# Evolution of learning strategies in changing environments

Bullinaria, John A.

[Link to publication on Research at Birmingham portal](#)

# Evolution of Learning Strategies in Changing Environments

John A. Bullinaria

School of Computer Science
University of Birmingham
Birmingham, B15 2TT, UK

*j.a.bullinaria@cs.bham.ac.uk*

**Abstract:** Learning is an important aspect of cognition that is crucial for the success of many species, and has been a factor involved in the evolution of distinct patterns of life history that depend on the environments in question. The extent to which different degrees of social and individual learning emerge follows from various species-dependent factors, such as the fidelity of information transmission between individuals, and that has previously been modelled in agent-based simulations with meme-based representations of learned knowledge and behaviours. A limitation of that previous work is that it was based on fixed environments, and it is known that different learning strategies will emerge depending on the variability of the environment. This paper will address that limitation by extending the existing modelling framework to allow the simulation of life history evolution and the emergence of appropriate learning strategies in changing environments.

# 1. Introduction

There is considerable evidence of environmental variability, such as climate change, driving the evolution of versatile behaviours (e.g., Potts, 1996, 1998, 2013; Grove 2011). In particular, the prevailing environment changes are known to affect the learning strategies that evolve (e.g., Rogers. 1988; Acerbi, & Parisi, 2006; McElreath & Strimling, 2008; Ehn & Laland, 2012). There are several potential approaches for studying this issue, including empirical studies of real populations, detailed grounded models of real populations, abstract mathematical models, and simplified agent-based simulations. Then within the realm of agent-based simulations are distinct approaches for representing the learned behaviours, such as in terms of neural network weights or as simplified sets of abstract memes. All these approaches have their own advantages and disadvantages, and comparing their results is important for exploring their limitations and arriving at reliable conclusions. An agent-based simulation framework based on abstract memes has previously been formulated (Bullinaria, 2017) and shown to provide a powerful approach for modeling the interaction of direct individual and social learning in general, and in the specific context of life history evolution, but that has so far only been developed and explored for static environments. The aims of this paper are to extend that framework to accommodate changing environments and to investigate how the form and degree of variability affects what emerges. In particular, it will explore the interaction between social and individual learning, the optimal learning strategies that emerge, and how they can affect the resulting patterns of life history evolution.

The modeling framework here is deliberately simplified to allow computationally-feasible simulations, with all the environmental and behavioural details totally abstracted out, limited meme-sets, and small population sizes. Nevertheless, the results found previously for static environments (Bullinaria, 2017) indicate that it is sufficient to capture enough relevant details to provide useful models. Moreover, the approach is purposely kept very general so it has the important advantage of being easily applicable to modeling a diverse range of species with only a few changes to parameter settings being required. To establish whether the same general framework is also sufficient for modelling species in changing environments, a series of simulations need to be carried out to establish suitable implementational details and parameter value ranges for the models. In parallel with that, simulation results need to be generated to test its correspondence with known biological populations, existing mathematical models and more-grounded simulations, and reasonable intuitions about what should occur. The remainder of this paper will present the results of doing that.

The interaction of learning and evolution has already been extensively studied in the past, and it is well known why at least some lifetime learning is generally required, even when the environment is stable and perfect genetic assimilation of previously learned behaviour is possible. For example, most species grow after birth and precise good adult behaviours will therefore often be inappropriate for new-born individuals, in the same way that learned adult neural connection strengths will generally be sub-optimal

for new-borns (Bullinaria, 2003). Similarly, optimal innate behaviours for a particular environment will be inappropriate for spatially or temporally changing environments (Crispo, 2007). What is not so clear is which forms of learning are most appropriate for particular forms of environmental changes. That is one of the key questions this paper needs to address, specifically in relation to aspects of behaviour for which genetic acquisition is either inappropriate or impossible. In principle, direct individual learning based on the current environment should work well under all conditions, but that can be costly in terms of time and effort, and may be unreliable or even deadly in some circumstances, so various forms of social learning are often more effective and have consequently evolved for some species when sufficient transmission fidelity is possible (e.g., Boyd & Richerson, 1985; Tomasello, 1999; Kameda & Nakanishi, 2002; Rendell et al., 2010).

The issue of different learning strategies emerging in environments which change significantly over individual lifetimes, or over generational timescales, has previously been explored in the simplified mathematical models of Feldman et al. (1996), Wakano and Aoki (2006) and Ehn and Laland (2012), the neural network controlled agent simulations of Acerbi and Parisi (2006), and in the individual-based stochastic models of Whitehead (2007) and Whitehead and Richerson (2009), and different time-weighted learning strategies are known to be superior in such situations (Rendell et al., 2010). One of the key advantages of agent-based simulation approaches is that such strategies can be encoded genetically and optimized by evolution along with everything else. Moreover, factors that vary with location, such as environmental conditions, are also easily simulated in this approach, and incorporating them into the models can allow the emergence of cultural differences between groups to be studied in more detail (Henrich & Boyd, 1998). Modelling all the relevant issues clearly requires a well-defined, flexible and tested simulation framework that can accommodate the various interacting factors and trade-offs as accurately as possible in line with particular realistic scenarios, and facilitate reliable and reproducible comparisons with minimal confounding factors. Formulating and testing such a modeling framework based on agents and memes was presented in an earlier paper (Bullinaria, 2017) and the same approach will be followed here to accommodate a range of forms of environmental variability.

In the next section the general meme- and agent-based modeling framework of Bullinaria (2017) will be outlined and extended to incorporate changing environments. The following section then presents a series of simulation results obtained using it that explore how environmental variability affects the trade-off between social and individual learning, the resulting probabilities of population collapse and strategies for avoiding it, the potential social learning strategies and how they may evolve, and some representative life history evolution issues. The paper ends with some discussion and conclusions.

## 2. Simulation Framework

The underlying agent and meme based simulation framework to be explored here for changing environments has already been described and tested in some detail for static environments (Bullinaria,

2017), so it should be sufficient to just outline the resulting framework here. The general idea is to maintain a population of individual agents that are each specified by a set of relevant innate parameters (such as brain size, individual learning rate, social learning rate, etc.) and have each acquired some subset of the available "memes" that represent the behaviours they have adopted for surviving in their environment. The overall performance of each individual is the sum of the "performance contributions" associated with their set of learned memes. Of course, mental representations are much more complex than small discrete sets of memes that can be copied directly, and memes will not really contribute to performance or fitness in a simple additive manner, but this simplification is a useful starting point for simulations that simply need a convenient way to "keep score" of how individually and socially learned information and behaviours are affecting individual performances (Aunger, 2002).

## 2.1 Meme-based agent simulations

The crucial underlying assumption here is that all the relevant knowledge and behaviours can be sufficiently well approximated by an overall set of $M$ memes $\{m_j : j = 1,\dots,M\}$ and each individual $i$ at each stage of their life has acquired and stored some subset of them in their brain that has maximum capacity $B_i$ memes. The models are deliberately kept very abstract so they can represent as many different species and real life scenarios as possible (e.g., Reader & Biro, 2010). Consequently, there is no need to specify at this stage exactly what the memes represent and how they are originally created from direct individual learning, nor worry about the details of the meme transfer processes such as imitation, emulation and teaching, and how those processes may have evolved (e.g., Thornton & Raihani, 2008; Marriott et al., 2010; Kline, 2015). Individual are born with a baseline performance level of zero (corresponding to their innate abilities), and throughout their life acquire memes that either improve or worsen their performance. Another simplification is that all memes are taken to have equal complexity and imitability, though to allow suitably realistic simulations it is important that different memes contribute unequally to the overall performance of the individuals possessing them. Having the meme performance contributions uniformly distributed in the range $[-1, +1]$ does that in an easy-to-implement manner with "good memes" contributing positively to performance, while corresponding "bad memes" contribute negatively and represent information that is incorrect or contradictory in some way to a good meme. This is consistent with the idea of an opposing pair of memes existing for each behavioural context, with performance contributions in line with the importance of that context, and acting as *cognitive attractors* in the way that has previously been studied in the models of Henrich and Boyd (2002). Then reconciling incompatible ideas and correcting harmful information can be dealt with simply by having corresponding pairs of memes of equal and opposite contribution (i.e., positive and negative) cancel each other out if they are both acquired by the same individual.

A key advantage of this approach is that it is possible to model the rather different *individual* and *social* learning processes within a single coherent framework using the same abstract memes. The

simplest form of social learning allows each individual $i$ to "copy" up to $\alpha_i \phi B_i$ memes each time-step from other individuals, where $\phi$ is a parameter that specifies the maximum rate at which memes can be copied by the given species, $\alpha_i$ is the evolved innate social learning rate in the range [0, 1], and the brain capacity $B_i$ is included to establish a scale independence to the learning rate. Similarly, individual learning has each individual $i$ learn $\delta_i \psi B_i$ random good memes directly from their environment each time-step, where $\delta_i$ is an evolvable direct learning rate and $\psi$ is a measure of ease of direct learning for the given species. Both forms of learning each involve a further parameter that potentially limits the associated learning rate that evolves.

For social learning there is a species dependent transmission fidelity $f$ such that a fraction $1–f$ of copied good memes are copied incorrectly. As previously discussed by Bullinaria (2017), it is not obvious how best to implement such copying errors in the current framework. A fully realistic set of memes would offer a whole range of behaviours and performance contributions for each context, and it would be natural to have a poorly copied good meme become a random poorer quality meme associated with the same context, with performance contribution reduced in proportion to the degree of copying error. However, that makes less sense for the simplified good/bad meme pairs adopted here for each behavioural context, in which case it is more natural to have a poorly copied good meme become the corresponding bad meme, and a poorly copied bad meme remain the bad meme. Bullinaria (2017) tested the various possibilities and confirmed that the best approach is to have each copied good meme with performance contribution $x$ switch to the meme with contribution $–x$ with probability $1–f$.

The corresponding issue for direct learning is that learning too quickly can also introduce errors, and a straightforward meme-based implementation of that, which is consistent with the pattern of errors implemented for the social learning, involves having a probability $\rho \delta_i$ of learning the bad meme with contribution $–x$ rather than the chosen good meme with contribution $x$, where $\rho$ is a species dependent measure of individual learning difficulty (Bullinaria, 2017). This individual learning cost is distinct from the cost related to the effort involved in carrying out the individual learning. For the purposes of the initial simulations of this paper, it will be assumed that there is roughly equal effort involved in both social and individual learning, so neither needs to be included explicitly in the models, though this is something that will need to be addressed more carefully in more realistic models in the future. However, the relative advantages of social and individual learning for particular values of $\rho$ and $1–f$, and the associated interaction with environmental change, are key factors to be explored in this paper.

Finally, an important extension that renders the simulations more realistic, that is adopted in all the simulations of this paper, is to introduce a degree of stochasticity by drawing the number of learned memes uniformly from the ranges $[0, 2\alpha_i \phi B_i]$ or $[0, 2\delta_i \psi B_i]$ so the numbers only average out to the previously specified $\alpha_i \phi B_i$ or $\delta_i \psi B_i$. This reflects the various kinds of variability found in real environments and real social structures, has the effect of maintaining more diverse populations, and tends to result in more robust evolved traits.

Next the interaction and evolution of populations of these meme-based agents need to be formulated. As in real life, the weakest individuals are more likely to die at any point in their lifetimes, either from a direct fight with another individual, or due to being an easier target for predators or less likely to find enough food to survive. A standard *tournament selection* approach (Eiben & Smith, 2015) is sufficient to model all such deaths, with an appropriate number of random pairs of individuals competing each simulated year by having their performance compared, and only the winners surviving. For the simplest simulations, such as in this paper, these performance comparisons will be based purely on what the individuals have learned so far, but they could also involve performance adjustments that depend on factors such as age or how many children they are looking after. Older individuals are also prone to dying of old age, and these deaths are taken to occur with a constant probability each year after individuals have exceeded their species' natural lifespan, though the framework is general enough to allow more realistic old-age death rates, potentially involving other factors including performance. The required number of children to maintain the fixed population size that the given environment can support are each produced from their two parents using the standard evolutionary operators of crossover and mutation, with the parents chosen each simulated year from the eligible individuals, again selected by taking the winners of performance comparisons of randomly chosen pairs of individuals (Eiben & Smith, 2015). Crossover corresponds to having each child's innate parameter values chosen randomly from the corresponding ranges spanned by their two parents, and the mutations are random constants added to each inherited parameter to allow a significant chance of its value falling outside the parental range.

The power and reliability of this simplified meme-based approach has previously been demonstrated (Bullinaria, 2017) by showing how it leads to: mimetic transition results in line with those of the earlier simulations of Higgs (2000); qualitatively the same life history evolution effects as in the full neural network-based simulations of Bullinaria (2009); and consistency with various special-case results that have been found using simplified mathematical models (e.g., Boyd & Richerson, 1985; Rogers, 1988; Ehn & Laland, 2012).

In practical terms, running the simulations here simply involves maintaining a list of parameter values and acquired memes for each individual in the population, updating them at each time-step, and over-writing the parameters of any individuals that die with those of the children that replace them. The time-step size obviously needs to be set so the life history of the modeled species is simulated at a sufficient level of granularity; for example, one update per simulated year is usually appropriate for humans. Similarly, particular distributions of other details such as litter sizes, delays between offspring, death rates and mutations also need to set in line with the species being modeled. Updating each individual just means incrementing their age and adding the required number of newly acquired memes to their list of known memes in line with their individual and social learning rates. Then some individuals die due to competition or old age, and are replaced by children with age zero, no known memes, and innate parameters (learning rates, etc.) determined by crossover and mutation from their selected parents. Each

simulation run starts with an appropriate un-biased random initial population, and needs to be executed for more than enough simulated years that the population averages of the various evolving traits, parameters, acquired memes sets, and resulting levels of performance stabilize. A few very long test runs need to be performed to establish a suitable simulation length, and then many runs of a sufficient set number of simulated years are carried out so that reliable variances across and during runs can be established. A complication is that for some levels of environmental variability it becomes increasingly common for populations to collapse leaving no ongoing stable state, but in such cases a fairly stable period normally exists earlier in the run. All the simulations presented in this paper were stopped after 10 million simulated years, or at population collapse if it occurred before then, which was found to be plenty to ensure a stable period was present in the middle portion of the run.

To ease comparisons with the existing simulation results for stable environments (Bullinaria, 2017), the same parameter values are used throughout the remainder of this paper: $M = 500$, $B_i = 100$, $\phi = 0.1$, $\psi = 0.1$ and $\rho = 0.04$, with fixed population size of 200 individuals, 10% of the population dying each year due to unsuccessful competition, and 20% of individuals aged over 60 years dying each year of old age. The initial populations have zero social learning rates $\alpha_i$ to avoid any bias towards social learning, and small random individual learning rates $\delta_i$ drawn from the uniform distribution $[0, 0.1]$ to provide an initial set of memes that any social learning can work with. This combination of parameters was identified by Bullinaria (2017) as resulting in populations with human-like age distributions, and timescales for the individual and social learning that are similar to those of human children, with simulation run-times that are short enough to allow a wide range of experiments to be carried out. The social learning will normally involve copying memes from the best other individuals, again chosen by performance-based tournament selection from random pairs of individuals, but that strategy choice is one of the factors to be explored later. For each social learning event, as many new memes as possible are copied in random order from the selected individual before moving on to the next selected individual, and this continues until the required number of new memes have been acquired, or the brain capacity is reached. Since attempting to copy from the whole population each year is unrealistic, and to terminate the copying process when there are not enough new memes in the whole population, the number of potential copied individuals allowed each year for each copier is limited to half the population size.

Finally, to determine the variations arising from the stochastic nature of the simulations, all the results presented are averaged over at least 20 independent runs (using different random number seeds). When there is a well-defined outcome for each run, such as the time to population collapse, the relevant graphs have error bars representing the standard deviations over the 20 runs. However, the variation of evolving traits, parameters and performances within runs is often considerably greater than the variation of run averages across runs, even when the initial settling-down periods are excluded. This means that taking only a single population to represent each run can be misleading, particularly when there are large random environmental changes. Moreover, in runs that end with a population collapse, the evolved values and

meme distributions near the end of the run can be rather different from those that exist during the relatively stable bulk of the run. The simulation outcomes are therefore most reliably presented across all cases by considering both the variation during the middle of each run *and* across the whole set of runs. Consequently, in this paper, all the results are shown as population means computed from 30 equally spaced points in the middle 20% of the 20 independent runs, and the error bars represent the associated standard deviations. That leads to a high probability of sampling representative successful populations separated by large numbers of generations (typically 100 to 1000).

## 2.2 Changing environments

The specification of the simulation framework has so far followed that of Bullinaria (2017) which assumed a static environment, and that will need to be modified to accommodate changing environments. The crucial factor is that the performance contributions of particular memes will depend on the state of the environment, and they will need updating if the environment varies over time or location. One of the advantages of the abstract meme-based framework here is that the details of the environment and associated behaviours (i.e. memes) never need to be specified. All that is needed to incorporate environmental change is a simple implementation of the idea that the goodness (i.e. performance contribution) of each meme will depend on the environment. Moreover, for selection purposes, only the relative contributions of the memes are important and need to be tracked. Since the simulations involve a simplified set of $M$ memes with performance contributions spread uniformly through the range $[-1, +1]$, the memes at each time-step can be conveniently numbered such that the meme with relative position $m$ has performance contribution $1-2m/M$. Then representing an environmental change simply means renumbering the set of memes held by the individuals in line with the new ranking of the $M$ memes. For example, a meme that becomes relatively less useful after an environmental change would swap its ranking with a meme that becomes relatively more useful. So there is no need to keep track of what the memes are, nor how they depend on the environment – all one needs to do is apply a simple permutation of the meme order to represent each environmental change. It is the simplicity of this process that renders the meme-based approach used here so powerful. Obviously, this approach is a gross simplification of what happens in real populations in real changing environments, and one of the key aims of this paper is to establish whether it captures enough of the crucial details to provide reliable simulations.

Small environmental changes will correspond to small changes to the meme order, perhaps a single pair of adjacent memes being swapped. A massive environmental change would be something closer to a totally random reordering. The previously learned behaviours of individuals do not change with the environment, just their performance rankings, so that is all the simulations need to keep track of. Fully realistic environmental changes would need an enormous number of parameters to represent them, but for the initial simulations here, three parameters will suffice: one to specify the magnitude of individual changes (i.e., the range of meme positions over which random position swaps take place), one to specify

the frequency of change (i.e., how many years there are between each environment change), and one to specify the overall size of the change (i.e., the number of individual swaps at each change). Larger values for any of these parameters represent greater environmental changes, but it remains to be tested how each of the three change parameters affect the details of what emerges in the simulations. It would be fairly straightforward to add more realism and stochasticity to the environmental changes by allowing each of the three change parameters $x$ to be replaced at each stage by one chosen randomly from the range $[0, 2x]$, but, to keep the results analysis simple, that will not be done for the simulations of this paper.

A gradual environmental change might correspond to a single swap every ten generations of one random pair of memes adjacent in the ranking order. Less gradual changes could be implemented by having such swaps occur more frequently, or by having more such swaps at each change, or by swapping a random meme with a random other meme within a larger window of a particular size around it. The most drastic changes would involve many swaps of totally randomly chosen memes, effectively with a magnitude or window size $M$. An appreciation of what such environment changes mean in practice can be achieved by plotting the correlation of the resulting meme order with the original meme order against the total number of swaps of each type. This is shown in Figure 1 for gradual changes (swaps of adjacent memes, i.e. magnitude 1), medium-sized changes (swaps within a window of +/- 30, i.e. magnitude 30), and drastic changes (swaps of totally random pairs, i.e. magnitude $M = 500$).

The main aims of this paper are to investigate how this abstract representation of environment change affects the results previously found for static environments (Bullinaria, 2017), and how consistent the new results are with what is already known about the effect of environment change on learning by real species and in other models. The next section presents simulation results that explore each of the key issues.

## 3. Consequences of Environmental Variability

There are several important factors here that are likely to have strong dependences on the type and degree of environmental variability. Foremost is the potential trade-off between social learning and direct individual learning, and the issue of how that trade-off depends on the transmission fidelity of the social learning. What happens will depend on the social learning strategies used and how those strategies evolve, and on the potential presence of competing species and other causes of population collapse and species extinction. Then all those factors will affect the broader patterns of life history evolution that emerge. Each of these issues will now be addressed in turn.

### 3.1 Trade-off between individual and social learning

As already noted, simulations of biological populations require simplifications and approximations to render them computationally feasible and easily analyzable. One such issue here is that, in reality, particular behaviours are not learned entirely by totally separated individual and social learning – a skill

initially acquired by direct learning may be improved by social learning (e.g., athletes have coaches), or a socially learned skill may be improved by direct experience (e.g., practice makes perfect), and the interaction will depend on the quality or fidelity of the social learning. In principle, behaviours may be broken down into components that are small enough to be leaned entirely as single "memes" acquired by a particular form of learning, but the complexities of this matter will not be addressed here. For an initial exploration, it will be sufficient to begin by only simulating the evolution of pure perfect-fidelity social learning in association with the simplest form of direct individual learning. There may be a trade-off between them, but an important feature of the models here is that there is always potential for the social and individual learning to co-exist and co-evolve.

The crucial learning issue here is that during periods of significant environmental change, individual learning will be able to take the changes into account, whereas pure social learning will result in inappropriate behaviours that were developed for conditions that no longer exist. One might therefore expect the emergence of more individual learning and less social learning in conditions of increased environmental change. However, if individual learning is able to correct inappropriate socially learned behaviours, it is likely to be better overall to maintain the social learning rates along with the increased individual learning. This intuition is broadly confirmed by explicit simulations with the environmental change represented by one meme swap of a particular magnitude per simulated year as seen in Figure 2. The social learning rates $\alpha \in [0, 1]$ are relatively unaffected and remain near the maximum allowed, while much higher individual learning rates $\delta$ (that are only indirectly limited by the parameter $\rho$) emerge for large magnitudes of environmental change, and low (but slowly increasing) rates persist when the changes are small to generate useful memes for the social learning to spread. As would be expected, the average performance levels fall with increasing magnitudes of change, and the numbers of good and bad memes both rise when a large amount of individual learning is taking place. What might be less expected is the relatively sudden increase in individual learning rates when the change magnitudes exceed around 150. The abstract nature of the models and simplified representation of the environmental changes make it difficult to map this transition to real world systems, and it is not even clear that such sharp transitions will arise for more realistic patterns of change, but it seems likely that in any system there will be a particular level of change at which individual learning starts to increase significantly.

For small and intermediate levels of environmental change (i.e., magnitudes up to around 150) the populations are able to survive for long periods with small amounts of individual learning sufficient to track the changes, so natural selection does not drive up those learning rates to high levels. However, this strategy allows the population performance levels to suddenly collapse when certain unfortunate patterns of change occur. If the simulations are allowed to continue beyond that point, the rate of social learning does drop and the low individual learning rates are sufficient to allow the population performance to recover, but eventually the cycle repeats. Figure 3 shows a typical simulation run exhibiting this pattern for an intermediate (magnitude 30) level of change. It can also be seen how the individual learning rates

10

are higher during the initial stages of evolution (usually the first half to one million years) – this quickly generates the memes later spread by social learning, but minimizes the costs of individual learning later. For higher levels of change (i.e., magnitudes above 150), increasing large individual learning rates are needed to successfully track the changes.

### 3.2  Population collapse and species extinction

A crucial feature of real populations missing from the simulations so far is competition with other species or different groups within the same species, and that makes the cycles seen in Figure 3 unrealistic. If there is only one species or group, they may be able to survive with very low performance levels while they learn to adapt to the environmental changes that caused those levels to drop. However, if they need to compete with other species or groups that can adapt more effectively to environmental change, they are likely to be driven to extinction or extirpation before they have had a chance to adapt themselves (e.g., Diamond, 2011). There will often also be minimum performance levels required to survive even when there are no competing groups or species. Whitehead and Richerson (2009) have already presented stochastic models that show how reliance on social learning can lead to population collapse in realistically variable environments. The ideal way to explore this within the current framework would be to run the simulations with multiple species or groups, or minimum performance levels for cooperating sub-populations, but a more computationally feasible way to proceed is to simulate all such effects by simply having a population become extinct if its average performance falls below some minimum level required for survival. As previously discussed in some detail by Whitehead and Richerson (2009), there are numerous difficulties in trying to tie such simplified models to the kinds of collapses known to have arisen in real societies, such as those described by Diamond (2011), but the simulations here are sufficient to confirm how cultural behaviours can interact with environmental change and lead to similar kinds of collapse, and provide a basis for more realistic models in the future.

As already seen in Figure 3, the collapses tend to be rather sudden, so the precise cutoff point taken to signify extinction actually makes little difference and a convenient average performance level of 20 can be used for it. In that case, the average number of years before population collapse for different magnitudes of environmental change is as plotted in the left graph of Figure 4, with the simulations stopped after 10 million years if there has not been a collapse by then. It is seen that the extinctions occur more quickly as the magnitude of change increases until the point is reached (at magnitudes around 150) where the changes are sufficient to drive the evolution of the much larger individual learning rates seen in Figure 2. The potential presence of such extinctions complicate the computation of evolved or learned properties because they are likely to be changing drastically near the end of any simulations which end in collapse. Since it makes no sense to include collapsing populations, or the initial stages of evolution, when averaging results for the evolved populations during their stable periods, all such averages will, as noted earlier, be computed from the middle 20% of the relevant runs.

Another complication here is that it is not obvious how the population collapses will depend on the population size, which has been set at an unrealistically small 200 in the simulations so far. If the whole population consisted of five non-interacting sub-populations of 200, each in their own randomly changing environment, each would evolve and potentially collapse in the same manner as the simulations seen so far, so the time to total population collapse would be the maximum of the five individual collapse times, and the average of those maximums would inevitably be more than the average of the sub-populations. However, if those five sub-populations were allowed to co-exist and interact in a single environment, it is possible that any risky strategies would propagate throughout the whole population of 1000, and any unfortunate environmental changes could then collapse the whole larger population just as easily as a smaller population. On the other hand, a larger population may well include more diversity, even if it is fully interacting, and that could mean more robustness to change and result in longer times to population collapse. Running further simulations to test this is straightforward, and Figure 4 shows how the times to population collapse differ for population sizes of 200, 5×200 and 1000. The left graph confirms the extent of the increased times for five non-interacting sub-populations, and shows that a five-fold increase in population size does lead to slight increases in the times to collapse for intermediate magnitudes of environmental change, but the differences are small compared to the variances of those times. The nature of the collapse time distributions and their degree of overlap for the three population sizes is shown in the right graph of Figure 4 for magnitude 30 environmental changes (computed from 1000 independent runs of populations size 200, and 300 runs of size 1000). The difference in the means for populations sizes 200 and 1000 is highly statistically significant ($t$ test $p < 10^{-16}$), but since the qualitative patterns of results have been found to be independent of the population size, the remainder of this paper will continue with the more computationally feasible population size of 200.

### 3.3 Adaptive learning strategies

It is not clear how many real species have become extinct due to changing environments in a similar way to that found in the simulations here, but it is certainly possible to increase their chances of survival by adopting more versatile behaviours that allow their learning strategy to change when the population average performance drops to dangerously low levels. For example, by boosting the individual learning rates when population collapse appears imminent, it should be possible to avoid extinction. Incorporating such an adaptive learning strategy into the earlier simulations gives results that are qualitatively the same for a wide range of threshold and rate changes: population collapse is avoided in all cases. Figure 5 shows what happens in a typical run for magnitude 30 changes when the all the individual learning rates are each increased by a random amount in the range [0, 0.5] whenever the population average fitness drops below 20. The population performance quickly recovers from each crash, and the learning rates gradually evolve back to the lower values that existed before the crash. Higher thresholds and smaller learning rate changes lead to similar, but smoother, graphs. And similar outcomes also emerge if the

individual learning boost is an additional component, distinct from the innate learning rate, that returns to zero once the difficulties introduced by the environmental change have been solved. The benefit of similar kinds of strategic switching between social and individual learning has been studied and discussed previously (e.g., Boyd & Richerson, 1985; Kendal et al., 2005, Ehn & Laland, 2012), and is consistent with the variability selection hypothesis (Potts, 1996, 1998, 2013; Grove, 2011) that more versatile behaviours will be adopted in more variable environments.

That leads on to the question of how such behaviours might evolve. An entire population which has not already adopted an adaptive strategy will simply become extinct leaving nothing to evolve. More diverse populations may include suitably adaptive sub-populations that are able to avoid collapse whenever the unfortunate environmental changes occur, and that will leave the entire remaining population with that trait. Both of these cases can easily be simulated in the current framework with the expected results. That leaves the question of how is it possible for sub-populations to evolve or learn to recognize an imminent collapse and adapt appropriately before it is too late if they cannot survive without that ability. It could be that a simple random diversity in behaviours will suffice. Another possibility is that, in reality, rather than there being a single performance level that determines population collapse, there will be numerous activities that are each important, but not crucial, for group survival, and recognizing the need to adapt as one of them fails could be learned or evolved for future application to other activities. There are numerous variations along these lines that could be simulated in the current framework, but they would not add to the key result that when suitable adaptive learning strategies can be employed it is possible to avoid all the difficulties caused by environmental change. The remainder of this paper will look at what happens when this is not possible.

## 3.4 Dependence on type of environmental change

All the simulations presented so far have involved one meme swap per simulated year and only the effect of the magnitude of the swaps has been investigated. The consequences of real environmental change will clearly depend on the type of change, as has previously been explored in the stochastic models of Whitehead (2007) and Borenstein et al. (2008), and the type of change will normally depend on the environment in question (e.g., Steele, 1985). There are also likely to be different forms of variation if the changes are purely spatial, such as associated with movements away from the parental habitat, rather than purely temporal, such as associated with climate change, or some mixture of both (Boyd & Richerson, 1985). Here the general patterns can be investigated by looking at two extreme variations.

First, simply increasing the number of swaps per year has a similar effect to increasing the magnitude of the swaps. For example, Figure 6 shows the effect of multiple swaps of magnitude 30 on the population collapse times and learning rates, for comparison with the right-hand portions of the left graphs in Figures 2 and 4. This is consistent with the expectation that many random swaps of an intermediate magnitude should be roughly equivalent to a single random swap of a larger magnitude.

Less obvious is the effect of having the same number and magnitude of swaps, but bunched up in time rather than once per simulated year. This may be more realistic than the environmental changes already considered because there is evidence that the critical real-world environmental variation is made up of relatively rare relatively large changes (Halley, 1996; Whitehead & Richerson, 2009). Figure 7 shows the evolved learning rates for intermediate change represented by swaps of magnitude 30 and drastic change represented by swaps of magnitude 500. For the intermediate changes there are no significant changes to the times to extinction, learning rates or performances as the frequency of the changes varies from once per year to once per thousand years. Whereas for the drastic changes the individual learning rates begin to fall when the changes occur less frequently than once every 64 years. However, checking the details of individual simulations reveals that the summary results here may be slightly misleading in that infrequent environmental changes tend to produce cyclic population performance levels rather than the type of population collapse seen in Figure 3, as shown in Figure 8 for the case of changes every 512 years. In these cases it will be important to consider more carefully what counts as an extinction and an average learning rate, particularly since the individual learning rates do not reach their "normal" range of values till long after the extinction is likely to have occurred. This complication will not be pursued further in this paper, but it is clearly going to require a more careful species-specific treatment when attempting to model evolution and learning in these types of environment.

### 3.5 Interaction with transmission fidelity

In static environments, an important factor affecting the trade-off between individual and social learning is the transmission fidelity of the social learning, i.e. the rate at which errors are introduced by the relevant social behaviour or knowledge transfer process, and that will clearly be species-specific. The pattern of simulation results is as found previously (Bullinaria, 2017) with the evolution of mostly social learning when the transmission fidelity is high, and mostly direct individual learning when it is low, as shown in the top-left graph of Figure 9. This pattern is shifted in cases of changing environment as seen in the other plots of Figure 9. For drastic changes (of magnitude 500), the individual learning rates are high for all fidelities, which is consistent with that seen previously in Figure 2 for the perfect-fidelity case. The bottom-left graph of Figure 9 shows that high rate only increases slightly for lower fidelities, and the social learning rate decreases for lower fidelities at an even faster rate than found in stable environments. For intermediate environmental changes (of magnitude 30) the evolved learning rates are remarkably similar to those of the stable environment case for all fidelities, as seen in the top-right graph of Figure 9, which is again consistent with the perfect-fidelity results of Figure 2. The big difference in this case is the decrease in times to population collapse for the high fidelity cases which are dominated by social learning, as seen in the bottom-right graph of Figure 9. Since social learning only plays a significant role when its transmission fidelities are high, perfect fidelity will be assumed for the remaining simulations of this paper that are primarily concerned with the social learning strategies.

14

### 3.6 Dependence on fixed social learning strategy

In regimes where social learning plays a significant role, success generally relies on more sophisticated strategies than simply imitating random other individuals at a constant rate. It is certainly known to make a big difference which other individuals are learned from, particularly in rapidly changing environments (Rendell et al., 2010). For example, learning from ones own parents seems to be the best strategy in the current framework for static environments (Bullinaria, 2017), but other types of models with changing environments have found different optimal strategies (e.g., Acerbi & Parisi, 2006; McElreath & Strimling, 2008). Fortunately, exploring the effect of learning strategy in changing environments is straightforward in the simulation framework of this paper. All the earlier simulations have had the social learning strategy involve copying the best of the other individuals using tournament selection, which has been known for some time to be a good strategy in general (Miller & Dollard, 1941). Two extreme alternative selection approaches should be sufficient to illustrate how varying that strategy affects the outcome, namely copying random other individuals and having individuals only copy their own parents. Intermediate strategies, such as coping parents first and then coping random other individuals, have been found to lead to intermediate outcomes.

A different kind of social learning strategy, that does not involve copying from particular selected individuals, involves "conformist transmission" in which individuals tend to adopt the behaviours that are already most common in the population (Boyd & Richerson, 1985; Henrich & Boyd, 1998; Whitehead & Richerson, 2009; Muthukrishna, Morgan & Henrich, 2016). On its own, this would lead to the most commonly held memes becoming the only held memes, but in general it will operate in conjunction with individual learning and other forms of social learning and act as a mechanism for establishing and maintaining cultural norms. Past models have already shown this strategy to be advantageous in environments with only small temporal changes (Henrich & Boyd, 1998), to cause population collapse in larger-magnitude temporally changing environments (Whitehead & Richerson, 2009), and to result in significant between-group differences in spatially changing environments (Boyd & Richerson, 1985; Henrich & Boyd, 1998). A full investigation of these issues in the current meme-based framework will have to be left for the future, but a simple extreme version of conformist learning has been implemented to show the general pattern of results that can emerge. In this case, each social learning event involves identifying and copying the relevant number of the most commonly held memes in the population that have not already been acquired by the given learning individual.

The top-left graph of Figure 10 shows how the average performance levels vary with magnitude of environmental variation for the four social learning strategies. Only imitating parents leads to the best performance levels when the change magnitudes are zero or very low, which is consistent with parents needing good memes to compete to survive and reproduce, as was found in previous simulations for non-changing environments (Bullinaria, 2017). However, that advantage diminishes with increasing

magnitudes of variation, though not in a clear monotonic manner. Imitating the best other individuals does lead to a monotonic decrease in performance with magnitude of variation, which is what one would expect given the nature of the variation, but for low intermediate magnitude changes it results in better performance levels than only imitating parents. This is broadly consistent with the results McElreath and Strimling (2008) obtained using a rather different modelling approach. Conformist learning leads to a non-monotonic pattern of performance similar to only imitating parents, and, contrary to the earlier models, does not appear to be the best strategy for low magnitudes of environmental change. However, the combination of high individual learning rates combined with the error-correcting effect of conformist social learning makes this the best overall strategy in situations of large environmental change.

Imitating random other individuals leads to the worst performance levels across most magnitudes of environmental change, but that does not necessarily mean it is the worst strategy. The top-right graph of Figure 10 shows the average times to population collapse. The imitate-parents strategy follows a similar pattern to that seen before for imitating the best others, but with an even more pronounced fall-off in times to collapse until increased individual learning rates becomes a profitable strategy. Imitating random others results in very rare population collapses for all change magnitudes, so that is actually probably the best strategy for intermediate levels of environmental change, at least in the sense of providing the most reliable continuation of the species, though it needs to be remembered that the social learning strategies always operate in conjunction with associated evolved learning rates, and understanding any differences may be complicated by rather different rates of individual learning.

The evolved social and individual learning rates are shown in the bottom-left and bottom-right graphs of Figure 10. They all remain level across small and intermediate magnitudes of change, and for large magnitudes all strategies show an increase in individual learning rates with an associated slight dip in social learning rates. The biggest outcome difference between strategies is the evolution of high individual learning rates for all magnitudes of change when imitating random others, which is what allows that approach to avoid population collapses. Perhaps the least expected result is that, unlike for the selection-based strategies, the times to population collapse for conformist learning begin to rise before the big increase in individual learning rates which starts at the same environmental change magnitude independent of the social learning strategy. To understand that and the other differences in outcomes across the social learning strategies, it is necessary to look in more detail at what distributions of memes are being learned by the populations in each case. That is the topic of the next section.

### 3.7 Distributions of the learned memes

The plots of the performances and times to population collapse in Figure 10 are informative, but they hide what is actually being learned in each case. A big advantage of the simulation framework here is that it is straightforward to keep track of what memes have been acquired by the populations at each point in time, and that makes it possible to investigate how the associated distributions vary with environmental change

magnitude and social learning strategy. The various interactions are rather complex, not least due to the effects of variations in the associated evolved individual learning rates, but some of the important features of the meme distributions and their key differences are illustrated in Figure 11.

The top-left graph presents the average number of different memes that are known by the whole population during the stable portions of the relevant simulations, and their dependency on the social learning strategy and environmental change magnitude. For all four strategies, there is a clear increase in diversity for larger magnitudes of change. Not surprisingly, random copying leads to much greater diversity than the other three strategies. The only-copy-parents and conformist strategies have similar outcomes, and for small magnitudes of change their numbers of known memes are close to the brain capacity (set at 100 in all the simulations here) meaning that all individuals are acquiring essentially the same set of memes. That is to be expected for the conformist strategy, but it is not so obvious that only copying parents would have a similar outcome. The diversity emerging with the copy-the-best strategy lies, as expected, in between the two extremes.

The detailed meme distributions underlying these averages vary significantly, and ultimately they are responsible for the different potentials for population collapse in each case. The bottom six graphs of Figure 11 each show the number of individuals in the population that know particular memes at a typical point in time well away from initialization and population collapse. To aid their counting, the memes are sorted in order of popularity for the whole meme-set together (All) and separately for each of the four performance contribution quartiles (Best, Good, Bad, Worst). For all cases, the selection pressures present are seen to be sufficient to remove any previously popular memes that have been turned bad by an environmental change, and to lead to the better memes for the current environment becoming more widely known, though the speed at which this happens depends on the social learning strategy employed. Comparing the strategies for intermediate (magnitude 60) environmental change gives a good idea of the differences generally found. For the random-copying strategy, all the good memes are roughly equally well known which means there must be wide diversity across the population. The copy-the-best strategy leads to the best quartile memes being known more than the other good memes, meaning there is less diversity and better performances, with almost all the memes in the best quartile known by at least one-third of the population, but there is still a wide variation between individuals with few memes known by more than half the population. The only-copy-parents strategy leads to even less overall diversity, with about half of the best memes known by more than half of the population. Finally, the conformist strategy has the lowest diversity, with the most popular ~97 memes known by the large majority of the population, but still there are enough individual differences for the selection pressures to act.

The problem of population collapse typically arises when there is insufficient diversity or individual learning to replace bad memes by good memes quickly enough when the environment changes. The top-centre graph of Figure 11 shows the full time-course of a typical simulation for magnitude 60 changes with copy-the-best strategy. There is a fairly long stable period where the population copes with the

environmental changes, followed by a relatively fast collapse bought on by some unfortunate combination of events. The top-right graph shows how the distribution of known memes contracts during the collapse in time-steps of 5000 years, starting from a similar distribution to that seen in the corresponding graph below and ending with the whole population failing to reach their brain capacity and knowing the same small set of memes. The last few time-steps here mimic the process of collapse in the case of conformist social learning with low individual learning. That leaves the question of why conformist learning is less prone to population collapse with magnitude 60 changes compared to magnitude 30 changes, even though there is very little difference in the individual learning rates. Comparing their meme distributions in Figure 11 reveals two obvious differences: for magnitude 60 there are relatively large numbers of individuals still carrying bad memes, and there is a relatively long tail of memes that are only known by a few individuals, so it is presumably this additional diversity that results in the increased robustness against collapse. When large magnitudes of change cause the adoption of high individual learning rates, the distribution tails thicken in all cases, leading to distributions similar to the random strategy case in Figure 11 for all the non-conformist strategies, though there remains a clear conformist component with the conformist strategy as seen in the bottom-right graph of Figure 11.

This brief overview of the meme distributions provides an idea of what can be done with the meme-based simulation framework, but there clearly remains much scope for more detailed future investigations including how the meme distributions vary with other social learning strategies, different distributions of environmental change, and other factors such as social learning fidelity.

## 3.8 Evolution of social learning strategies

Given how different outcomes emerge for the various social learning strategies, the obvious next question is what happens if the strategy itself is allowed to evolve by natural selection. In principle, that can easily be answered by incorporating further evolvable parameters into the existing simulations to represent the social learning strategy. There are many ways that can be done, but to obtain the clearest results the above simulations are best augmented by a single evolving parameter that simply chooses between two extreme cases. Given the results in Figure 10, the choice likely to give the clearest results has that additional parameter specify the probability of imitating best individuals rather than random individuals at each social learning point. However, there is a complication here that was not present in the case of static environments studied previously by Bullinaria (2017). This issue is related to the complication noted in the discussion of Figure 8 and concerns the way the timescales of evolutionary change are effectively set by the mutation rates used. In static environments there are no population collapses, so the simulations can be run for long enough for all the evolved parameters to reach stable states whatever mutation rates are used. In changing environments, the evolution of the parameters may be terminated by a population collapse before their optimal values have been reached, so more care is needed in the choice of mutation rates. Moreover, the specification of the initial populations also affects the speed of evolution at the early

stages, so that needs to be set in line with the mutation rates, but for high enough mutation rates it has little effect on what eventually evolves.

The evolutionary selection pressures for the learning rates result in them reaching their final value ranges very quickly, even with the low mutation rates used in the earlier simulations, and well before any population collapses. Indeed, it is the evolved learning rates that lead to the population collapses, so changing the mutation rates there will have little effect on the outcomes. The social learning strategy parameter tends to be much slower to evolve, and for low mutation rates it may still be rising towards its optimal value when the population collapses. This is evident in the typical low-mutation-rate run shown in the left graph of Figure 12 for intermediate (magnitude 30) environmental changes, starting from initial populations with probabilities drawn randomly from the intermediate range [0.4, 0.6]. The same graph shows how this problem can be avoided by simply increasing the mutation rate by a factor of five. The right graph of Figure 12 presents the final evolved social learning strategy parameters for different magnitudes of environmental variability with the low and high mutation rates. For high enough mutation rates, natural selection leads to the adoption of an almost pure "learn from the best" strategy for all environmental change magnitudes, even though that leads to population collapse for intermediate levels of environmental change. Even when the mutation rates are lower, the tendency to evolve from the unbiased starting population average probability of 0.5 towards that pure strategy is still evident for all change magnitudes.

Having confirmed that evolution will select for copying the best rather than random individuals, the natural next question is what happens when the strategy parameter instead specifies the probability of copying only ones own parents rather than the best individuals. The performances here have high variances compared to the differences in the means, as shown in Figure 10, and that is reflected in the simulation outcomes shown in Figure 13. The evolved parameter values presented in the right graph only correspond to clear strategies (i.e., probabilities close to 0 or 1) for the high magnitudes of change that result in high levels of individual learning. For intermediate magnitude changes there is a tendency towards copying parents in line with the peak in performance seen in Figure 10. For lower levels of change, there is a tendency towards copying the best. Individual simulation runs are very variable and rarely have a fast or slow drift towards an optimal strategy like those seen in Figure 12. There are often repeated flips between strategies, as seen in the left graph of Figure 13, each presumably related to the stochastic aspects of what is happening at each stage. One thing that is quite consistent, however, is a rapid adoption of the copy-parents strategy early on, which later becomes less pronounced. This is to be expected since only copying parents has been shown previously (Bullinaria, 2017) to be the best way to limit the spread of bad memes early on in evolutionary runs with limited individual learning. Then that strategy will naturally dominate in runs that experience early population collapse, and that can account for the observed tendency in that direction for intermediate magnitudes of change. Such complications mean that predicting what might evolve by simply comparing the performances in simulations with different

19

fixed strategies is likely be unreliable – full evolutionary simulations are required.

As elsewhere in this paper, this section has merely illustrated the kinds of things that can be done within the proposed framework, leaving much more to be done in the future, such as evolution of all the other combinations of social learning strategies, or evolution of less extreme conformist strategies that use local groups rather than the whole population to determine which memes to acquire.

## 3.9 Life history evolution

There are clearly numerous wider aspects of life history evolution, that depend on learning and the stability of the environment, which can be explored using the simulation framework proposed here. One of the problems, though, is that many of the relevant traits in real populations are known to be correlated and to have co-evolved, and that often makes it difficult to untangle the various causes and effects. For example, the study of Walker et al. (2006) illustrates the issues involved in understanding the evolution of longer juvenile periods and increased brain sizes in primates, and how they relate to social complexity, while controlling for covariance with factors such as body size, life span and home range. Ultimately, of course, one would want to model all such traits together, along with their whole co-evolutionary process. However, sometimes it is also useful to know which traits would still evolve even if all the various "helper traits" were not present, and in that case running simulations has a clear advantage over studying real populations. One such isolated trait, that has already been explored in some detail for static environments (Bullinaria, 2009, 2017), should suffice to illustrate how life history evolution models would work more generally within the proposed framework, namely how learning affects the period of protection that the parents of some species offer their children. The underlying issue here is that protection periods can evolve to allow the avoidance of fast-but-risky learning strategies because it removes the pressure to learn quickly to compete with older individuals to survive. If there is no disadvantage to the protection, the protection periods quickly evolve to such high values that all individuals are protected till their parents die, which is clearly not biologically realistic. In practice there would at least be a cost to the protecting parents which prevents that. A cost to the children such as not being able to reproduce themselves while being protected also prevents that, and that trade-off alone is sufficient to result in more realistic protection periods emerging (Bullinaria, 2009, 2017).

The earlier simulations in this paper have already shown how environmental variation affects two crucial underlying factors here – the learning rates that evolve and the times to population collapse. The obvious way to proceed here would be to simply incorporate the protection period as another evolvable trait and rerun the earlier simulations to determine what values emerge. The protection can easily be approximated by assuming that children are not allowed to die in any way that might reasonably be prevented by their protectors (e.g., by predators or other individuals killing them, or by dying through lack of food or shelter). Implementing that in the simulations just means the protected children cannot die as a result of the performance-based competition for survival until they have reached the end of their

protection period, with the length of that period evolving in the same way as the other innate parameters such as learning rates. Naturally, the protection will rarely be that effective in real populations, but assuming this extreme case allows a clear exploration of the relevant issues, leaving more realistic models for the future. This approach worked well in the earlier static environment models (Bullinaria, 2009, 2017), but there is an additional complication here, similar to that found for the evolved social learning strategies, that renders this less than straightforward – in changing environments the evolution of a life history trait may be prematurely terminated by a population collapse even when that trait is advantageous. As before, the learning rates evolve quickly and reach their optimal values well before any population collapse, but to establish the range of potential evolved protection periods, low and high mutation rates for it need to be tested for appropriate initial conditions.

A series of evolutionary simulations were run starting from initial populations with small random protection periods drawn from [0, 4] years, and (as before) small random individual learning rates drawn from [0, 0.1] and zero social learning rates. The key simulation results as a function of magnitude of environmental change are presented in Figure 14 for the case of no reproduction while protected. The left graph shows the evolved learning rates for comparison with Figure 2, and the right graph shows the corresponding evolved protection periods for low and high mutation rates. The region of relatively low protection periods for low mutation rates corresponds to the short times to population collapse which, as noted above, leaves insufficient time for optimal values to be reached from their low starting point. For sufficiently high mutation rates, the protection period remains unaffected by increasing magnitudes of environmental change until it dips for very high magnitudes of change. That dip is in line with the faster learning times resulting from the simultaneous adoption of both individual and social learning, as has been observed previously for intermediate levels of transmission fidelity in static environments (Bullinaria, 2017). The reduction in individual learning rates compared to Figure 2 for high magnitudes of change is the expected consequence of the reduced need for fast learning resulting from the parental protection. Thus, it seems that the approach for simulating changing environments proposed in this paper also gives appropriate results when other life history factors are evolved.

## 4. Discussion and Conclusions

In an earlier paper (Bullinaria, 2017), an abstract meme-based simulation approach was developed and tested for studying the evolution of learning strategies that can potentially involve both direct individual learning and imitative social learning. The simulations presented in the earlier sections of this paper were designed to explore the various effects of environmental variability on that approach. They tested the robustness and reliability of the general framework and the chosen methodology for modelling the variability, and also extended the results in the earlier literature on learning in changing environments. The outcomes showed that, despite all the abstractions and approximations involved, appropriate results do emerge across a range of model details and values of the key parameters.

Lifetime learning is known to be advantageous even in static environments when complex behaviours are required (e.g., Bullinaria, 2003, 2009, 2017), and it is clear that it will be even more important when the environment is changing faster than the genetic evolution of the relevant species can track. Moreover, the idea that survival in increasingly variable environments requires the adoption or evolution of increasingly versatile behaviours and learning strategies is also well established (e.g., Potts, 1996, 1998, 2013; Grove, 2011). The various simulations presented in this paper showing the evolution of different variability-dependent learning strategies (e.g., Figures 2, 9 and 10) and the advantage of adaptive learning strategies (Figure 5) are consistent with that. Indeed, all the results presented in this paper appear to be consistent with the broad trends that have been identified in biological evolution.

However, there remains considerable scope for future studies using the same meme-based framework and for exploring further extensions of it. Some of that future work will primarily be concerned with the species-specific life history evolution factors that have already been discussed in some detail by Bullinaria (2009, 2017). Some will be more directly related to the learning issues of this paper that have also been discussed previously, such as improved parameterization for the individual learning (Bullinaria, 2017), more sophisticated strategies for the social learning (e.g., Henrich & Gil-White, 2001; Henrich & Boyd, 2002; Laland, 2004; Rendell et al., 2010), associations between good and bad memes and the creation of memeplexes (e.g., Blackmore, 1999), and the processes for dealing with inconsistent memes (e.g., Shultz & Lepper, 1996). Others concern the interaction of learning and evolution and genetic assimilation via the Baldwin effect (e.g., Best, 1999; Bullinaria, 2003; Crispo, 2007; Jones & Blackwell, 2011). Of particular relevance here are those specific to the details of the meme representations used, and how various forms of environmental change can be accommodated in them.

Many aspects of the simulations in this paper have deliberately been treated rather abstractly with the intention that they will be specified in more detail for particular species and environments later. The simple abstract meme representation used, and the associated mapping from memes to performance of the individuals who have acquired them, could certainly be improved to match specific and more realistic scenarios. For example, having a whole range of memes for each behavioural context rather than simple good/bad pairs. Also, some types of meme will be easier to learn directly and others will be easier to learn socially, but the details of that will depend on the species and environment involved. In particular, none of the simulations so far have allowed the memes themselves to evolve, and it is known that some memes have evolved to the extent that they can no longer be acquired directly by individuals (Boyd & Richerson, 1996). Representing the memes and how their performance contributions are affected by environmental changes in a totally realistic manner in computationally-viable simulations will be challenging, but a feasible first step would be to give the environmental variability a structure better aligned with what is known about the real world (e.g., Halley, 1996), or to formulate it to more closely match particular scenarios that have already been modelled in other ways to see how much more complexity is required to achieve consistent results.

Comparing the results from the current meme-based approach against other simulation studies is a little problematic since the few that have so far been carried out with comparable levels of detail have generally represented individual behaviour in rather different ways. Most other agent-based models of social learning in variable environments have employed simple binary string genotypes and phenotypes with fitness given by their Hamming distance from a bit-string representing the environment (e.g., Best, 1999; Jones & Blackwell, 2011; Borg & Channon, 2012). These are extensions of the groundbreaking work of Hinton and Nowlan (1987) and enable clear experiments concerning the interaction of individual and social learning with genetic evolution and the potential for genetic assimilation of learned behaviours. However, they are far removed from the memes that are learned in the current study, that are designed to represent more complex environments and behaviours that are likely to be beyond genetic assimilation. Moreover, no account has been taken of crucial details like transmission fidelity for the social learning or the speed-accuracy trade-off inherent in real individual learning. Nevertheless, the latest of that series of studies (Borg & Channon, 2012) does provide some scope for comparison of results.

For static environments with perfect transmission fidelity, Borg and Channon (2012) found a clear advantage of social learning over individual learning, as seen in the right-most data points in the top-left graph of Figure 9 in this paper. Of course, the rest of that graph shows how that result is changed for even modest reductions in the transmission fidelity. For environments with low levels of variability they again find social learning is more useful than individual learning as seen for small and medium magnitude variability in Figure 2 of this paper, and for increasing variability they find the individual learning rates rise, again as seen in Figure 2 of this paper. Finally, they found population collapse results consistent with Figure 4 of this paper, and agree with the general finding of this paper that individual learning is necessary for effective social learning, particularly in highly variable environments. This consistency is reassuring for both approaches, and it would be interesting to see if any of the other results of this paper could be reproduced using the simpler bit-string approach.

There is also broad consistency with the individual-based stochastic models of Whitehead (2007) which were designed to investigate the relative advantages of 15 different learning strategies and their dependence on the spectral characteristics of the environmental variation, though further variations of both approaches would need to be implemented to perform a detailed comparison. Similarly, many of the key general results of Whitehead and Richerson (2009), such as the risk of population collapse being reduced by increased individual learning or less conformist social learning, are also consistent with the results of this paper. However, for some magnitudes of environmental change, their results appear to be inconsistent, though checking that will require the learning strategies, learning costs and environmental changes to be brought closer into line. The study by Acerbi and Parisi (2006), based on neural network controlled agents, compared inter- and intra-generational social learning and found that intra-generational transmission increased diversity and improved performance in changing environments. That is broadly consistent with the results for magnitude 2 to 8 changing environments in the meme-based approach here,

but their strict generational structure and rather different environment and performance representations make it difficult to be sure how consistent that is with the non-monotonic performance changes seen here for parents-only transmission in the top-left graph of Figure 10. Again, it would be useful if future work could minimize any unnecessary differences in these two modelling approaches so that more direct comparisons could be made.

A rather different approach has involved the use of mathematical models, which generally allow a more precise formulation of the relevant issues, but usually at the expense of requiring even more abstraction and simplification than the agent-based models to render them tractable (e.g., Feldman et al., 1996; Wakano et al., 2004; Aoki et al., 2005; Wakano & Aoki, 2006). These studies are in broad agreement that greater environmental stability leads to more dependence on social learning and less on individual learning, which is what is found in the simulations of this paper (e.g., Figures 2 and 4) and provides further support for the reliability of its meme-based approach. They also show that in extremely stable environments, innate behaviour will dominate, but that, of course, assumes the required behaviours can be accommodated genetically in realistic timescales which is unlikely to be the case for many complex behaviours, and typical animal development is likely to require lifetime learning however stable the environment is (Bullinaria, 2003). Borenstein et al. (2008) extended the earlier mathematical models to explore different types of environmental change and found some interesting dependencies that would be worth trying to replicate in the agent-based models of this paper. The models of Grove (2011) have aimed to test the variability selection hypothesis (Potts, 1996, 1998) that changing environments lead to adaptations that increase behavioural versatility. The outcomes of the meme-based models presented in this paper (e.g., Figures 2 and 5) are already consistent with that hypothesis, but extending those models to replicate the more detailed results of Grove (2011) would also be worthy of future work. Finally, Henrich and Boyd (1998) found widespread advantages for conformist learning in their models that were not found in the meme-based models here, and that clearly needs investigation in the future.

In conclusion, this paper has presented and tested a general flexible meme-based framework for simulating individual and social learning and the evolution of appropriate learning strategies in variable environments, and has identified a number of potentially fruitful directions for future research using it. Hopefully others will find the approach promising enough to explore it further themselves.

## References

Acerbi, A. & Parisi, D. (2006). Cultural transmission between and within generations. *Journal of Artificial Societies and Social Simulation, 9, 1,* 9.

Aoki, K., Wakano, J. Y. & Feldman, M. W. (2005). The emergence of social learning in a temporally changing environment: A theoretical model. *Current Anthropology*, *46*, 334-340.

Aunger, R. A. (2002). *The Electric Meme: A New Theory of How We Think*. New York, NY: Simon and Schuster/Free Press.

Best, M. L. (1999). How culture can guide evolution: An inquiry into gene/meme enhancement and opposition. *Adaptive Behavior*, *7*, 289-306.

Blackmore, S. (1999). *The Meme Machine*. Oxford, UK: Oxford University Press.

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.

Borg, J. & Channon, A. D. (2012). Testing the variability selection hypothesis: The adoption of social learning in increasingly variable environments. *Proceedings of the Thirteenth Artificial Life Conference*, 317-324.

Boyd, R. & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

Boyd, R. & Richerson, P. J. (1996). Why culture is common, but cultural evolution is rare. *Proceedings of the British Academy of Science*, *88*, 73-93.

Bullinaria, J. A. (2003). From biological models to the evolution of robot control systems. *Philosophical Transactions of the Royal Society of London A, 361*, 2145-2164.

Bullinaria, J. A. (2009). Lifetime learning as a factor in life history evolution. *Artificial Life*, *15*, 389-409.

Bullinaria, J. A. (2017). Imitative and direct learning as interacting factors in life history evolution. *Artificial Life*, *23*, 375-405.

Crispo, E. (2007). The Baldwin effect and genetic assimilation: Revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution, 61,* 2469-2479.

Diamond, J. (2011). *Collapse: How societies choose to fail or succeed*. New York, NY: Penguin.

Ehn, M. & Laland, K. (2012). Adaptive strategies for cumulative cultural learning. *Journal of Theoretical Biology, 301*, 103-111.

Eiben, A. E. & Smith, J. E. (2015). *Introduction to Evolutionary Computing*. Berlin, Germany: Springer.

Feldman, M. W., Aoki, K. & Kumm, J. (1996). Individual versus social learning: Evolutionary analysis in a fluctuating environment. *Anthropological Science*, *104*, 209-232.

Grove, M. (2011). Speciation, diversity, and Mode 1 technologies: The impact of variability selection. *Journal of Human Evolution*, *61*, 306-319.

Halley, J. M. (1996). Ecology, evolution and 1/f noise. *Trends in Ecology and Evolution*, *11*, 33-37.

Henrich, J. & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, *19*, 215-241.

Henrich, J. & Boyd, R. (2002). On modeling cultural evolution: Why replicators are not necessary for cultural evolution. *Journal of Cognition and Culture*, *2*, 87-112.

Henrich, J. & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*, 165-196.

Higgs, P. G. (2000). The mimetic transition: A simulation study of the evolution of learning by imitation. *Proceedings of the Royal Society B: Biological Sciences, 267*, 1355-1361.

Hinton, G. E. & Nowlan, S. J. (1987). How learning can guide evolution. *Complex Systems*, *1*, 495-502.

Jones, D. & Blackwell, T. (2011). Social learning and evolution in a structured environment. *Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems (ECAL 2011)*, 380-387.

Kameda, T. & Nakanishi, D. (2002). Cost–benefit analysis of social/cultural learning in a nonstationary uncertain environment: An evolutionary simulation and an experiment with human subjects. *Evolution and Human Behavior*, *23*, 373-393.

Kendal, R. L., Coolen, I., van Bergen, Y. & Laland, K. N. (2005). Trade-offs in the adaptive use of social and asocial learning. *Advances in the Study of Behavior, 35*, 333-379.

Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences, 38*, e31.

Laland, K. N. (2004). Social learning strategies. *Animal Learning and Behavior*, *32*, 4-14.

Marriott, C., Parker, J. & Denzinger, J. (2010). Imitation as a mechanism of cultural transmission. *Artificial Life, 16*, 21-37.

McElreath, R. & Strimling, P. (2008). When natural selection favors imitation of parents. *Current Anthropology*, *49*, 307-316.

Miller, N. E. & Dollard, J. (1941). *Social Learning and Imitation*. New Haven, CT: Yale University Press.

Muthukrishna, M., Morgan, T. J. H. & Henrich, J. (2016). The when and who of social learning and conformist transmission. *Evolution and Human Behavior, 37*, 10-20.

Potts, R. (1996). Evolution and climate variability. *Science*, 273, 922-923.

Potts, R. (1998). Variability selection in hominid evolution. *Evolutionary Anthropology, 7*, 81-96.

Potts, R. (2013). Hominin evolution in settings of strong environmental variability. *Quaternary Science Reviews*, *73*, 1-13.

Reader, S. M. & Biro, D. (2010). Experimental identification of social learning in wild animals. *Learning and Behavior*, *38*, 265-283.

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., Fogarty, L., Ghirlanda, S., Lillicrap, T. & Laland, K. N. (2010). Why copy others? Insights from the social learning strategies tournament. *Science*, *328*, 208-213.

Rogers. A. R. (1988). Does biology constrain culture? *American Anthropologist*, *90*, 819-831.

Shultz, T. R. & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Psychological Review*, *103*, 219-240.

Steele, J. H. (1985). A comparison of terrestrial and marine ecological systems. *Nature*, *313*, 355-358.

Thornton, A. & Raihani, N. J. (2008). The evolution of teaching. *Animal Behaviour, 75*, 1823-1836.

Tomasello, M. (1999). The human adaptation for culture. *Annual Review of Anthropology*, *28*, 509-529.

Wakano, J. Y. & Aoki, K. (2006). A mixed strategy model for the emergence and intensification of social learning in a periodically changing environment. *Theoretical Population Biology*, *70*, 486-497.

Wakano, J. Y., Aoki, K. & Feldman, M. W. (2004). Evolution of social learning: A mathematical analysis. *Theoretical Population Biology*, *66*, 249-258.

Walker, R., Burger, O., Wagner, J. & Von Rueden, C. R. (2006). Evolution of brain size and juvenile periods in primates. *Journal of Human Evolution, 51*, 480-489.

Whitehead, H. (2007). Learning, climate and the evolution of cultural capacity. *Journal of Theoretical Biology*, *245*, 341-350.

Whitehead, H. & Richerson, P. J. (2009). The evolution of conformist social learning can cause population collapse in realistically variable environments. *Evolution and Human Behavior*, *30*, 261-273.
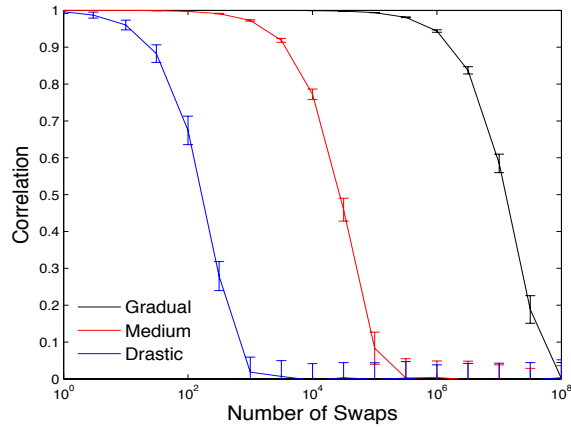
Figure 1: Correlation of meme performance contributions as the environment changes in gradual (magnitude 1), medium-sized (magnitude 30) and drastic (magnitude 500) steps. Actual environment changes will depend on the distribution of the relevant types of swaps over time.
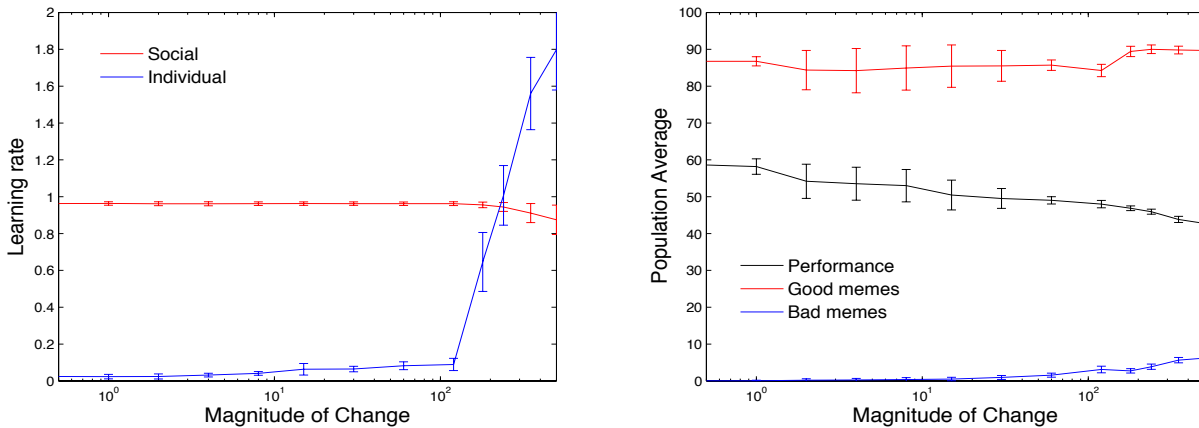


Figure 2: Evolved social and individual learning rates for different magnitudes of environmental change (left), and the associated performance levels and numbers of good and bad memes carried by the population (right), for perfect fidelity social learning.
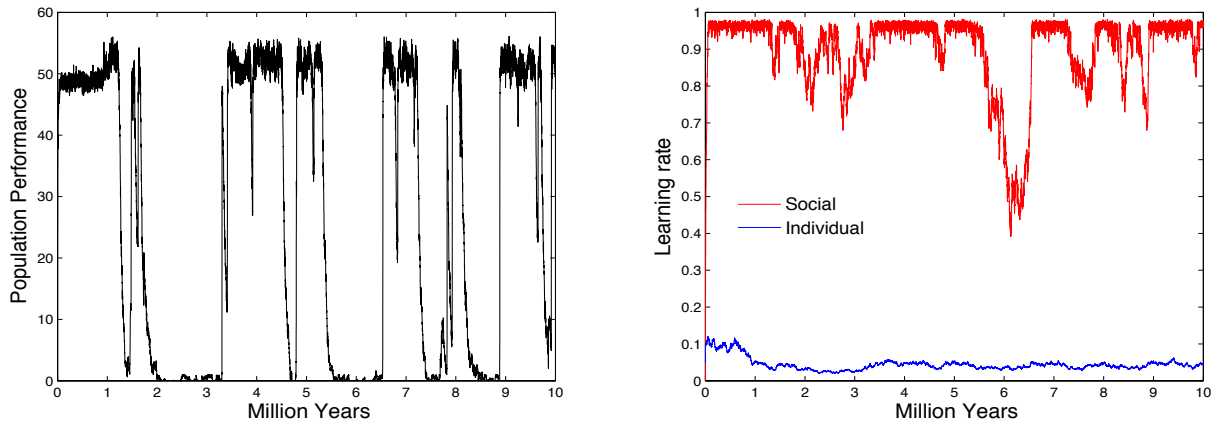
Figure 3: Evolution of population average performance (left) and social and individual learning rates (right) during a typical simulation with intermediate (magnitude 30) environmental changes.
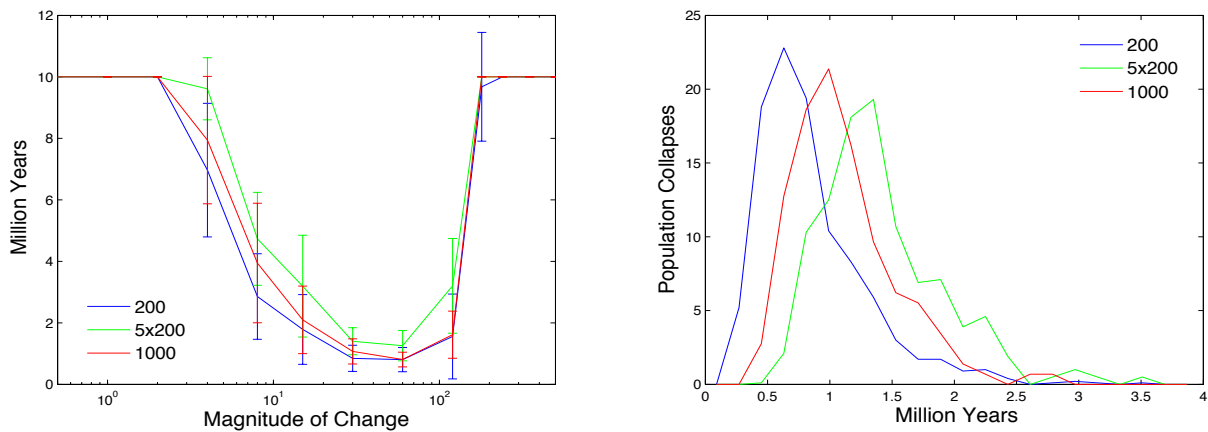


Figure 4: Average times to population collapse and extinction as a function of the magnitude of environmental change (capped at maximum run length of 10 million simulated years) for population sizes 200, 5×200 and 1000 (left), and the associated normalized distributions of population collapse times for intermediate (magnitude 30) environmental changes (right).
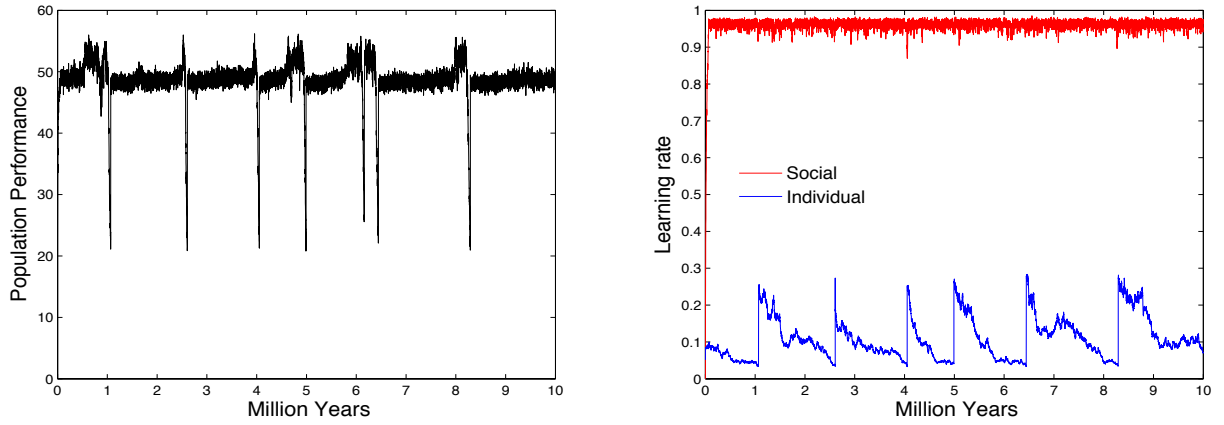
Figure 5: Evolution of population average performance (left) and social and individual learning rates (right) during a typical simulation with intermediate (magnitude 30) environmental changes when a simple adaptive learning strategy is used to avoid population collapse (for comparison with Figure 3).
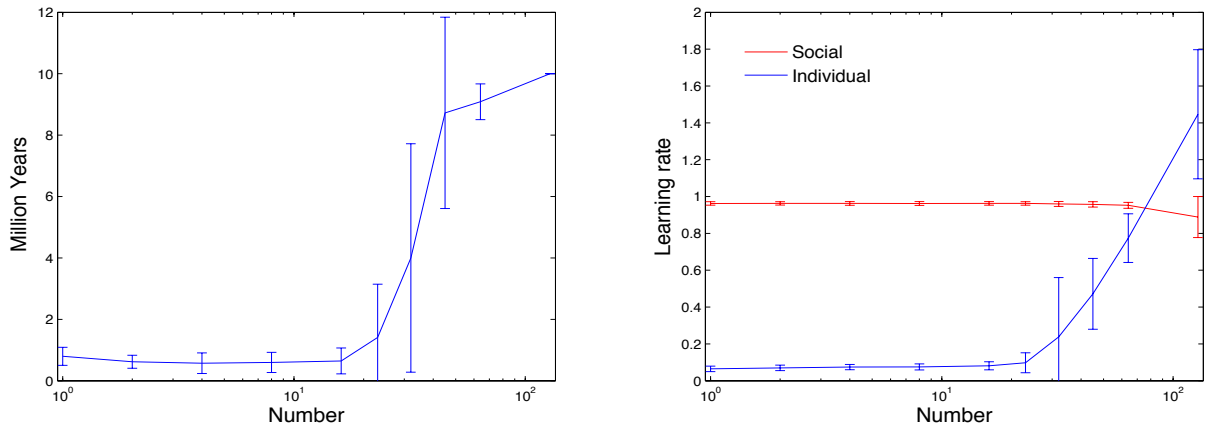


Figure 6: Changes to the average time to population collapse (left) and learning rates (right) when intermediate (magnitude 30) environmental changes occur $N$ times per simulated year instead of only once per year.
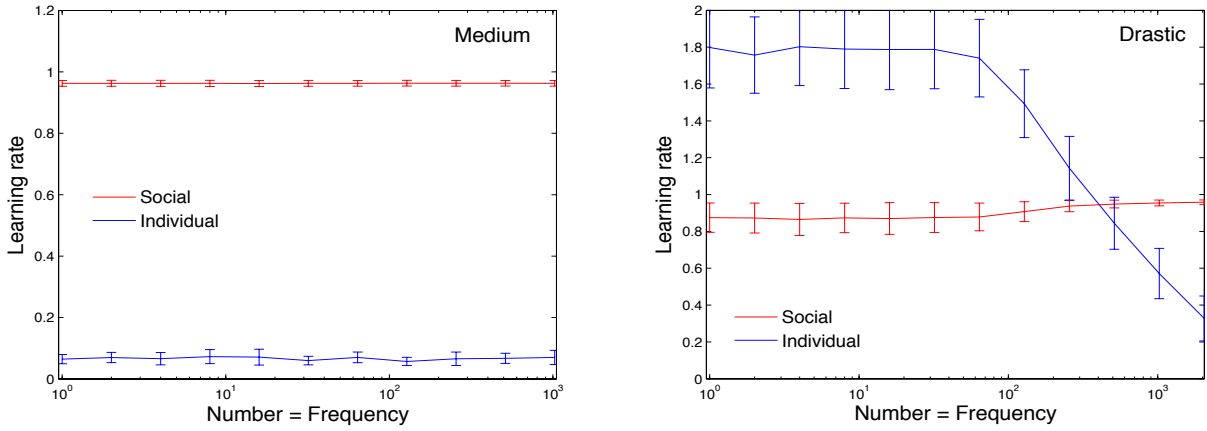
Figure 7: Average learning rates when intermediate (magnitude 30) and drastic (magnitude 500) environmental changes are bunched up with a larger number *N* changes occurring once every *N* simulated years instead of one change per year.
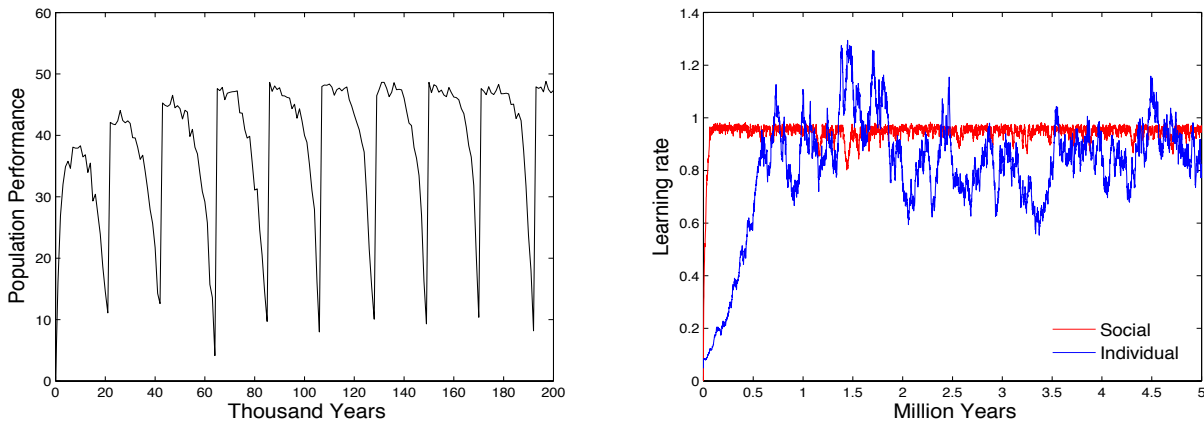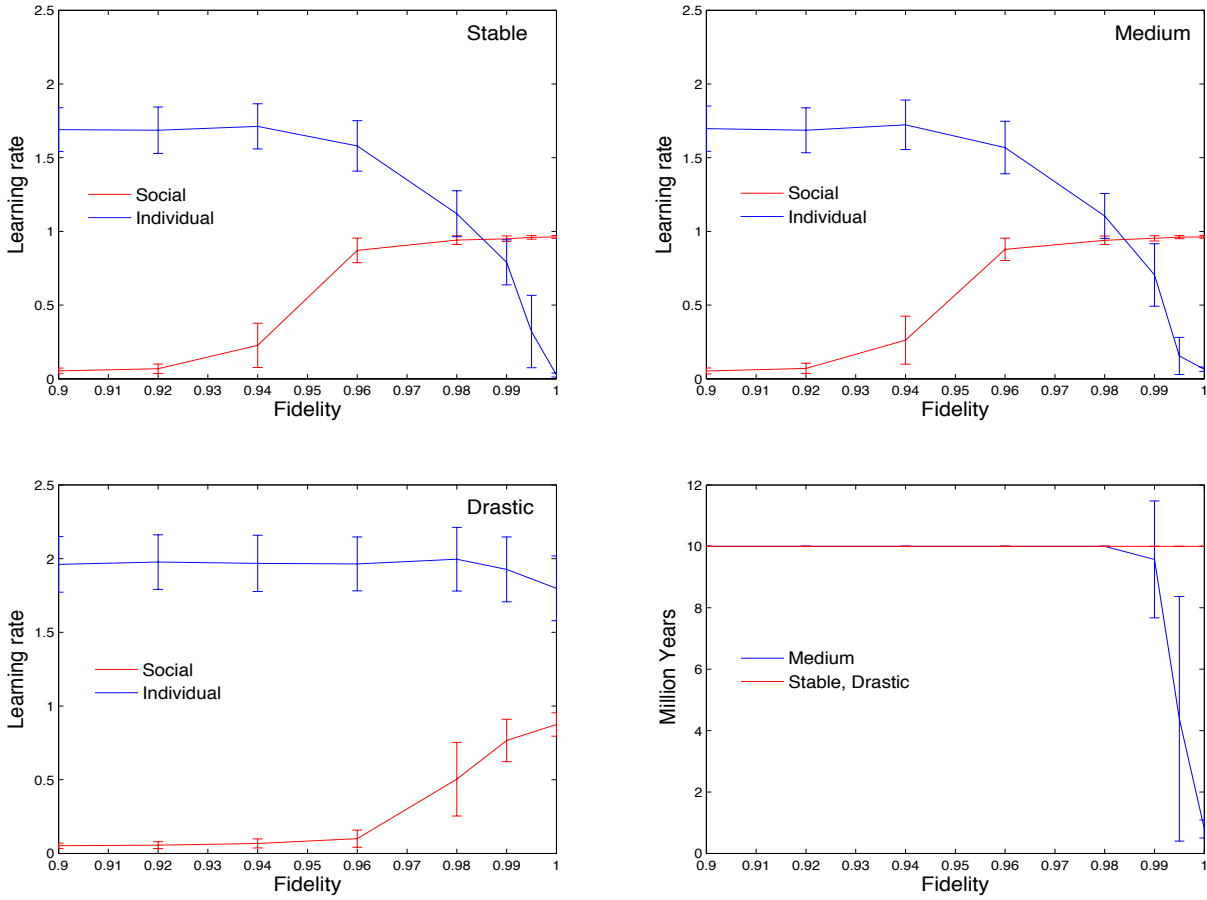


Figure 8: Typical evolution of population performance (left) and average learning rates (right) when drastic (magnitude 500) environmental changes are bunched up with 512 changes occurring once every 512 simulated years instead of one change per year.

31

Figure 9: Evolved social and individual learning rates as a function of transmission fidelity for stable environments (top left), intermediate (magnitude 30) changing environments (top right), and drastic (magnitude 500) changing environments (bottom left). The key difference between the stable and intermediate cases is in the average times to population collapse (bottom right).
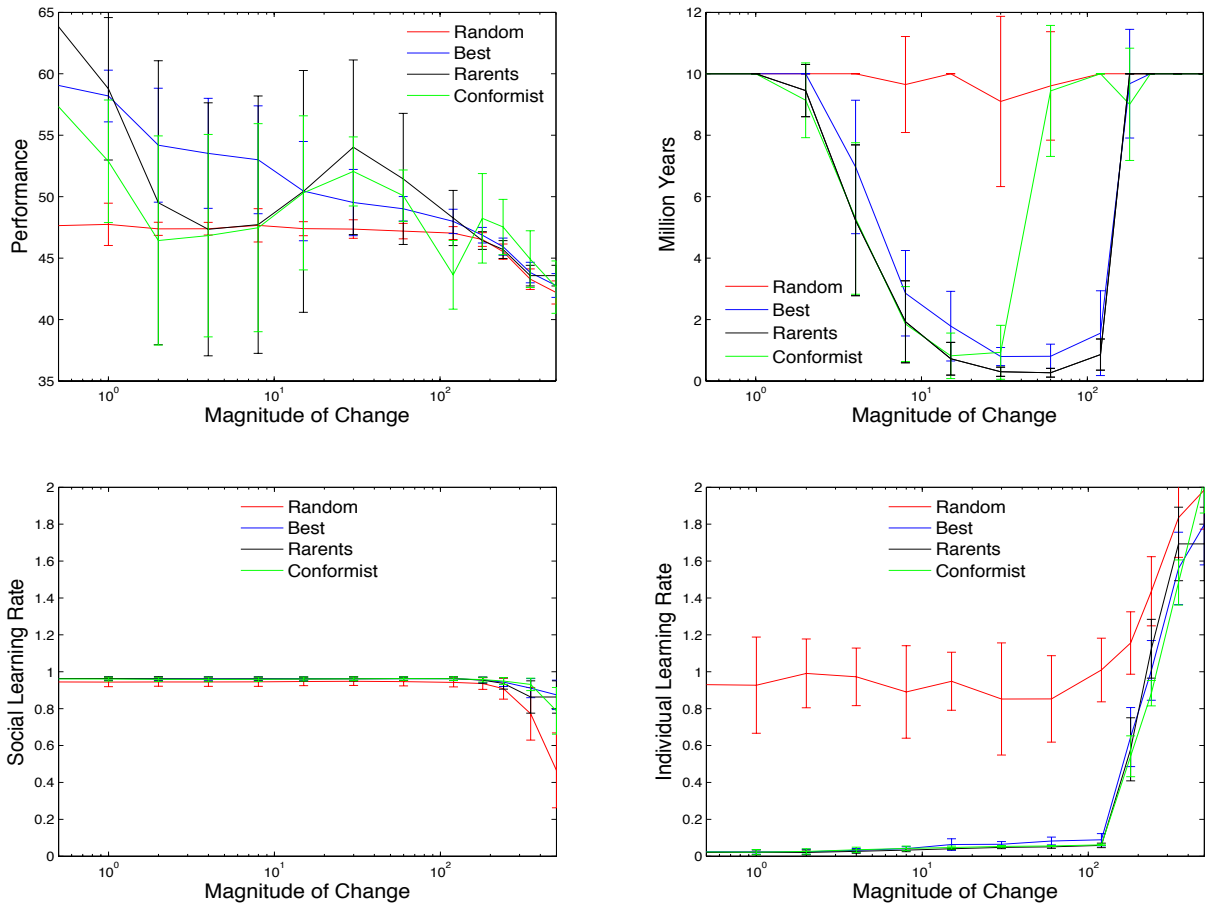
Figure 10. Effect of social learning strategy (copy random individuals, copy best individuals, only copy parents, and conformist) on the average population performances (top left), times to population collapse (top right), social learning rates (bottom left), and individual learning rates (bottom right), for a range of environmental change magnitudes.
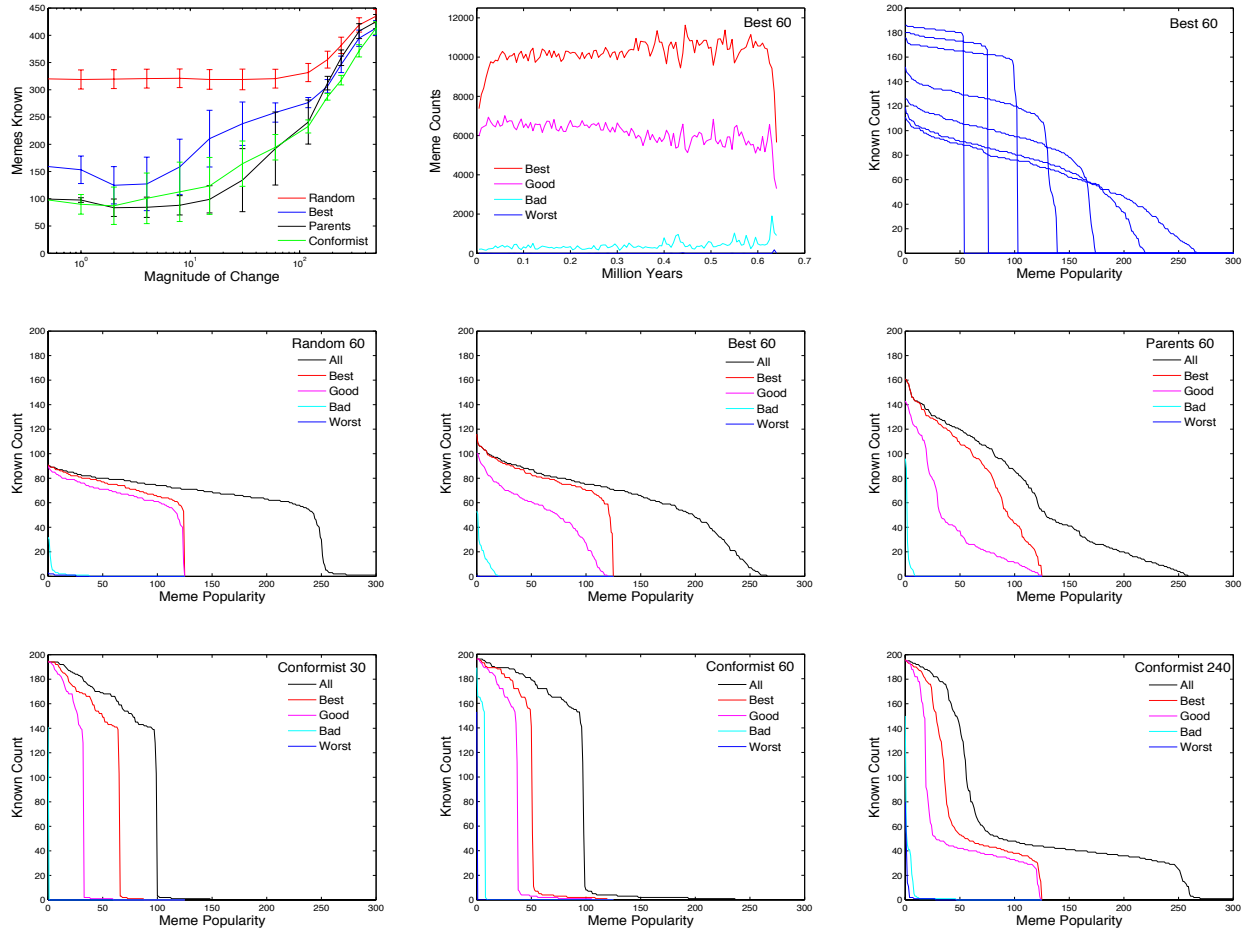
Figure 11. Distributions of the learned memes. Diversity of the known memes for the four social learning strategies (top left). Time-course of meme counts during a typical copy-the-best magnitude 60 simulation (top centre) and the eventual collapse of the meme distribution (top right). Typical steady-state meme distributions for each of the four strategies with magnitude 60 changes, and, for comparison, the conformist strategy with magnitude 30 and 240 changes (middle and bottom rows).
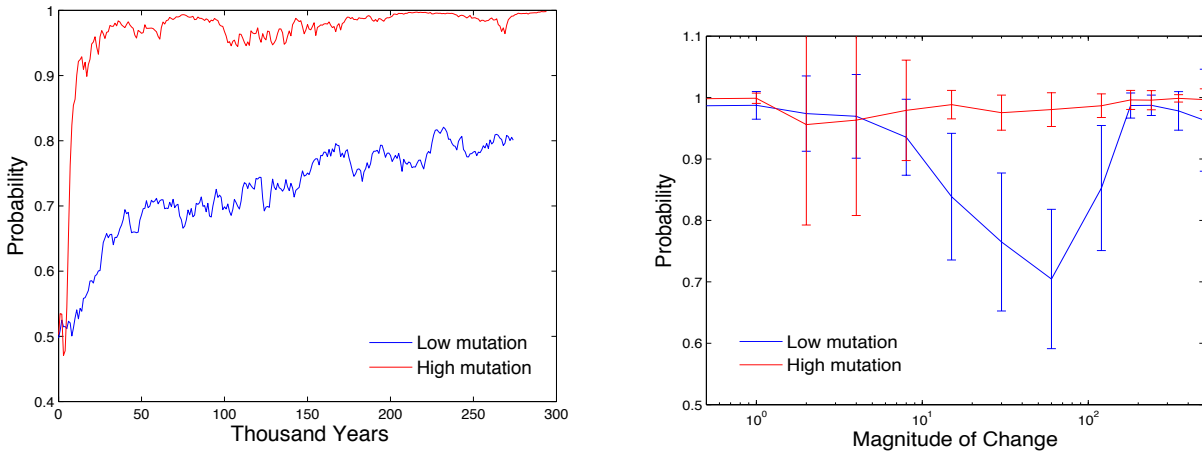
Figure 12. Evolution of social learning strategy specified by probability of copying best rather than random individuals. Typical runs till population collapse with low and high mutation rates for intermediate (magnitude 30) environmental change (left), and the final evolved strategy probabilities as a function of mutation rate and magnitude of environmental change (right).
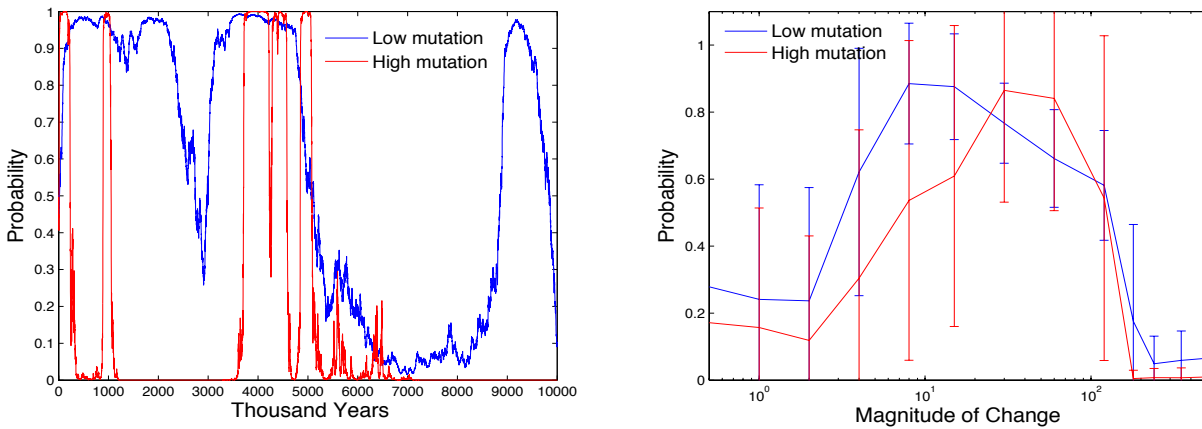


Figure 13. Evolution of social learning strategy specified by probability of copying parents rather than best individuals. Typical runs with low and high mutation rates for low (magnitude 1) environmental change (left), and the final evolved strategy probabilities as a function of mutation rate and magnitude of environmental change (right).
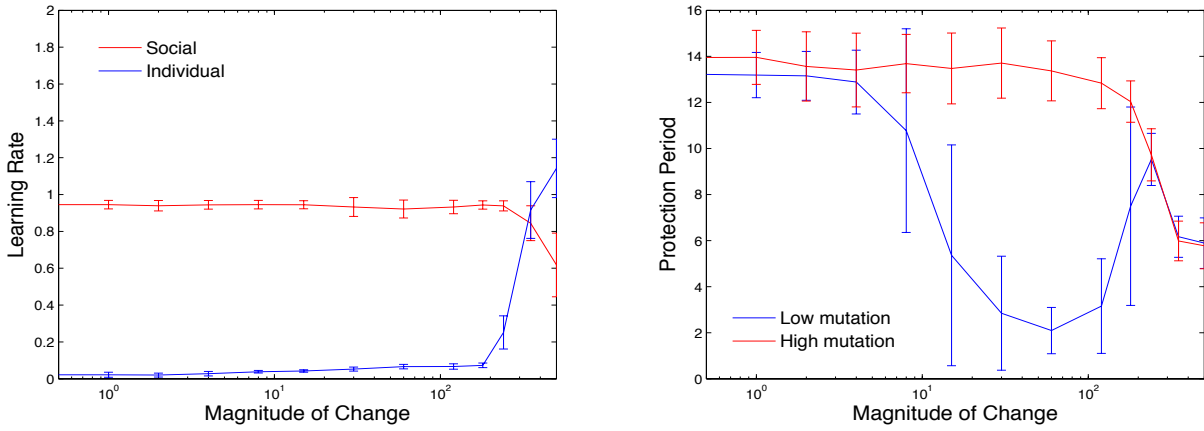
Figure 14. Evolved social and individual learning rates for different magnitudes of environmental change when protection periods are allowed to evolve (left), and the associated protection periods that evolve with low and high mutation rates (right).