (as a producer) or socially (as a scrounger) by observing or joining a producer, over many time steps during the agents’ lifetimes. When the number of learning steps per lifetime was large enough, the social learners were able to learn to forage on the high-payoff-low-probability food sources, which resulted in their ultimately taking over the population. Other such studies have shown that learners can invade a population of non-learners (Dubois et al., 2010) and that if the producer–scrounger dimorphism is not perfectly heritable, learners

could come to dominate the population when the amount of resource was either fixed or fluctuating (Katsnelson et al., 2011).

Although these agent-based simulations involve complexities of finite population size, multiple phenogenotypes and complex environments, which preclude their representation by the recursion systems that we would like to analyze formally, they provide an informative complement to the mathematical analyses reviewed here. They also suggest avenues along which the mathematical models might profitably be developed, albeit in a more abstract form.

### Acknowledgments

We are grateful to many colleagues. In particular, Wataru Nakahashi and Magnus Enquist provided clarification on their models, Laurent Lehmann and Jeremy Kendal made helpful comments on the first draft, and Shripad Tuljapurkar contributed Appendix C. This research was supported in part by Monbukagakusho grant 22101004 to KA and NIH grant GM28016 to MWF.

### Appendix A. Infinite-states $l$ -cycle model (Model 1)

We summarize the local stability analysis for the two monomorphic equilibria.

1.  $\hat{z}^{(i)} = 1$  for  $0 \leq i \leq l - 1$ .

When the environment changes between generations,

$$x' + y' = \frac{1-s}{1-c} (x + y).$$

When the environment does not change,

$$x' + y' = \frac{1}{1-c} (x + y).$$

Hence after  $l$  generations,

$$x^{(l)} + y^{(l)} = \frac{1-s}{(1-c)^l} (x^{(0)} + y^{(0)}).$$

Thus, this equilibrium is stable if  $\frac{1-s}{(1-c)^l} < 1$ , i.e., if  $l < \frac{\log(1-s)}{\log(1-c)}$ .

2.  $\hat{z}^{(i)} = 0$ ,  $\hat{x}^{(i)} = 0$ ,  $\hat{y}^{(i)} = 1$  for  $0 \leq i \leq l - 1$ .

Assume that  $\hat{z}^{(i)} = 0$ . Then Eq. (1.1) gives  $\frac{\hat{x}^{(i)}}{\hat{y}^{(i)}} = 0$ , and Eq. (1.2) gives  $\frac{\hat{x}^{(i+1)}}{\hat{y}^{(i+1)}} = \frac{1}{1-s} \cdot \frac{\hat{x}^{(i)}}{\hat{y}^{(i)}}$  for  $1 \leq i \leq l - 1$ . Hence,  $\hat{y}^{(i)} = 1$  for  $0 \leq i \leq l - 1$ , when social learners are fixed.

After an environmental change,

$$x^{(1)} = 0, \quad z^{(1)} = \frac{1-c}{1-s} z^{(0)}.$$

Without an environmental change,

$$x^{(i+1)} = \frac{1}{1-s} (x^{(i)} + z^{(i)}), \quad z^{(i+1)} = \frac{1-c}{1-s} z^{(i)}.$$

Hence,

$$\begin{aligned} \begin{pmatrix} x^{(l)} \\ z^{(l)} \end{pmatrix} &= \begin{pmatrix} \frac{1}{1-s} & \frac{1}{1-s} \\ 0 & \frac{1-c}{1-s} \end{pmatrix}^{l-1} \begin{pmatrix} 0 & 0 \\ 0 & \frac{1-c}{1-s} \end{pmatrix} \begin{pmatrix} x^{(0)} \\ z^{(0)} \end{pmatrix} \\ &= \begin{pmatrix} 0 & * \\ 0 & \left(\frac{1-c}{1-s}\right)^l \end{pmatrix} \begin{pmatrix} x^{(0)} \\ z^{(0)} \end{pmatrix} \end{aligned}$$

where  $*$  indicates a non-zero quantity. The eigenvalues do not depend on the order of matrix multiplication (Caswell, 2001, pp. 350–351), and since  $c < s$  the leading eigenvalue is larger than 1, and the equilibrium is unstable.

### Appendix B. Two-state $l$ -cycle model (Model 2)

We identify the genetically monomorphic equilibria and determine their local stability conditions.

1.  $\hat{z}^{(i)} = 1$  for  $0 \leq i \leq l - 1$ .

After an environmental change,

$$x' + y' = \frac{1-s}{1-c} (x + y).$$

Without an environmental change,

$$x' + y' = \frac{1}{1-c} (x + y).$$

Hence after  $l$  generations,

$$x^{(l)} + y^{(l)} = \frac{1-s}{(1-c)^l} (x^{(0)} + y^{(0)}).$$

Thus stable if  $\frac{1-s}{(1-c)^l} < 1$ , i.e., if  $l < \frac{\log(1-s)}{\log(1-c)}$ .

2.  $\hat{x}^{(i)} + \hat{y}^{(i)} = 1$  for  $0 \leq i \leq l - 1$ .

The natural period of the periodic solutions is  $2l$  as shown below. After the first environmental change,

$$\frac{x^{(1)}}{y^{(1)}} = \frac{1}{1-s} \cdot \frac{y^{(0)}}{x^{(0)}}.$$

During the subsequent  $l - 1$  generations without change,

$$\frac{x^{(i)}}{y^{(i)}} = \frac{1}{1-s} \cdot \frac{x^{(i-1)}}{y^{(i-1)}}$$

where  $2 \leq i \leq l$ . Similarly, after the second environmental change

$$\frac{x^{(l+1)}}{y^{(l+1)}} = \frac{1}{1-s} \cdot \frac{y^{(l)}}{x^{(l)}},$$

and during the following  $l - 1$  generations of stasis,

$$\frac{x^{(l+i)}}{y^{(l+i)}} = \frac{1}{1-s} \cdot \frac{x^{(l+i-1)}}{y^{(l+i-1)}}$$

for  $2 \leq i \leq l$ . Hence,

$$\frac{x^{(2l)}}{y^{(2l)}} = (1-s)^{-l} \cdot \frac{y^{(l)}}{x^{(l)}} = (1-s)^{-l} \cdot (1-s)^l \cdot \frac{x^{(0)}}{y^{(0)}} = \frac{x^{(0)}}{y^{(0)}},$$

which entails that there exist an infinite number of periodic solutions of period  $2l$ , each of which is neutrally stable (see eigenvalue  $\lambda_1$  below).

Next, we consider the local stability of each of these periodic solutions. Set  $y^{(i)} = \hat{y}^{(i)} + \varepsilon_y^{(i)}$ ,  $z^{(i)} = \varepsilon_z^{(i)}$ . After the first environmental change,

$$\begin{aligned} W^{(0)} &= (1-c) \varepsilon_z^{(0)} + [1-s(1-\hat{y}^{(0)} - \varepsilon_y^{(0)})] (1-\varepsilon_z^{(0)}) \\ &\approx 1-s(1-\hat{y}^{(0)}) + s\varepsilon_y^{(0)} + O(\varepsilon_z^{(0)}), \\ \therefore \varepsilon_y^{(1)} &\approx \frac{(1-s)(1-\varepsilon_z^{(0)})(1-\hat{y}^{(0)} - \varepsilon_y^{(0)})}{1-s(1-\hat{y}^{(0)}) + s\varepsilon_y^{(0)} + O(\varepsilon_z^{(0)})} - \frac{(1-s)(1-\hat{y}^{(0)})}{1-s(1-\hat{y}^{(0)})} \\ &\approx -\frac{1-s}{[1-s(1-\hat{y}^{(0)})]^2} \varepsilon_y^{(0)} + O(\varepsilon_z^{(0)}) \\ \varepsilon_z^{(1)} &= \frac{1-c}{1-s(1-\hat{y}^{(0)})} \varepsilon_z^{(0)}. \end{aligned}$$

Without change,

$$\begin{aligned} W^{(i-1)} &= (1-c) \varepsilon_z^{(i-1)} + [1-s(\hat{y}^{(i-1)} + \varepsilon_y^{(i-1)})] (1-\varepsilon_z^{(i-1)}) \\ &\approx 1 - s\hat{y}^{(i-1)} - s\varepsilon_y^{(i-1)} + O(\varepsilon_z^{(i-1)}), \\ \varepsilon_y^{(i)} &\approx \frac{(1-s)(1-\varepsilon_z^{(i-1)})(\hat{y}^{(i-1)} + \varepsilon_y^{(i-1)})}{1 - s\hat{y}^{(i-1)} - s\varepsilon_y^{(i-1)} + O(\varepsilon_z^{(i-1)})} - \frac{(1-s)\hat{y}^{(i-1)}}{1 - s\hat{y}^{(i-1)}} \\ &= \frac{1-s}{(1-s\hat{y}^{(i-1)})^2} \varepsilon_y^{(i-1)} + O(\varepsilon_z^{(i-1)}), \\ \varepsilon_z^{(i)} &= \frac{1-c}{1-s\hat{y}^{(i-1)}} \varepsilon_z^{(i-1)}. \end{aligned}$$

Analogous equations hold for the next  $l$  generations. It follows that the eigenvalues of the local stability matrix are

$$\lambda_1 = (1-s)^{2l} \mu^2 \quad \text{and} \quad \lambda_2 = (1-c)^{2l} \mu,$$

where

$$\mu = \frac{1}{1-s\hat{x}^{(0)}} \left( \prod_{i=1}^{l-1} \frac{1}{1-s\hat{y}^{(i)}} \right) \cdot \frac{1}{1-s\hat{x}^{(l)}} \left( \prod_{i=1}^{l-1} \frac{1}{1-s\hat{y}^{(l+i)}} \right).$$

Note that,  $\mu$  is the reciprocal of the product of the mean fitnesses over the  $2l$  generations.

To evaluate these eigenvalues, we need to obtain the explicit solution for the periodic equilibrium. Set  $K_0 = \frac{\hat{y}^{(0)}}{\hat{x}^{(0)}}$  and  $K_l = \frac{\hat{y}^{(l)}}{\hat{x}^{(l)}}$ . Then

$$\begin{aligned} 1 - s\hat{y}^{(i)} &= (1-s) \frac{1 + (1-s)^{-(i+1)} K_0}{1 + (1-s)^{-i} K_0} \\ 1 - s\hat{y}^{(l+i)} &= (1-s) \frac{1 + (1-s)^{-(i+1)} K_l}{1 + (1-s)^{-i} K_l} \end{aligned}$$

for  $1 \leq i \leq l$ . Moreover,

$$\begin{aligned} 1 - s\hat{x}^{(0)} &= \frac{1-s+K_0}{1+K_0} \\ 1 - s\hat{x}^{(l)} &= \frac{1-s+K_l}{1+K_l}. \end{aligned}$$

Hence,

$$\mu = (1-s)^{-2l} \frac{1+K_0}{1+(1-s)^{-l}K_0} \cdot \frac{1+K_l}{1+(1-s)^{-l}K_l}.$$

But

$$K_l = \frac{\hat{y}^{(l)}}{\hat{x}^{(l)}} = (1-s)^l \frac{\hat{x}^{(0)}}{\hat{y}^{(0)}} = (1-s)^l K_0^{-1}.$$

Hence,

$$\mu = (1-s)^{-2l} \frac{1+K_0}{1+K_l^{-1}} \cdot \frac{1+K_l}{1+K_0^{-1}} = (1-s)^{-2l} K_0 K_l = (1-s)^{-l},$$

which entails that the geometric mean of the mean fitnesses is  $\sqrt{1-s}$ .

Thus, we finally have

$$\lambda_1 = 1 \quad \text{and} \quad \lambda_2 = \frac{(1-c)^{2l}}{(1-s)^l}.$$

Since the corresponding eigenvectors are orthogonal, we conclude that the fixation of social learners is stable to invasion by individual learners if  $(1-c)^2 < 1-s$ .

### Appendix C. Probability of an even number of environmental changes in $t$ generations (Model 3)

Let  $q$  be the probability of an environmental change per generation, and set  $p = 1 - q$ . The probability of an even number of environmental changes in  $t$  generations, including the case of no change, is

$$\pi = \binom{t}{0} q^0 p^t + \binom{t}{2} q^2 p^{t-2} + \dots + \binom{t}{t} q^t p^0$$

when  $t$  is even, and

$$\pi = \binom{t}{0} q^0 p^t + \binom{t}{2} q^2 p^{t-2} + \dots + \binom{t}{t-1} q^{t-1} p^1$$

when  $t$  is odd. In either case, we can rewrite  $\pi$  as

$$\pi = \frac{1}{2} [(p+q)^t + (p-q)^t].$$

In particular, if we set  $q = \varepsilon$  and  $p = 1 - \varepsilon$ , we obtain

$$\pi = \frac{1}{2} [1 + (1-2\varepsilon)^t].$$

### Appendix D. Fully polymorphic equilibria of the information decay model (Model 4)

From Eqs. (4.1c) and (4.1d), we have at equilibrium

$$\hat{z}(1-\hat{z})\{c-s[\varepsilon+\hat{y}(1-\varepsilon)]\} = 0.$$

Since  $0 < \hat{z} < 1$ ,

$$\hat{y} = \frac{c-s\varepsilon}{s(1-\varepsilon)},$$

which is valid for  $s\varepsilon < c$ . Next, from Eqs. (4.1a) and (4.1b), we have

$$\frac{\hat{x}}{\hat{y}} = \frac{(1-\hat{y})(1-\varepsilon)}{(1-s)[\varepsilon+\hat{y}(1-\varepsilon)]},$$

from which we obtain

$$\hat{x} = \frac{c-s\varepsilon}{s(1-\varepsilon)} \cdot \frac{s-c}{c(1-s)}.$$

Hence,

$$\hat{z} = 1 - \hat{x} - \hat{y} = \frac{(s-c)(\varepsilon-c)}{c(1-s)(1-\varepsilon)},$$

which is valid for  $c < \varepsilon$ . In summary, the fully polymorphic equilibrium exists if  $s\varepsilon < c < \varepsilon$ .

Next, using  $y = \hat{y} + \varepsilon_y$  and  $z = \hat{z} + \varepsilon_z$  as the variables for the local stability analysis, the characteristic polynomial is

$$M(\lambda) = \begin{vmatrix} \frac{c-s\varepsilon}{c} \cdot \frac{1-\varepsilon-(s-c)}{(1-\varepsilon)(1-s)} - \lambda & -\frac{c(1-s)}{s(1-c)} \\ \frac{s(c-s\varepsilon)}{(1-\varepsilon)c(1-s)} & 1-\lambda \end{vmatrix}.$$

Assuming real eigenvalues, we can write this as

$$\begin{aligned} M(\lambda) &= \left[ \frac{c-s\varepsilon}{c} \cdot \frac{1-\varepsilon-(s-c)}{(1-\varepsilon)(1-s)} - \lambda \right] (1-\lambda) \\ &\quad + \frac{c(1-s)}{s(1-c)} \cdot \frac{(s-c)(\varepsilon-c)}{(1-\varepsilon)c(1-s)} \cdot \frac{s(c-s\varepsilon)}{c(1-s)}. \end{aligned}$$

By assumption  $c < s$ , and the condition for existence is  $s\varepsilon < c < \varepsilon$ . Hence,

$$\begin{aligned} 0 &< \frac{c-s\varepsilon}{c} \cdot \frac{1-\varepsilon-(s-c)}{(1-\varepsilon)(1-s)} < 1 \quad \text{and} \\ \frac{c(1-s)}{s(1-c)} \cdot \frac{(s-c)(\varepsilon-c)}{(1-\varepsilon)c(1-s)} \cdot \frac{s(c-s\varepsilon)}{c(1-s)} &> 0, \end{aligned}$$

and both eigenvalues will be positive and smaller than 1. Assuming complex eigenvalues, on the other hand, we can rewrite the characteristic polynomial as

$$M(\lambda) = \lambda^2 - \left[ 1 + \frac{c - s\varepsilon}{c} \cdot \frac{1 - \varepsilon - (s - c)}{(1 - \varepsilon)(1 - s)} \right] \lambda + \frac{c - s\varepsilon}{c(1 - c)},$$

where the constant term can be shown to satisfy  $0 < \frac{c - s\varepsilon}{c(1 - c)} < 1$ . Hence, the eigenvalues will be less than 1 in absolute value. Thus, whether the eigenvalues are real or complex, the existence of the equilibrium implies its stability.

#### Appendix E. Analysis of mixed strategy model with infinite-states $l$ -cycle (Model 5)

To prove that the periodic equilibrium of the resident type is given by Eq. (5.3), we note from Eqs. (5.1a) and (5.1b) that

$$\frac{1}{\hat{x}^{(1)}} = \frac{L(1 - c) + (1 - L)(1 - s)}{(1 - L)(1 - s)} = \alpha - \frac{s}{1 - s},$$

and from Eqs. (5.2a) and (5.2b) that

$$\begin{aligned} \frac{1}{\hat{x}^{(i)}} &= \frac{L(1 - c) + (1 - L)(1 - s\hat{x}^{(i-1)})}{(1 - L)(1 - s)\hat{x}^{(i-1)}} \\ &= \alpha \cdot \frac{1}{\hat{x}^{(i-1)}} - \frac{s}{1 - s} \end{aligned}$$

for  $2 \leq i \leq l$ . Hence

$$\begin{aligned} \frac{1}{\hat{x}^{(i)}} &= \alpha^{i-1} \frac{1}{\hat{x}^{(1)}} - \frac{s}{1 - s} \sum_{j=0}^{i-2} \alpha^j \\ &= \alpha^i - \frac{s}{1 - s} \sum_{j=0}^{i-1} \alpha^j \\ &= \frac{1 - \alpha + \beta(1 - \alpha^i)}{1 - \alpha}. \end{aligned}$$

Next, we obtain the condition for invasion of the mutant type. Adding Eqs. (5.1c) and (5.1d) gives

$$y^{(1)} + \bar{y}^{(1)} = A(y^{(0)} + \bar{y}^{(0)}),$$

where

$$A = \frac{(L + dL)(1 - c) + (1 - L - dL)(1 - s)}{L(1 - c) + (1 - L)(1 - s)} > 0.$$

Similarly, adding Eqs. (5.2c) and (5.2d) gives

$$y^{(i)} + \bar{y}^{(i)} = B^{(i-1)}(y^{(i-1)} + \bar{y}^{(i-1)}),$$

where

$$B^{(i)} = \frac{(L + dL)(1 - c) + (1 - L - dL)(1 - s\hat{x}^{(i)})}{L(1 - c) + (1 - L)(1 - s\hat{x}^{(i)})} > 0.$$

Hence the residents will be invaded by mutants if

$$\lambda = A \prod_{i=1}^{l-1} B^{(i)} > 1.$$

#### Appendix F. Sign of selection gradients in Model 9

The selection gradients are

$$\begin{aligned} s_0 &= \frac{\partial F}{\partial u_0^m} \Big|_{u_0^m=u_0, u_1^m=u_1} \\ &= - \frac{2(\hat{z} - z^*)}{w} \exp \left\{ - \frac{(\hat{z} - z^*)^2}{w} \right\} \frac{\partial z^m}{\partial u_0^m} \Big|_{u_0^m=u_0, u_1^m=u_1}, \\ s_1 &= \frac{\partial F}{\partial u_1^m} \Big|_{u_0^m=u_0, u_1^m=u_1} \\ &= - \frac{2(\hat{z} - z^*)}{w} \exp \left\{ - \frac{(\hat{z} - z^*)^2}{w} \right\} \frac{\partial z^m}{\partial u_1^m} \Big|_{u_0^m=u_0, u_1^m=u_1}. \end{aligned}$$

It can readily be shown that

$$z^* - \hat{z} = z^* - z^m \Big|_{u_0^m=u_0, u_1^m=u_1} = \frac{z^*(1 - A - C)}{1 - A} > 0,$$

where  $A$  and  $C$  are given by Eq. (9.3). Moreover,

$$\begin{aligned} (z^*)^{-1} [1 - u_1\alpha_1 - (1 - u_1)\beta_1]^{-1} (1 - A) \frac{\partial z^m}{\partial u_0^m} \Big|_{u_0^m=u_0, u_1^m=u_1} \\ = - \{u_1\alpha_1\beta_0(1 + \alpha_0) + \alpha_0(1 - \beta_0)[1 - (1 - u_1)\beta_1]\} \end{aligned}$$

and

$$\begin{aligned} (z^*)^{-1} (1 - A) \frac{\partial z^m}{\partial u_1^m} \Big|_{u_0^m=u_0, u_1^m=u_1} \\ = \alpha_1(1 - \beta_1)[1 + u_0\alpha_0 - (1 - u_0)\beta_0]. \end{aligned}$$

Clearly,

$$\frac{\partial z^m}{\partial u_0^m} \Big|_{u_0^m=u_0, u_1^m=u_1} < 0 \quad \text{and} \quad \frac{\partial z^m}{\partial u_1^m} \Big|_{u_0^m=u_0, u_1^m=u_1} > 0,$$

from which we conclude  $s_0 < 0$  and  $s_1 > 0$ .

#### References

- Aoki, K., 2010. Evolution of the social-learner-explorer strategy in an environmentally heterogeneous two-island model. *Evolution* 64, 573–581.
- Aoki, K., Lehmann, L., Feldman, M.W., 2011. Rates of cultural change and patterns of cultural accumulation in stochastic models of social transmission. *Theor. Popul. Biol.* 79, 192–202.
- Aoki, K., Nakahashi, W., 2008. Evolution of learning in subdivided populations that occupy environmentally heterogeneous sites. *Theor. Popul. Biol.* 74, 356–368.
- 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, 300–308.
- Arbilly, M., Motro, U., Feldman, M.W., Lotem, A., 2011. Evolution of social learning when high expected payoffs are associated with high risk of failure. *J. R. Soc. Interface* 8, 1604–1615.
- Barnard, C.J., Sibly, R.M., 1981. Producers and scroungers: a general model and its application to captive flocks of house sparrows. *Anim. Behav.* 29, 543–550.
- 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*. University of Chicago Press, Chicago, IL.
- Boyd, R., Richerson, P.J., 1995. Why does culture increase human adaptability? *Ethol. Sociobiol.* 16, 125–143.
- Boyd, R., Richerson, P.J., 1988. An evolutionary model of social learning: the effect of spatial and temporal variation. In: Zentall, T., Galef, B.G. (Eds.), *Social Learning*. Erlbaum, Hillsdale, NJ, pp. 29–48.
- Bush, R.R., Mosteller, F., 1955. In: Bush, R.R., Estes, W.K. (Eds.), *A Comparison of Eight Models*. In: *Studies in Mathematical Learning Theory*, Stanford University Press, Stanford, CA, pp. 293–307 (Chapter 16).
- Caswell, H., 2001. *Matrix Population Models*, second ed. Sinauer, Sunderland, MA.

- Chudek, M., Heller, S., Birch, S., Henrich, J., 2012. Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evol. Hum. Behav.* 33, 46–56.
- Claidière, N., Bowler, M., Whiten, A., 2012. Evidence for weak or linear conformity but not for hyper-conformity in an everyday social learning context. *PLoS One* 7 (2), 1–8.
- Dubois, F., Morand-Ferron, J., Giraldeau, L.-A., 2010. Learning in a game context: strategy choice by some keeps learning from evolving in others. *Proc. R. Soc. Lond. (Biol.)* 277, 3609–3616.
- Eliassen, S., Jørgensen, C., Mangel, M., Giske, J., 2009. Quantifying the adaptive value of learning in foraging behavior. *Amer. Nat.* 174, 478–489.
- Enquist, M., Eriksson, K., Ghirlanda, S., 2007. Critical social learning: a solution to Rogers's paradox of nonadaptive culture. *Am. Anthropol.* 109, 727–734.
- Eriksson, K., Coultas, J.C., 2009. Are people really conformist-biased? An empirical test and a new mathematical model. *J. Evol. Psych.* 7, 5–21.
- Eriksson, K., Enquist, M., Ghirlanda, S., 2007. Critical points in current theory of conformist social learning. *J. Evol. Psych.* 5, 67–87.
- Feldman, M.W., Aoki, K., Kumm, J., 1996. Individual versus social learning: evolutionary analysis in a fluctuating environment. *Anthropol. Sci.* 104, 209–213.
- Feldman, M.W., Cavalli-Sforza, L.L., 1984. Cultural and biological evolutionary processes II: gene-culture disequilibrium. *Proc. Natl. Acad. Sci.* 81, 1604–1607.
- Giraldeau, L.-A., Caraco, T., 2000. *Social Foraging Theory*. Princeton University Press, Princeton, NJ.
- Hanania, M.I., 1959. A generalization of the Bush–Mosteller model with some significance tests. *Psychometrika* 24 (1), 53–68.
- Henrich, J., 2004. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *Am. Antiquity* 69, 197–214.
- Henrich, J., Boyd, R., 1998. The evolution of conformist transmission and the emergence of between-group differences. *Evol. Hum. Behav.* 19 (4), 215–241.
- Henrich, J., Broesch, J., 2011. On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Phil. Trans. R. Soc. Lond. B Biol.* 366, 1139–1148.
- Izquierdo, L.R., Izquierdo, S.S., 2008. Dynamics of the Bush–Mosteller learning algorithm in 2–2 games. In: Weber, C., Elshaw, M., Mayer, N.M. (Eds.), *Reinforcement Learning: Theory and Applications*. I-Tech Education and Publishing, Vienna, Austria, pp. 199–224.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47 (2), 263–292.
- Katsnelson, E., Motro, U., Feldman, M.W., Lotem, A., 2011. Individual-learning ability predicts social-foraging strategy in house sparrows. *Proc. R. Soc. Lond. (Biol.)* 278, 582–589.
- Kendal, J., Giraldeau, L.-A., Laland, K., 2009. The evolution of social learning rules: payoff-biased and frequency-dependent biased transmission. *J. Theoret. Biol.* 260, 210–219.
- Kobayashi, Y., Aoki, K., 2012. Innovativeness, population size and cumulative cultural evolution. *Theor. Popul. Biol.* 82, 38–47.
- Lachlan, R.F., Janik, V.M., Slater, J.B., 2004. The evolution of conformity-enforcing behavior in cultural communication systems. *Anim. Behav.* 68, 561–570.
- Lehmann, L., Feldman, M.W., 2009. Coevolution of adaptive technology, maladaptive culture, and population size in a producer–scrounger game. *Proc. R. Soc. Lond. (Biol.)* 276, 3853–3862.
- 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 (11), 2356–2369.
- Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- McElreath, R., Strimling, P., 2008. When natural selection favors imitation from parents. *Curr. Anthropol.* 49, 307–316.
- Mesoudi, A., 2009. How cultural evolutionary theory can inform social psychology and vice versa. *Psychol. Rev.* 116 (4), 929–952.
- Mesoudi, A., 2011. Variable cultural acquisition costs constrain cumulative cultural evolution. *PLoS One* 6 (3), 1–10.
- Nakahashi, W., 2010. Evolution of learning capacities and learning levels. *Theor. Popul. Biol.* 78 (3), 211–224.
- Nakahashi, W., 2007. The evolution of conformist transmission in social learning when the environment changes periodically. *Theor. Popul. Biol.* 72, 52–66.
- Nakahashi, W., Wakano, J.Y., Henrich, J., 2012. Adaptive social learning strategies in temporally and spatially varying environments. *Hum. Nat.* 23, 386–418.
- Powell, A., Shennan, S., Thomas, M.G., 2009. Late pleistocene demography and the appearance of modern human behavior. *Science* 324, 1298–1301.
- Rendell, L., Fogarty, L., Laland, K.N., 2010. Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. *Evolution* 64 (2), 534–548.
- Rodriguez-Gironés, M.A., Vázquez, R.A., 1997. Density-dependent patch exploitation and acquisition of environmental information. *Theor. Popul. Biol.* 52 (1), 32–42.
- Rogers, A.R., 1988. Does biology constrain culture? *Am. Anthropol.* 90 (4), 819–831.
- Stephens, D.W., 1991. Change, regularity, and value in the evolution of animal learning. *Behav. Ecol.* 2 (1), 77–89.
- 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 (4), 486–497.
- Wakano, J.Y., Aoki, K., 2007. Do social learning and conformist bias coevolve? Henrich and Boyd revisited. *Theor. Popul. Biol.* 72 (4), 504–512.