



Learning to cope: the evolution and degradation of complex traits

Sergio González-Mollinedo



**university of
groningen**

Learning to cope: the evolution and degradation of complex traits

MSc. Thesis

Sergio González-Mollinedo

S4524179

Master Ecology & Evolution
Track Evolutionary Biology (MEME)

September 2021 to February 2022

Supervised by:
Prof. Dr. F.J. Weissing and Dr. M.A. Kozielska-Reid

Theoretical Research in Evolutionary Life Sciences

University of Groningen

Abstract

Adapting to a challenging environment is key for the survival of endangered populations. Fields like Conservation Genetics or Evolutionary Rescue Theory investigate the conditions under which small and declining populations can adapt to changing conditions. However, virtually all models in this field focus on simple traits and a simple relationship between genotypes and phenotypes. Accordingly, our current understanding of how complex traits involved in adaptation are affected by genetic erosion is poor. Here I study whether and when a complex adaptation, learning from experience, will evolve in a large population and under what conditions this adaptation will persist when the population is experiencing permanent or cyclic bottlenecks. In a first step, I investigated the evolution of decision-making (with or without learning) in a large population that is subject to fluctuating environmental conditions. To this end, I run individual-based simulations with an already existing neural network model for learning. It turned out that learning only evolved if the environmental regime was unstable enough to benefit from learning, but consistent enough so that previous environmental states are informative, similar to environmental regimes that favour the evolution of plasticity. In those scenarios where learning did evolve, one of two learning styles evolved, which I be characterized as 'reckless learning' and 'fine-tuned learning', respectively. Reckless learners devote only a short fraction of their lifetime to learning, and during this learning period they learn in large steps. In contrast, Fine-tuned learners devote several timesteps to learning and carefully adjust their decision-making curve to the environment. learning styles. The two learning styles present differences in the number of resources gained and in the performance of decision-making, or the proportion of the maximum number of resources gained if the highest quality environmental cue was chosen in all of the exploitation phase. Reckless learning had a mean performance of 0.49 while the fine-tuned learners approached performances of 0.6-0.8, which signified an increase of about 50% in mean resource collection. In a second step, I analysed the effect of population bottlenecks in the evolution and maintenance of evolved learning styles. To understand the effect of permanent bottlenecks on the evolved styles, I took the best networks from ten separate replicates from a single environmental regime where both styles could evolve. I used each of these networks to create monomorphic populations of different sizes, ranging from 50 individuals to the original size of 1,000. I could observe a decrease in the performance and mean resource collection in population sizes of 50 and 100, suggesting that complex traits decrease their performance in long and harsh bottlenecks. I also analysed the effect of cyclic bottlenecks in the evolution of populations in different environmental regimes. I subjected populations to bottleneck treatments by reducing their size to 50 and 100 individuals for 50 or 100 generations, every 500 and 1000 generations, respectively. In these cyclic bottleneck treatments, I could not see a clear signal of the degradation of learning styles or populations that adopted other strategies such as adaptive tracking or bet hedging. In conclusion, populations that have evolved complex traits will maintain their performance during cyclic bottlenecks, but it will require long and harsh bottlenecks to actually decrease their performance.

Table of Contents

I.	Introduction	4
II.	Materials and methods	6
III.	Results	9
IV.	Discussion	22
V.	Conclusions	26
VI.	Literature Cited	27
VII.	Appendix 1	29
VIII.	Appendix 2	34
IX.	Appendix 3	39
X.	Appendix 4	40

I. Introduction

The loss of biodiversity is one of the most prominent global crises that is currently unfolding due to human activities (Ceballos *et al.*, 2015). Local extinctions have increased by habitat perturbation and climate change, also increasing the amount of small and isolated populations that might go extinct in the future (Matthies *et al.*, 2004; Wiens, 2016; Staude *et al.*, 2019; Román-Palacios and Wiens, 2020). The persistence of small and isolated populations facing environmental disturbances has become a priority for conservation efforts. Management of these populations has focused on reducing the effects of stochastic events in the population and in avoiding inbreeding depression (Frankham *et al.*, 2014; Frankham *et al.*, 2017). In this context, population persistence then relies strongly on the ability of individuals to adjust and adapt to the changing environmental conditions and other disturbance that they might experience despite low effective population sizes and loss of genetic diversity (genetic erosion). Adaptation might appear in a novel environment through beneficial mutations spreading in the population, but the pace of environmental change and degradation coming from human activities might be too fast for some populations to cope. Therefore, mechanisms of rapid adaptation to disturbed environments becomes crucial, such as adaptive phenotypic plasticity (Pazzaglia *et al.*, 2021).

Genetic erosion and inbreeding depression are concerning consequences of bottleneck events and fragmentation of wild populations, which results in loss of fitness and higher extinction risk (Frankham *et al.*, 2017). The main explanation given to this loss of fitness is through the expression of deleterious alleles or suboptimal phenotypes that become more likely to persist in the population due to loss of heterozygosity and genetic drift (Fredrickson *et al.* 2007; Ehlers *et al.* 2008; Funk *et al.* 2010; Johnson *et al.* 2010; Frankham *et al.*, 2017). However, the effects of genetic erosion and inbreeding depression are rarely related to phenotypic plasticity despite its prominent role in the rapid adaptation of populations to environmental change. The literature on the effects of inbreeding depression focusses on loss of fitness, but not in the possible loss of regulation mechanisms or plastic responses. This is reflected in genetic management strategies such as genetic rescue, which has proved to be an effective measure to restore the genetic diversity and evolvability of endangered populations but pays little attention to the underlying mechanisms that it is restoring by introducing genetic diversity into the population (Tallmon *et al.*, 2004, Frankham *et al.*, 2014). However, understanding these underlying mechanisms beyond genetic diversity may help to refine the management strategies that can be undertaken to increase the chances of survival of endangered populations.

Some studies have taken interest in the effects of genetic erosion on the ability of populations to maintain plastic responses, but usually, these conclusions are drawn from studies in controlled conditions and the interaction with environmental fluctuation is rarely considered (Luquet *et al.*, 2011). The studies that have analysed the interactions of genetic erosion and adaptive phenotypic plasticity have found conflicting evidence, with some populations losing their adaptive plasticity (Maynard Smith, 1956; Auld and Relyea, 2010) and others maintaining it (Luquet *et al.*, 2011). These studies have suggested that the complex genetic architecture that encodes the traits of interest might have a role in their decay in a context of genetic erosion, mentioning that traits that are encoded by a larger number of loci have a higher probability of expressing deleterious alleles (Auld and Relyea, 2010; Luquet *et al.*, 2011). However, it has been observed that traits with complex genetic architectures evolve to have a high robustness against the expression of deleterious mutations and maintain stable expression of the phenotype, especially when selection for the trait is strong (Wagner, 2005; Wagner, 2008; Masel and Siegal, 2009; Zheng *et al.*, 2020).

This is a property that allows complex traits to explore multiple genetic configurations which will produce the same phenotype even under environmental noise (Wagner, 2005; Wagner, 2008; Stewart *et al.*, 2012).

The inconsistencies in the effect of genetic erosion on plastic traits then raises interesting and relevant questions not only on how organisms can adapt to challenging conditions, but on the genetic architecture and the evolutionary aspects that determine which processes might decay or not when populations decline. Adaptive phenotypic plasticity can evolve under certain environmental conditions in which the frequency of change is moderate and environmental cues are correlated or relatively good predictors of the future environmental state (Botero *et al.*, 2015). Even though Botero *et al.* (2015) used a model comprised of evolving reaction norms, they were able to determine that the genetic architecture of the traits encoded in their populations was also key in determining how populations could adapt to different environmental regimes, identifying evolutionary tipping points that caused populations to collapse when their environmental regime was shifted. Therefore, the environmental regime where a population has evolved can not only dictate their ability to cope with novel environmental conditions but could also be a determining factor for the robustness of the traits that are evolved and their future evolvability (Masel and Siegal, 2009; Zheng *et al.*, 2020).

In this study, I try to assess the effect of population declines, possibly leading to the sustained loss of genetic diversity (i.e. genetic erosion), on a complex trait using an individual-based model (IBM) in which each individual has an artificial neural network (ANN) that can evolve to assess the quality of the environment where they forage by learning from environmental cues. I also draw parallels between learning and adaptive phenotypic plasticity, as the learning episodes in the model allowed for the modification of individual behaviour and the possibility to perform better within a lifetime, given that learning evolves, as in the initial proposition of the Baldwin Effect, that suggested that learning was a form of adaptive behavioural plasticity (ontological development) that allowed populations to climb closer to fitness peaks (Baldwin, 1896). In the first part of my study, I address the environmental conditions, such as magnitude and frequency of environmental shifts, that favour the evolution of learning. In the second part, I test if the evolved learning mechanisms are robust to cyclic and permanent reductions of population size.

In the first phase of this study, I explored the environmental conditions that allow learning to evolve and be maintained in a population, comparing different combinations of frequency and magnitude of environmental change, asking which environmental regimes would favour the evolution of learning. I expected learning to evolve in relatively unstable, but predictable, environmental regimes as those found to favour the evolution of plasticity by Botero *et al.* (2015) showing that the evolution of learning would contribute to the ability of populations to express a phenotype that adjusts to the environment (i.e. environmental robustness) in moderately unstable environmental regimes. Extremely unstable environments seemed to favour the evolution of bet-hedging. I explored the effect of population bottlenecks by reducing the population size in the simulation for an extended period, but also cyclic bottlenecks that would submit the populations to a period of increased genetic drift. In terms of robustness to population bottlenecks, I expected the populations where learning evolved to have an increased robustness against the accumulation of deleterious mutations, as selection has been observed to increase environmental and mutational robustness (Ciliberti *et al.*, 2007a). My results allow to draw some general conclusions on the maintenance of complex traits in reduced populations and how an

interplay of robustness to mutational and environmental noise might play out during population declines. To achieve this, I take simple definitions of mutational and environmental robustness that relate to general properties of the evolved networks and analyse their outcome during population bottlenecks. They could also contribute to design better strategies for evolutionary rescue, as complex traits might not respond as expected under genetic management.

II. Materials and methods

Model description

I adjusted an individual-based model (IBM) that had been previously developed to explore the evolution of learning. Every individual in the population can go through three phases during its lifespan: learning, exploitation of environmental resources, and reproduction. Each generation time consists of the total number of timesteps that amount to the lifespan of the individuals. In this case, simulations were run with a lifespan of 100 timesteps for all individuals.

Learning is an evolvable property of the individuals of the simulated population. An individual may evolve to allocate a certain number of timesteps of its lifespan as learning episodes to learn from the environment to better exploit it later during the timesteps it allocates to collect resources. The number of learning episodes that an individual evolves to have determines how long it will spend learning. After the individual carries out its learning episodes, it begins to exploit the environmental resources for the rest of its lifespan. It is important to note that the individuals are unable to exploit the environmental resources while learning, which creates a trade-off between learning and exploitation during their lifespans. Once their lifespans are completed, individuals will produce a number of offspring proportional to the resources they have collected (more resources result in a higher fitness and a higher number of offspring in the new generation) and a new population replaces the previous one, as another generation of the simulation begins (see Table S1 for an overview of the model parameters).

Environment

The environment where the populations evolve could be considered as a spectrum of nutrient sources of different quality. Individuals of the population receive cues that are completely reliable to the quality of the nutrient source that is found in the environment. The cues received by the individuals make these nutrient sources discrete, but the quality of the nutrient sources is given by a continuous wrapped gaussian function on an environmental axis delimited from -1 to 1, which is the range of possible cues (Figure 1). The environmental function is described by the environmental mean (or peak) and the environmental standard deviation (or width). All simulations in this study were carried out with an environmental SD of 0.1. The environment corresponds to a range from which cues for learning are drawn. The environmental mean is the value of the cue that has the highest possible quality, with a value of 1. The environmental mean can change with a given frequency during a simulation, but the standard deviation remained fixed.

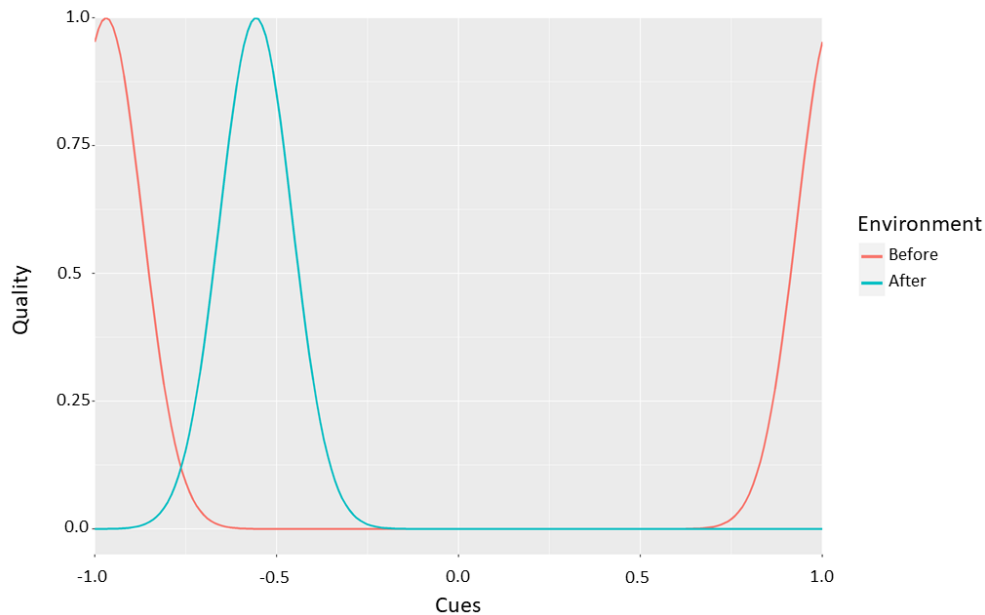


Figure 1. Structure of environmental quality and its temporal shifts. The environment in the simulation is represented as a range of cues that are associated with the quality given by a wrapped gaussian function, which can shift along the environmental axis. The red curve represents the environmental function before a shift and the blue curve shows its new position.

Artificial neural networks

Each individual in the model has an ANN, which allows them to receive environmental cues as input and process them to obtain an expected quality for the nutrient source represented by that cue. The network can evaluate cues in the entire range of the environmental function and return an expected quality associated to each environmental cue it receives. With this evaluation, individuals can potentially discriminate between low- and high-quality cues and decide which is the one that has the maximum quality.

The ANN consists of an input node which receives the environmental cue, two hidden layers of four nodes each, and an output node which returns the expected quality associated to the environmental cue received. It includes nine node biases and 24 weights connecting its nodes (Figure 2). Each weight can vary in value and sign, simulating the connection strength between the nodes of the ANN. The ANN imitates a neuron cluster that can process and interpret environmental information. Biases are intrinsic activation levels that could also be found in biological systems and the strength of the connection between neurons relates to the weights of the network (similar to neural pathways). The activation function of nodes is a clamped rectified linear unit (ReLU, Figure S11) constrained between 0 and 1. This function receives the activation signal from the nodes upstream in the network and then assigns a value between 0 and 1 that is sent to the nodes downstream of that node. The network can evolve by the occurrence of mutations in its biases and weights that can modify its expectations curve.

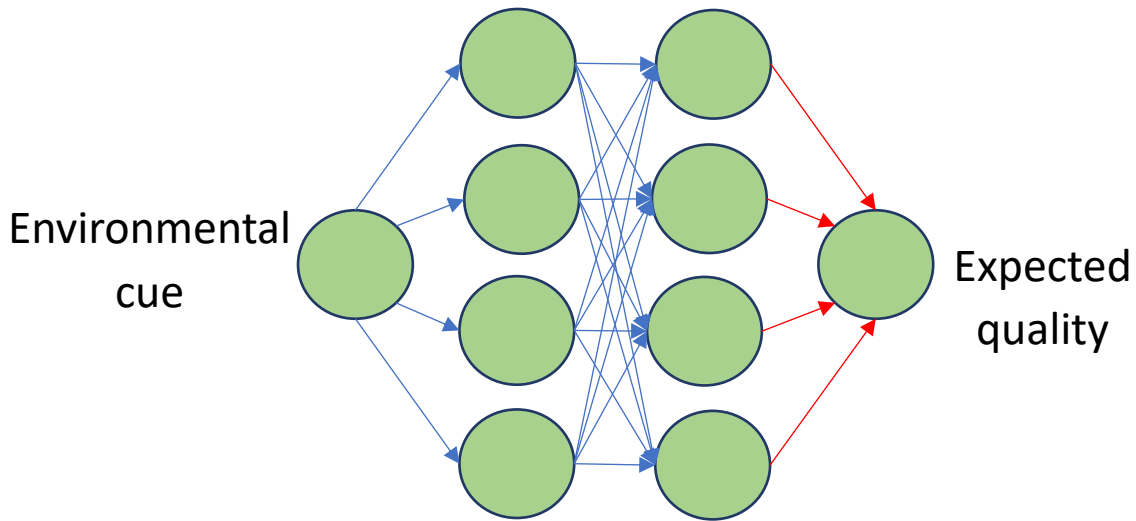


Figure 2. Structure of the neural networks considered in this study. Schematic representation of a neural network that is used for decision-making and learning. The single node to the left (the ‘input node’) is fed by the current environmental cue. This information is further processed (in a ‘feed-forward’ manner) in two layers of four neurons each. The information accumulated in the second-layer neurons is processed by the single node to the right (the ‘output node’), which returns a value of the expected environmental quality associated with the given cue. This expected quality is used in learning to adjust the weights that lead to the output node and in foraging, the network will evaluate each cue it is presented and select the one that will provide the maximum expected quantity of resources (i.e. quality). Weights modifiable by learning are shown in red.

Learning

Learning occurs during each learning episode that the individual has evolved to have. In every learning episode, an individual is presented with an environmental cue (input) from the environmental range and the ANN assesses the expected quality associated to that cue. It is later presented with the real quality, and it adjusts the four weights that connect to the output node (coloured red in Figure 2) with the following equation:

$$\Delta w_{ij} = \alpha(q_j - e_j) * x_i$$

The equation corresponds to the delta rule where each of the four weights (w_i) is modified when presented with the cue j . The weight is modified by Δw_{ij} which is the product of the learning rate (α), the difference between the real environmental quality of cue j (q_j) and the expected quality generated by the network for that same cue (e_j), and the input generated by the node connected to weight i (x_i). For each timestep in a lifetime the same cues are presented to all individuals, but the amount of learning episodes is a trait that can vary among them. The individual ANNs will learn depending on their own learning rate and weights for as many learning episodes as they have evolved to have. After the individual has finished its learning period, it proceeds to exploit the environment for resources.

Exploitation

After the individual has finished learning (if the individual is allocating timesteps to learn), it proceeds to exploit the environment to collect resources. During each timestep of the exploitation period, an

individual is presented with five random environmental cues as input. The individual performs an evaluation of each cue using its network function and then chooses the cue that has the highest expected quality. After this, the individual collects a quantity of resources that corresponds to the real quality of the environmental cue that it chose. Exploitation episodes continue to happen until the lifespan of the individual is over. By this point, individuals in the population have collected a quantity of resources that will influence their probability of reproduction.

The way cues are given to the individuals as input does not mean that the best cue guarantees the individuals to obtain the highest possible quantity of resources (1), but the highest possible quantity among the cues/environments sampled in each timestep. Therefore, performance in the exploitation phase is measured as the proportion of resources collected out of the maximum possible quantity of resources available during the exploitation period of the individual (Performance is 1 when the individual correctly chooses the best cue every time).

Reproduction

At the end of every generation, the population undergoes a reproduction event. The individuals are haploid and transfer their genome with a chance of mutation to their progeny. The total quantity of resources collected by the individual during its lifetime is directly proportional to the number of offspring it leaves to the new generation, therefore I refer to the quantity of resources collected as fitness from now on, since it translates directly into reproductive success.

At the start of each generation, a new population is drawn from the pool of individuals in the previous generation, sampling genomes with a probability directly proportional to their fitness. Individuals of the next generation inherit the ANN, the number of learning episodes and the learning rate of their parents. This new population is assembled from this sampling, and mutations may occur in the ANNs, the number of learning episodes and the learning rate of the new population with a probability of 1% per generation. The mutation steps for the weights and biases of the ANN, and the learning rate, are drawn from a Cauchy distribution with a location of 0.001. The mutation steps for learning episodes are drawn from a Poisson distribution with a mean of 1. These settings allow for large, but rare mutations, which could allow a rapid evolution of learning mechanisms. When using a normal distribution to draw mutational steps in a different project, learning would take much longer and would be more unlikely to evolve (Kozielska, personal communication) The base population size for the model I ran is 1,000 individuals, therefore, on average one third of individuals carries mutations every generation.

III. Results

Experiment 1: Learning in different environmental regimes

Learning depends on the ability of organisms to extract information from their environmental context to better understand its state and use this information in later actions. Therefore, the efficiency of learning is closely linked to the degree of predictability of future environmental states and to the frequency of environmental change.

To explore the environmental conditions where learning evolves, I created an experiment with various configurations of environmental change, or environmental regimes. Different magnitudes of the displacement of the mean of the environmental function were interpreted as the predictability of the

future environmental state. With greater displacement, it would be difficult to keep track of the future position of the environmental peak, but smaller displacements would make it easier to track. The mean of the environmental function was allowed to shift randomly left or right along the environmental axis, with shifts of 0.1, 0.2, 0.3 and 0.4. Then, another key feature for the evolution of learning is the frequency of environmental change. In Botero *et al.* (2015), the frequency of environmental change is found to also be an important factor affecting the evolution of plasticity. In this study, I implemented simulations where the environment changes every 125, 25, 5 and in every generation. I ran simulations for 16 different parameter combinations to test if learning could evolve in those environmental regimes. A population of 1,000 individuals was subjected to each environmental regime, and I allowed them to evolve for 50,000 generations. A control for these experiments was running the same simulation parameters, but in replicates where the evolution of learning was not possible, by fixing learning episodes to 0 and not allowing them to mutate. A comparison between learning and non-learning populations allowed me to identify the different strategies that evolved in different environmental regimes.

To assess the evolved mechanisms, I took the mean of fitness, performance, learning episodes, and learning rate for the last 5,000 generations of the simulations and plotted each replicate of both learning and non-learning simulations for each parameter combination. This allowed me to identify the parameter combinations where learning could provide a fitness benefit and could also be a trait that is selected for in the population. I summarized the results in composite plots where rows represent the magnitude of the environmental shifts and columns represent the frequency of environmental change (Figure 3).

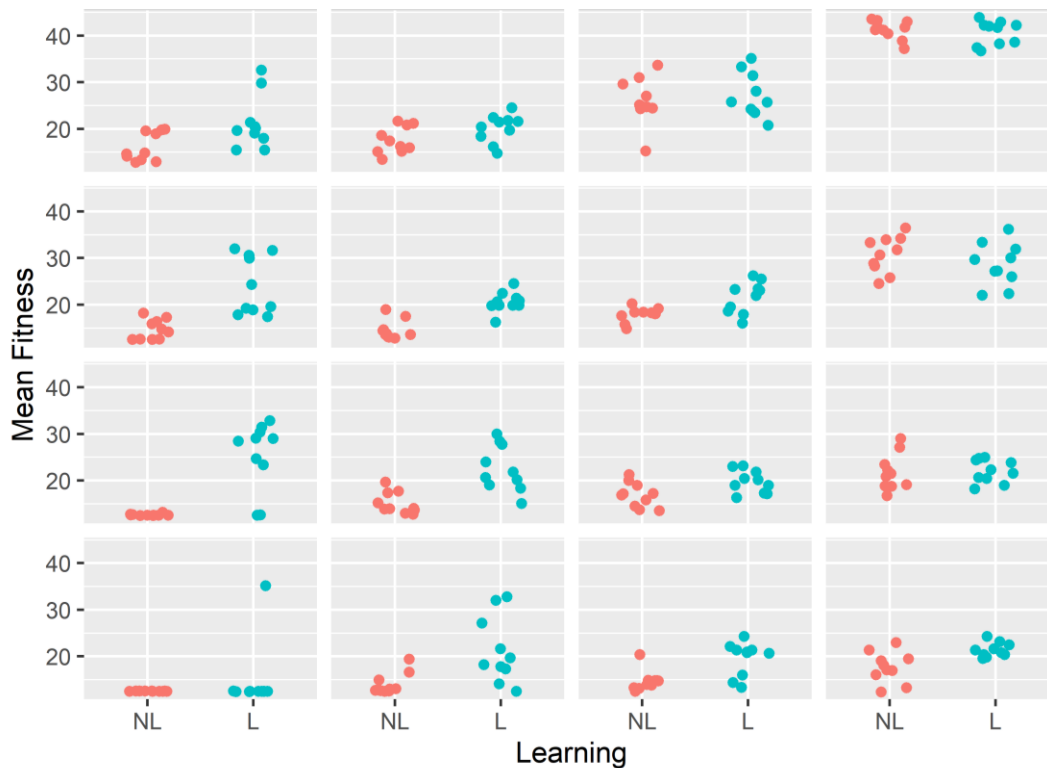


Figure 3. Mean fitness comparison between learning (L) and non-learning (NL) populations evolved in different environmental regimes. The rows represent the different magnitudes of the environmental shifts (0.1, 0.2, 0.3, and 0.4, from top to bottom) and columns represent the different frequencies of

environmental change used in the simulations (every generation, 1/5 generations, 1/25 generations and 1/125 generations, from left to right). L populations usually obtain a higher fitness (average over the last 5,000 generations of the simulation period) than NL populations in relatively unstable environmental regimes.

In this set of simulations, I expected to see different general strategies arise. First, when comparing the more stable environmental regimes (towards the top right corner), we can see that learning and non-learning populations are obtaining a similar mean fitness. These populations are extremely susceptible to environmental changes and experience a fitness drop when the environmental mean shifts (Supplementary Material, Figure S1) and sometimes evolve learning episodes but tend to lose them in the long run, suggesting that they are falling under an adaptive tracking of the environmental mean.

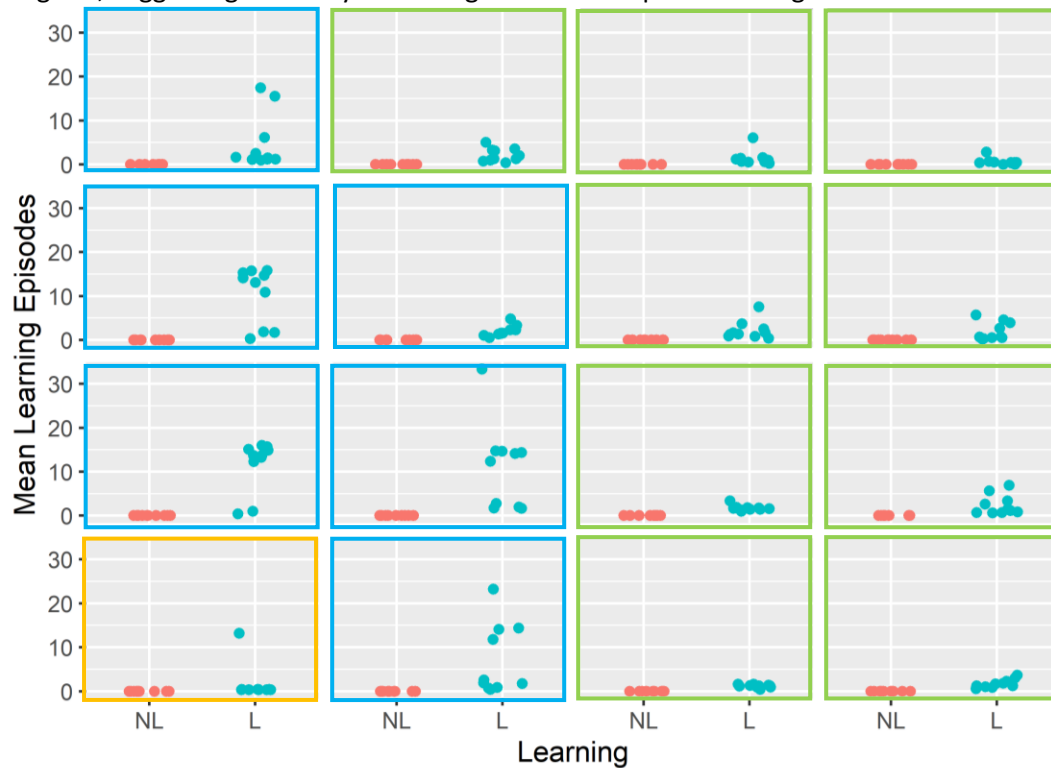


Figure 4. Mean number of learning episodes that evolve in different environmental regimes for populations evolving learning (L) and non-learning control (NL). Control non-learning replicates show 0 learning episodes and are compared to replicates capable of evolving learning. The rows represent the different magnitudes of the environmental shifts (0.1, 0.2, 0.3, and 0.4, from top to bottom) and columns represent the different frequencies of environmental change used in the simulations (every generation, 1/5 generations, 1/25 generations and 1/125 generations, from left to right). Learning episodes (averaged over the last 5,000 generations of the simulation period) evolve and are maintained especially in those environmental regimes where the fitness difference between L and NL populations is most evident (highlighted in blue). The most unstable environmental regime (highlighted in orange) favoured networks that were making random choices during exploitation, although one replicate was able to maintain learning episodes. The more stable environments favoured adaptive tracking (highlighted in green).

In the most unstable of environments, the extreme case of an environmental change of 0.4 every single generation in the simulation, the populations have a low fitness. The model does not allow for extinction, but the populations that find themselves foraging in such an unpredictable environmental regime would likely go extinct. In this extremely unstable environment, the individuals seem to be choosing cues randomly. The populations in the random choice (or beet-hedging corner) obtained performance around 0.26 (Supplementary Material, Figure S2).

In some intermediate environmental regimes, we can see some replicate populations obtaining a higher fitness than the non-learning populations. Upon closer inspection, these populations have evolved and maintained learning that allowed them to acquire higher levels of fitness in a moderately changing environmental regime. However, not all the populations evolved the same learning mechanism.

There are at least two different learning styles that evolved in the simulation. The first one is characterized by having 10 or more learning episodes and a low learning rate, that allows the ANNs to fine-tune their preference curve and find the cues with the highest quality most often in every generation (performance of around 0.8). These ANNs also displayed weights close to 0 in the connections to the output node, which were the ones that could be adjusted by learning. Therefore, their genetically given decision curve was often more of a flat line that started to shift as learning episodes occurred (Supplementary Materials, Figure S4). I decided to call these ANNs as fine-tuned learners. The second strategy that evolved is characterized by having two or three learning episodes and a high learning rate, sometimes reaching extreme values close to 10. These networks are characterized by having a decision curve that is capable of making great shifts in a single learning episode due to their large learning rate (Supplementary Material, Figure S3). This learning style is not as precise as the fine-tuners and usually performs worse (performance around 0.5) but it would perform better than the rest of individuals in a population of random pickers in moderately changing environments (Figure 5). In theory, this strategy would be outcompeted by fine-tuned learners.

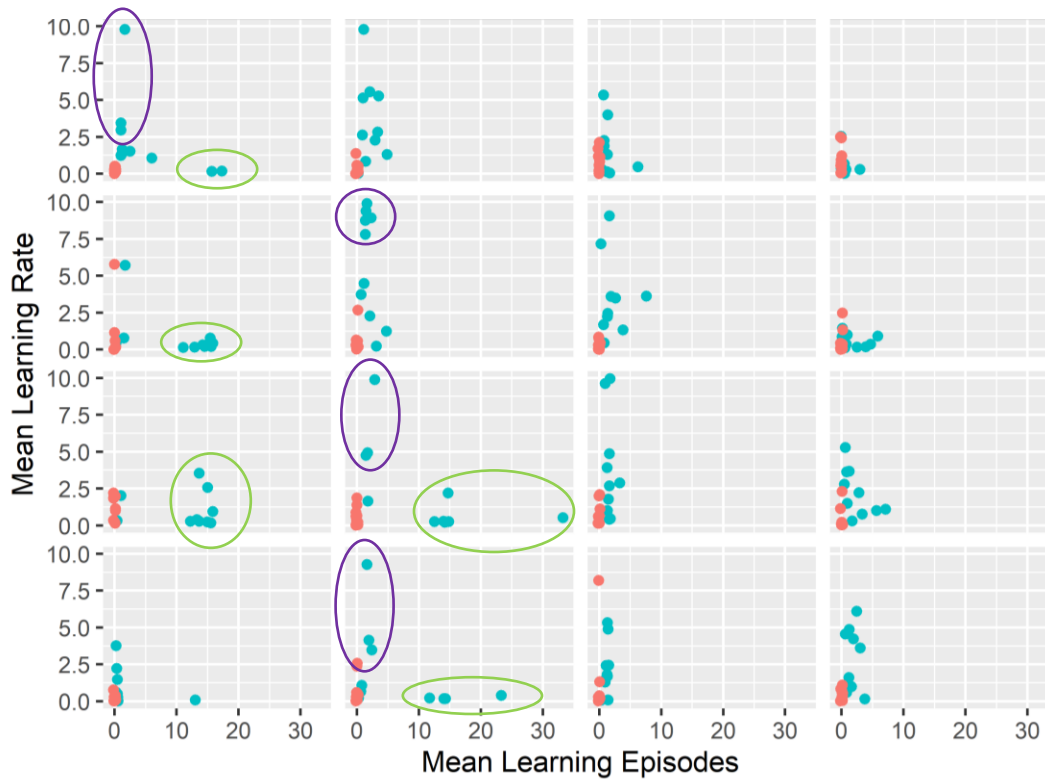


Figure 5. Evolved mean learning rate and mean learning episodes show a negative correlation in environmental regimes that favour the evolution of learning. The rows represent the different magnitudes of the environmental shifts (0.1, 0.2, 0.3, and 0.4, from top to bottom) and columns represent the different frequencies of environmental change used in the simulations (every generation, 1/5 generations, 1/25 generations and 1/125 generations, from left to right). The plot shows combinations of mean learning rate and mean learning episodes that evolved in different environmental regimes. A larger number of learning episodes seems to be coupled with lower learning rates. Also, replicates with fewer learning episodes seem to evolve towards a higher learning rate. These combinations seem to correspond to two learning styles: fine-tune learners (green ellipses) and reckless learners (purple ellipses).

In the case of the environmental regime with a 0.3 shift in the environmental mean every 5 generations, we can find a great diversity of strategies. Here we find both learning styles, but also one population that is performing adaptive tracking (blue dot outside the marked learning styles in Figure 5). This shows that some environmental regimes can allow for several strategies to occur, but upon closer inspection, the populations themselves do not host diverse strategies within them. Despite that, the strategy that is consistently performing better in the environmental regimes with intermediate change is the fine-tuned learning, followed by reckless learning and finally the bet-hedgers, across all unstable environments where these strategies evolve.

Experiment 2: The effect of reduced population size

Based on the previous simulations, I identified the environmental regime where the environmental mean shifts 0.3 in the environmental axis every 5 generations as one that allowed the evolution of both learning styles. I selected the fittest network for each replicate during the last generation of the simulation. I obtained a total of 10 networks. These individual ANNs were used to create 5 replicates of a new

monomorphic population of different sizes: 50, 100, 250, 500 and 1,000. This means that each population consisted of N individuals with the same ANN, learning rate, and learning episodes, which came from one replicate of the previous simulation. These populations were then allowed to evolve for another 50,000 generations in the same environmental regime.

By the end of the simulation, I checked whether the learning mechanism was maintained and if their fitness was similar to that of the populations that remained in their constant size of 1000 as a control. I obtained the mean fitness and learning episodes in the last 5,000 generations of the second phase of the simulation for this.

The effects observed on the mean fitness obtained by the different replicates was variable between strategies. Fine-tuned learners achieved a higher mean fitness at the end of the simulation than reckless learners, as it had been observed before.

However, some replicates experienced a reduction in the mean fitness when experiencing harsh reductions of population size to 50 or 100 individuals. Some replicates even experienced it at a population size of 250, but the prevalent pattern is that the acquired mean fitness is capable of being sustained even in such low population sizes for some of the replicates (Figure 6).

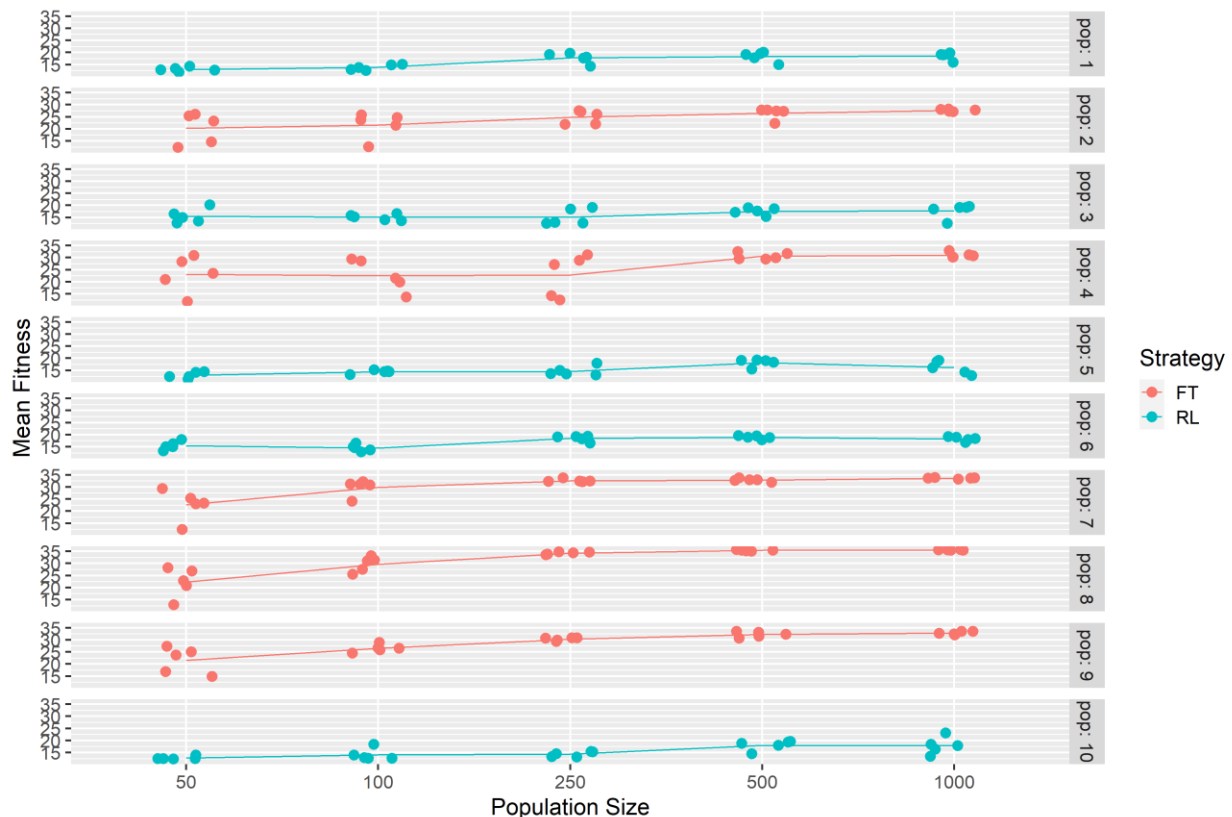


Figure 6. Mean fitness at the end of a bottleneck period for monomorphic populations of fine-tuned learners (red) and reckless learners (blue). Permanent and harsh declines in population size cause the loss of mean fitness for both learning styles. Fine-tuned learners (in red) and reckless learners (in blue) experience a decline in mean fitness when their populations are decreased to 100 individuals or less. Some

replicates are able to maintain relatively stable fitness values, but a general decline is observed in the mean (plotted with coloured lines).

The fine-tuned learners did not lose the large number of learning episodes in most of the replicates (Figure 7), which means that the learning mechanism was maintained, but it could not maintain its performance. In some rare instances, reckless learner populations seemed to evolve to have a higher number of learning episodes, but this did not result in higher fitness (see one of the replicates in population 1 and 3 in Figure 7).

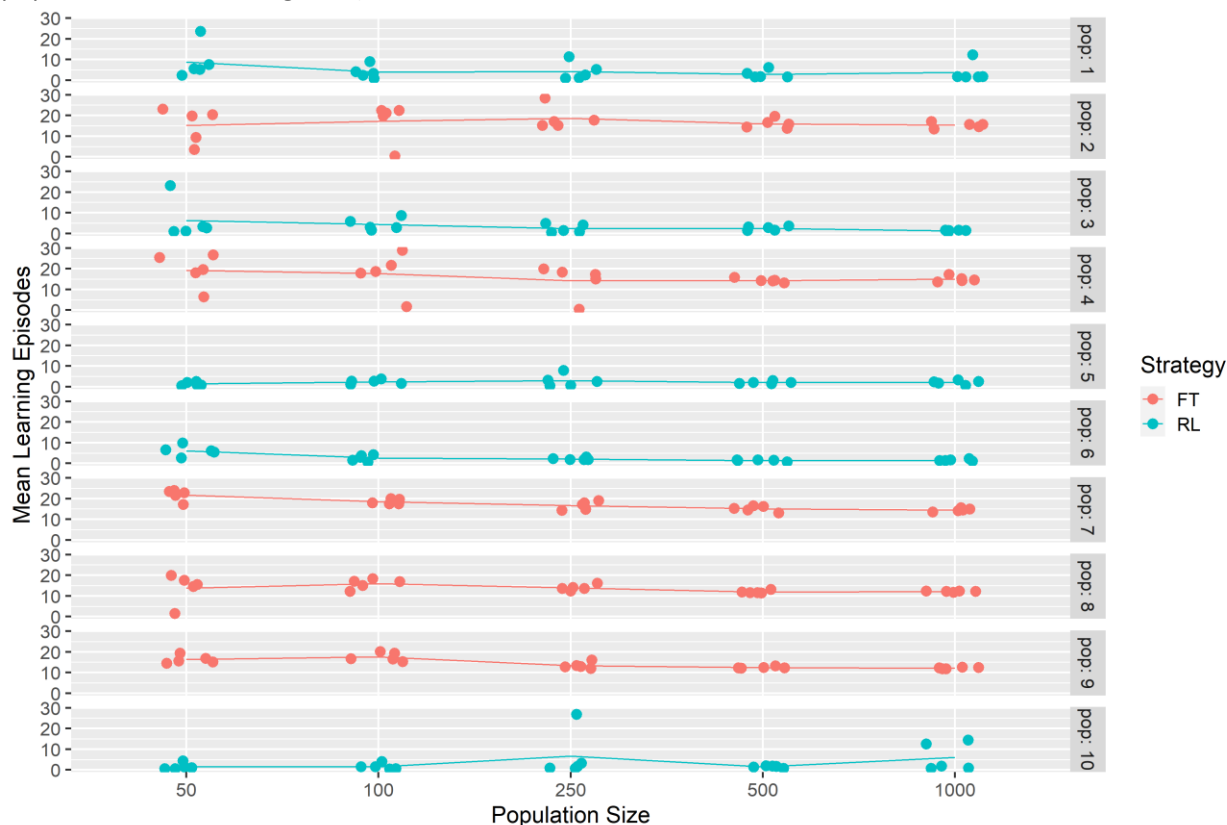


Figure 7. Mean learning episodes at the end of a bottleneck period for monomorphic populations of fine-tuned learners (red) and reckless learners (blue). Reductions in population size do not change the number of learning episodes associated to different learning styles. Despite loss of mean fitness, most fine-tuned learners (in red) maintain a high number of learning episodes while most reckless learners (in blue) maintain a low number of learning episodes.

This reduction in population size was tested in three scenarios. The previous results come from simulation with relaxed constraint weights (possible weight range from -10 to 10). However, I also tested the same experimental design for ANNs that evolved with their weights constrained between -1 and 1 and the same mutation rate. The third scenario took the same resulting individuals with constrained weights but increased the mutation step size in weights from 0.001 to 0.01 in a Cauchy distribution and allowed the weights to take values between -10 and 10 in the second phase of the simulation. These conditions resemble those that occur during stress-induced mutation. The results of these two additional scenarios can be found in Supplementary Materials. The stress-induced mutagenesis scenario showed that reckless learners would suffer mutational meltdown under at the early stages of the simulation (see Figure S6)

harsh bottlenecks and even at constant populations size. Fine-tuned learners did not experience such a mutational meltdown but were robust to some slight population bottlenecks. This shows a difference of robustness to population bottlenecks and stress between both learning styles.

Experiment 3: Evolution of learning under recurring bottleneck events

To understand the effect of recurring bottlenecks in the evolution of a complex trait such as a learning style in my model, I implemented different cyclic bottleneck regimes in the same environmental regimes than in experiment 1 (described above). I introduced bottleneck events with different duration, frequency, and size after a stable period of 10,000 generations that allowed the populations to evolve without disruptions and then the populations were allowed to evolve for another 40,000 generations under the bottleneck treatments. Three scenarios were run: first, the population was reduced for 100 generations every 1,000 generations; second, the population was reduced for 50 generations, every 500 generations. The third scenario reduced the population permanently after the stable period. I present the results of a population reduction to 50 individuals, which was the harshest reduction I explored. A less extreme bottleneck of 100 is presented in Supplementary Materials (Figure S10).

As with the first round of simulations in this setup, I obtained the mean fitness, learning episodes and learning rates, and performance for the last 5,000 generations of the simulation, assuming the populations had reached a stable state. This allowed me to identify the final strategies evolved by each replicate population in each environmental parameter combination and check if this was consistent or not with the evolution of populations that did not experience bottlenecks. However, closer examination across the simulation time allowed me to identify the effect of the bottleneck events, analysing the state of the population strategies before and after experiencing bottlenecks.

None of the simulations that experienced cyclic bottleneck treatments seemed to behave substantially different to the simulations where no bottlenecks occurred. Some replicates seem to not follow this general pattern. Some of the replicates in the environmental regime of a shift of the environmental mean of 0.2 or 0.3 every generation with the most frequent bottlenecks show more replicates with low fitness (Figure 8). However, observing the evolution of these replicates showed that the learning mechanisms were not able to evolve throughout the simulation, instead of being lost due to cyclic bottlenecks. For the rest of strategies, the apparent effect of the cyclic bottlenecks seems to be negligible at the end of the simulation period. All previously observed strategies are still found despite the bottleneck events and the variability in the outcomes of each replicate seems to follow the expectations provided by the populations that evolve at a stable size in most cases.

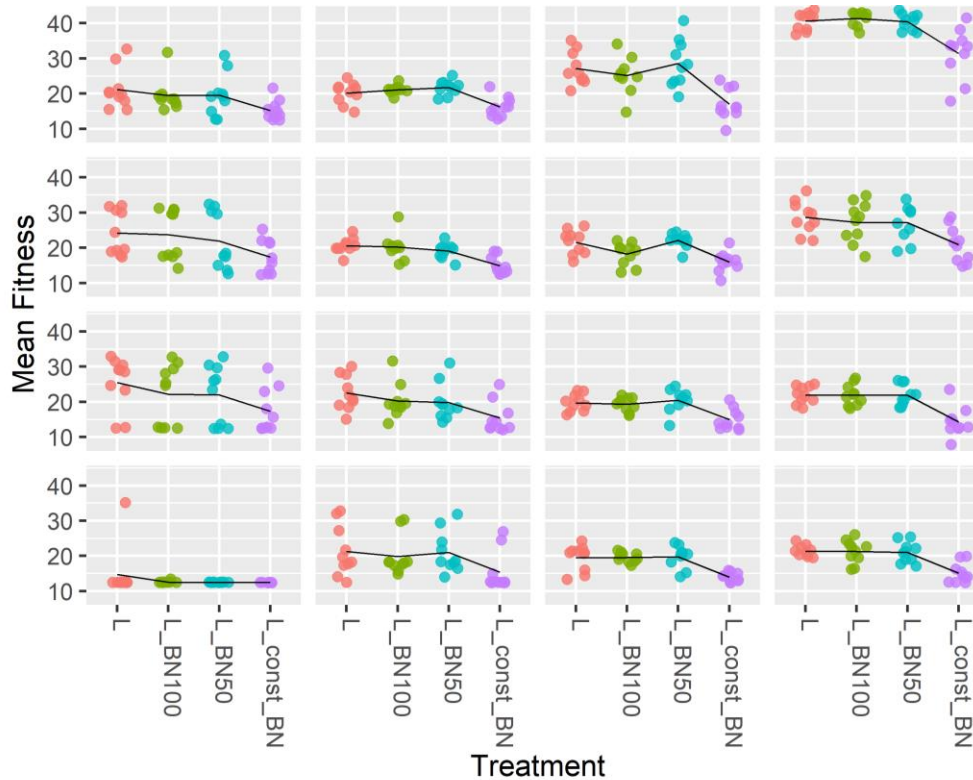


Figure 8. Mean fitness of learning replicates that evolved at constant population size and different bottleneck treatments. The rows represent the different magnitudes of the environmental shifts (0.1, 0.2, 0.3, and 0.4, from top to bottom) and columns represent the different frequencies of environmental change used in the simulations (every generation, 1/5 generations, 1/25 generations and 1/125 generations, from left to right). Cyclic bottlenecks do not hinder the evolution of learning mechanisms in the environmental regimes where it is favoured, i.e. cyclic bottleneck treatments (in green and blue) evolved similarly to the populations under stable population size (red/orange). The populations that experienced a constant bottleneck after a stable period of 10,000 generations (purple) did experience a decline in mean fitness.

When analysing the effect of constant bottlenecks, it is evident that the populations under this treatment were the ones that achieved the lowest fitness across environmental regimes. Not only the populations in environmental regimes that favoured the evolution of learning experienced lower fitness levels, but also populations that had evolved in environmental regimes that favoured adaptive tracking seemed to be negatively affected (Figure 8). The first case that is worth noticing is the environmental regime with shifts of 0.1 in the environmental mean every 25 generations. Here, most of the replicates are doing worse than the ones that are experiencing a cyclic bottleneck or the ones experiencing a stable population size. Upon closer examination, these replicates also evolved a larger number of learning episodes (Figure 8). I also ran simulations where learning was not allowed to evolve, and these populations exhibited better fitness than the ones that were able to learn (Figure 9). Another environmental regime that showed an interesting pattern with constant bottlenecks was the most stable one, with shifts of 0.1 every 125 generations. Here, most of the replicates under a constant bottleneck performed much worse than all other populations in all other treatments evolving in this environmental regime, but they did not evolve any learning episodes, with the exception of one replicate (Figure 8).

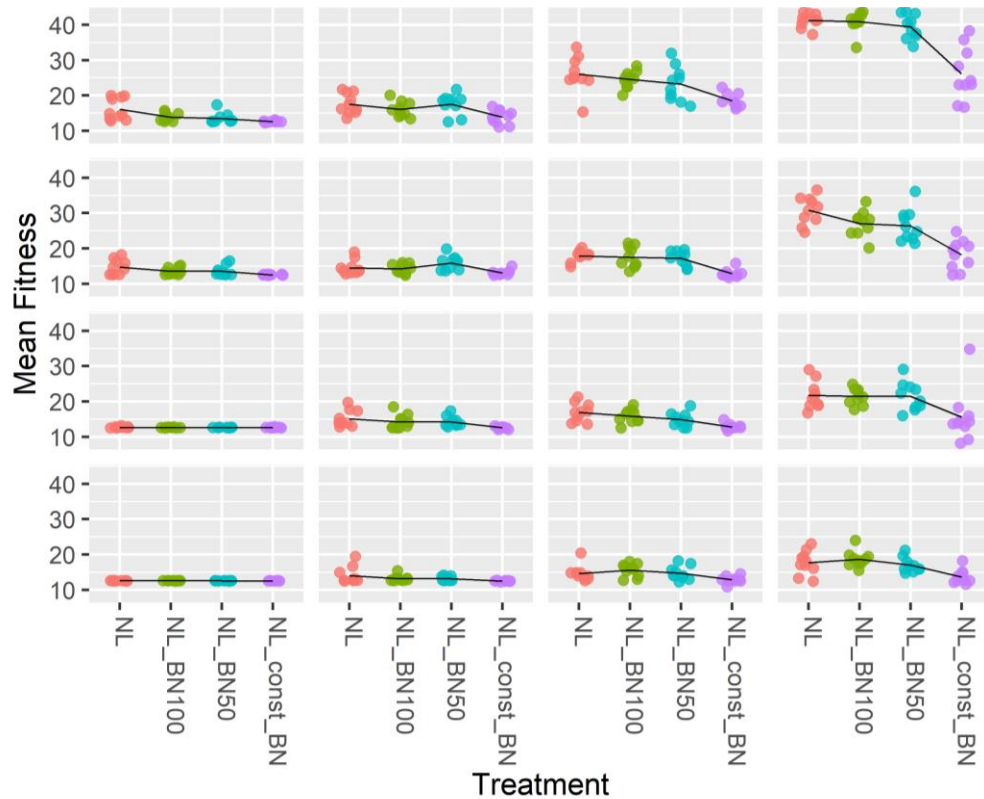


Figure 9. Mean fitness of non-learning replicates that evolved at constant population size and different bottleneck treatments. The rows represent the different magnitudes of the environmental shifts (0.1, 0.2, 0.3, and 0.4, from top to bottom) and columns represent the different frequencies of environmental change used in the simulations (every generation, 1/5 generations, 1/25 generations and 1/125 generations, from left to right). Non-learning control for the cyclic bottleneck treatments evidence decline in the mean fitness in some replicates, but overall similar results than populations evolving at a constant size for most of the environmental regimes. Populations evolving under a constant bottleneck after the stable period of 10,000 generations shows a decline in mean fitness, especially for regions where adaptive tracking is favoured.

I plotted the mean fitness gained by the different replicates in all environmental regimes against the mean learning episodes that they had evolved to have to better understand the prevalence of learning episodes in the replicates under constant bottlenecks in regions that favour adaptive tracking (Figure 10). This revealed that for environmental regimes with intermediate change, which favour the evolution of learning mechanisms, learning episodes usually have a positive correlation to fitness gain. In the case of replicates that have evolved and maintained learning episodes in stable environmental regimes that favour adaptive tracking, learning episodes show the opposite trend on fitness gain, although there are a few replicates that seem to gain as much fitness as those populations evolved without learning. The clear pattern that becomes evident is that having a large number of learning episodes in a stable environment is not beneficial.

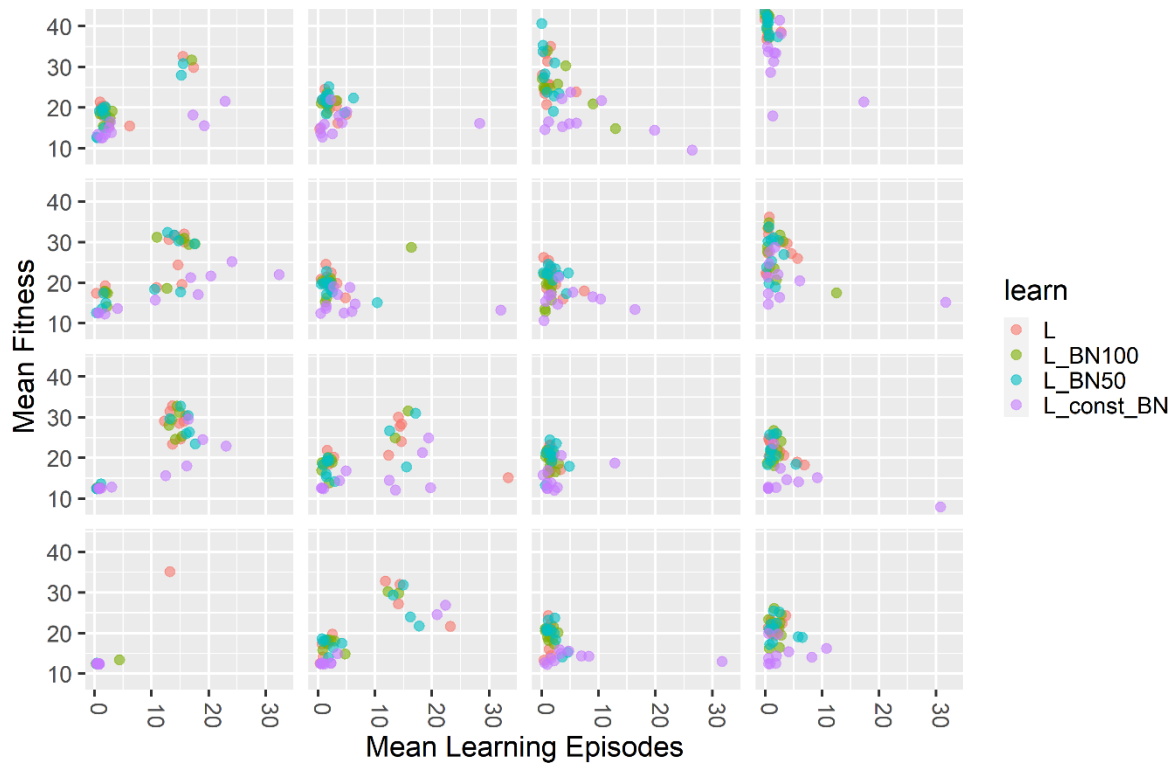


Figure 10. Mean fitness versus mean learning episodes of learning populations evolved under constant population size and different bottleneck treatments shows the contribution of learning to population fitness. Environmental regimes found to favour learning maintain a positive trend between mean fitness and mean learning episodes. The constant bottleneck treatment reveals that even with many learning episodes, these populations are unable to gain fitness at the level of populations that have evolved with cyclic bottlenecks or stable size in the environmental regimes that favour learning in most replicates. Some of the replicates under a constant bottleneck maintain high levels of fitness as it was also observed during Experiment 2. Some of the populations that are found in environmental regimes that favour adaptive tracking evolved to have a large number of learning episodes, and this hinders their fitness.

The pattern observed in Figure 10 also shows a variability even within learning styles, environmental and populations bottleneck regimes. As described before, fine-tuned learners have a large amount of learning episodes, but the fitness gain that they can achieve per learning episode is different even for replicates under the same environmental regimes. This variability may be observed in the populations evolving at a constant size, but also for the cyclic bottleneck treatments, which might make it hard to conclude that the cyclic bottlenecks cause a lower performance. However, the clearest pattern observed is again a degradation of the fitness gained by different learning styles under a constant bottleneck. Some populations under a constant bottleneck still seem to maintain a fitness gain similar to populations under different bottleneck treatments, but they are generally in the lower values of the fitness axis. For the environmental regimes of shifts of 0.3 in the environmental mean every five generations and every generation, the replicates under a constant bottleneck show a lower fitness gain even when having more or the same amount of learning episodes as replicates that evolved with a stable population size and with cyclic bottlenecks.

To identify if the learning styles were indeed degrading or simply did not evolve in some of the replicates, I explored the evolution of the populations throughout the simulation. I focused on all the bottleneck treatments in two environmental regimes: a shift in the environmental mean of 0.3 every five and every single generation.

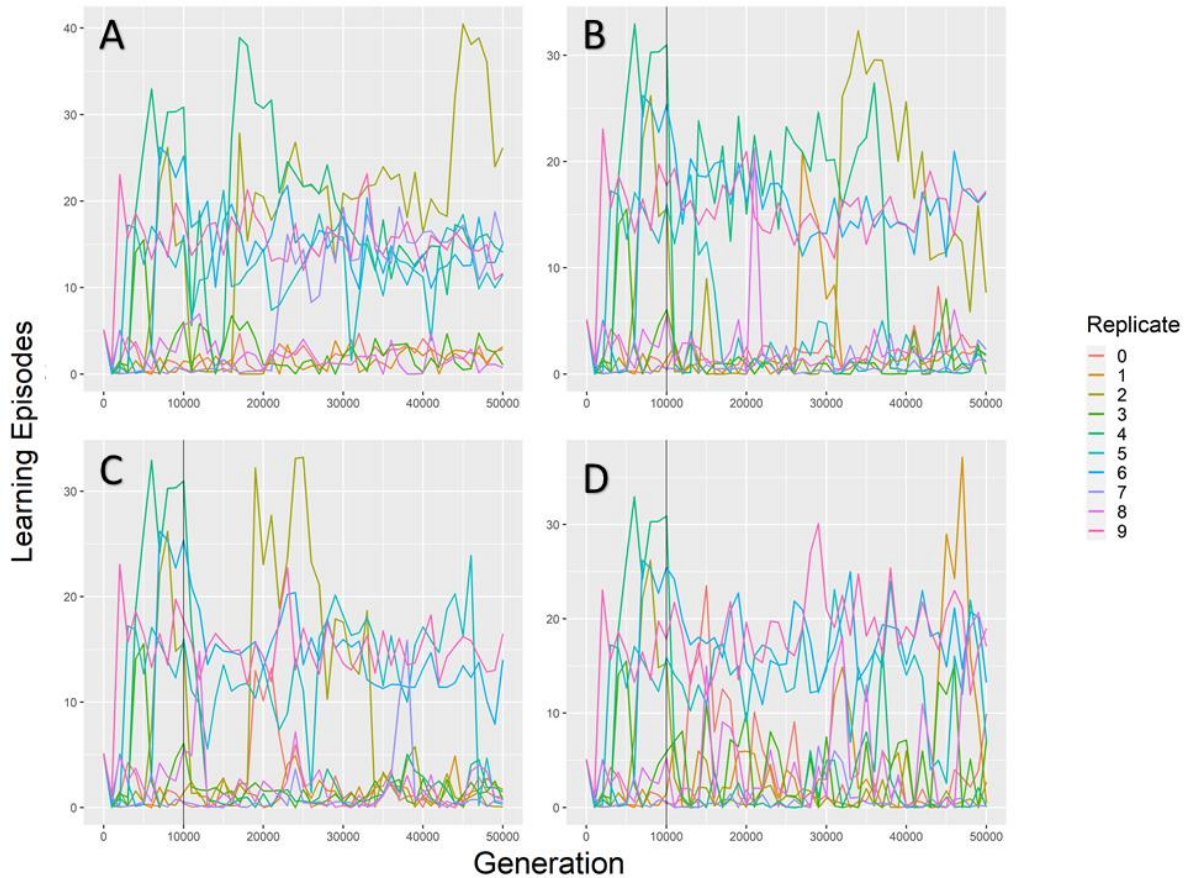


Figure 11. Time-course plots of learning episodes for replicates evolving under constant population size (A) and different bottleneck treatments (B-D) experiencing environmental shifts of 0.3 every five generations. The plot shows fluctuations in the number of learning episodes for all populations, but clustering of both learning styles seems to be maintained despite cyclic bottlenecks. Learning episodes can behave erratically in the simulation even at constant size (A). When comparing the bottleneck treatment of reductions every 1000 generations (B) and reductions every 500 generations (C), these fluctuations seem to be maintained. Fluctuations are also observed in the constant bottleneck scenario (D).

In the case of the environmental shifts of 0.3 every 5 generations (Figure 11), I could observe fluctuations in the amount of learning episodes through the simulations, but some replicates would maintain a constant fine-tuner strategy. Some replicates would also fluctuate among very few learning episodes, and some had occasional spikes that would finally decay. When comparing to the simulations where no bottleneck events occurred, I could identify similar fluctuations. And as seen previously, some replicates maintain a high number of learning episodes under the constant bottleneck, but at the end of the simulation, this did not necessarily translate into a larger fitness. The observed fluctuations for the cyclic bottlenecks does not seem to be substantially different from the simulations with stable population size.

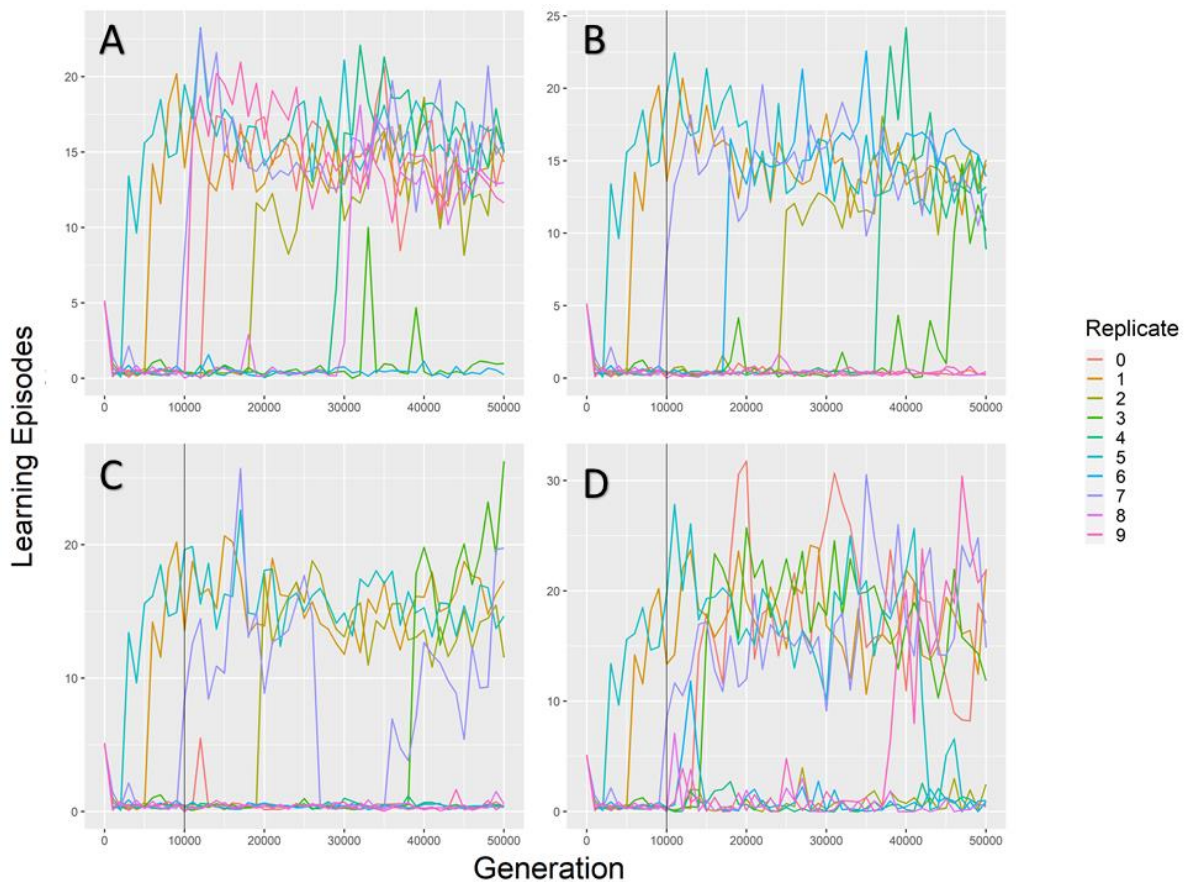


Figure 12. Time-course plots of learning episodes for replicates evolving under constant population size (A) and different bottleneck treatments (B-D) experiencing environmental shifts of 0.3 every generation. The plot shows fluctuations in the number of learning episodes. The predominant learning style that seems to evolve in this environmental regime are fine-tuned learners. Learning episodes can behave erratically in the simulation even at constant size (A), but always remain high. When comparing the bottleneck treatment of reductions every 1000 generations (B) and reductions every 500 generations (C), these fluctuations seem to be maintained, but it is clear that a large number of learning episodes are strongly selected for. The pattern remains even during the constant bottleneck treatment (D), with only one replicate out of ten losing the fine-tuned strategy.

In the case of environmental shifts of 0.3 every single generation (Figure 12), the evolution of fine-tuned learning is favoured as the most robust learning style, while reckless learning does not evolve. The replicates that end the simulation without learning were not able to acquire it at any point during the simulation and those which did acquire it and maintain it in all cyclic bottleneck treatments. Two replicates even evolved fine-tuned learning during the constant bottleneck period (pink and green). Only one replicate lost completely their fine-tuned learning during the constant bottleneck after acquiring a large amount of learning episodes during the stable period. This confirms that cyclic bottlenecks are rarely able to degrade completely the learning mechanisms evolved during these simulations.

IV. Discussion

Environmental conditions that benefit the evolution of learning

The first experiment has allowed me to confirm that learning has evolved in similar environmental regimes as plasticity did in the Botero *et al.* model, even though Botero *et al.* modelled the evolution of reaction norms, while I have used an ANN that can evolve learning to adapt to novel environmental conditions. Stable regions of the environmental parameter space that I explored favoured networks that tracked the environmental peak through mutation. The most unstable environmental regime was the only one that favoured random choice as analogous to bet-hedging (Figure 3). Other environmental regimes of intermediate stability favoured the evolution of learning and actually helped the population obtain higher level of fitness when compared to replicate populations that could not evolve learning. Therefore, the evolution of learning in this model appears to follow the general principles of the evolution of phenotypic plasticity in changing environments.

The determinant factor to evolve and maintain learning was how the trade-off between learning episodes and the exploitation period played out. If the benefit in performance and fitness gain obtained from a number of learning episodes was greater than the cost of investing that time to learn and not to exploit the environment, learning would be selected for. Therefore, we find the evolution and maintenance of learning mechanisms in environmental regimes of intermediate stability where the correlation between learning episodes and fitness is positive (Figure 10). However, this could be the case in other parameter combinations, but selection will favour learning if the trade-off between learning and exploiting is compensated by fitness gain.

The results I obtained for the environmental conditions that favour the evolution of learning highly resemble the results of Botero *et al.* (2015) for the evolution of phenotypic plasticity. This link suggests that plastic responses of variable complexity in their genetic architecture might evolve under the same environmental conditions.

Evolution of different learning styles

The closer inspection of the populations that evolved learning revealed two learning styles that I called fine-tuned learners and reckless learners. These seem to be two different attractor states that emerge in the environmental regimes that favour the evolution of learning. Generally, as learning styles that evolved in the same environmental regimes, they seemed to have a better performance than non-learning populations but differ considerably. The most striking difference was that fine-tuned learners maintain values very close to zero in the last four weights, which are the ones modified by learning, and had generally lower weight values in all the network connections, coupled with a low learning rate. In comparison, reckless learners had far fewer learning episodes, a larger learning rate, but also generally higher values in their weights.

This difference in the network configuration in the two strategies also was reflected in the behaviour of the learned expectation curves. Reckless learners displayed an impressive ability to modify greatly their expectation curve after two or three learning episodes due to the large value of their weights and their large learning rate. The shape of the inherited decision curve in some of these networks was similar to a sigmoid function and could flip easily the higher expectations from one side to the other with a single cue. Despite this impressive malleability of their decision curve, their performance was considerably lower

when compared to that of fine-tuned learners (around 0.5 vs. 0.8). Therefore, reckless learners can shift their decision curves greatly and quickly but lack the precision that fine-tuned learners have.

The emergence of two learning styles in this model suggests the evolution of several possible solutions. This also implies that there are several peaks in the fitness landscape given by each of the changing environmental regimes that favour the evolution of learning. The most likely scenario is that the populations find one of these fitness peaks and obtain an immediate fitness benefit that allows the learning mechanism to quickly fixate in the population. In this case, entire populations that have not evolved learning might find themselves in a fitness valley and either of the solutions would represent a fitness peak, as any learning style will perform better than random choice. Therefore, populations quickly fixate into one of the solutions, probably the one that is reached first given the initial conditions of the simulation. However, the number of learning episodes, the learning rates, and the characteristics of the network that each strategy requires is very different, and transitions from one strategy to the other within one population is unlikely without a period of lower fitness or a generation where the other learning style outperforms the dominating one in such a way that the population flips from one learning style to the other. This would occur if an alternative learning style existed in the population at lower frequencies, but I was unable to record the state of the population with enough resolution to confirm that this is the case. It is likely that in the final state of the simulation, one of the learning styles will have been fixated in the population.

Robustness in the evolved learning mechanism

The main questions of this study focus on the robustness of complex traits when facing several challenges. However, robustness remains an unclear term and it can refer to several properties of a complex trait. Here, I will discuss the robustness to environmental change as the ability of an individual or population to perform well even when facing environmental variation, which corresponds to the environmental change regime in my simulations. I will refer to robustness to population bottlenecks as the ability of a population to maintain a learning mechanism that performs well even when the population faces cyclic or permanent reduction in size (second and third experiments).

The environmental robustness of the populations in the simulation is given by their ability to exploit the correct cues in their environment despite the shifts in the environmental function over time. Therefore, a good way to quantify the environmental robustness of the populations that have evolved is through their performance as the proportion of the maximum fitness gain that they are able to obtain during the exploitation phase of their lifetime, having a maximum value of 1 if the individual always chooses the cue with maximum quality associated to it. The maximum possible fitness gain is given by the cues available during the lifetime. Performance is similar to other measurements used to quantify the adjustment to environmental challenges in other models that explore complex traits (see Stewart *et al.*, 2012). This metric allows us to tease apart both learning styles, since the fine-tuned learners perform best, achieving a performance that can exceed 80% of maximum achievable fitness in almost all replicates. In contrast, reckless learners achieve in average a performance of around 50% of the maximum achievable fitness during exploitation. This difference between the two strategies is considerable, but it exceeds the 26% of the maximum achievable fitness obtained if choice is made at random during the whole exploitation phase when learning does not evolve in the population. Both learning styles perform better than non-learners in changing environments and the fitness gained during their lifespan, even accounting for the

learning/exploiting trade-off built into the model, will cause the learning mechanisms to be selected for. This means that even though they are not able to exploit the environment for as long, the populations that have evolved learning outperform the ones that did not evolve learning. Also, the benefit in performance from fine-tuned learning seems to allow these populations to achieve a mean fitness gain considerably higher than that of the reckless learners even with a shorter exploitation period. This is true only in moderately unstable environments, as was shown in Experiment 1. The trade-off between learning and exploitation can lead to strategies like bet-hedging and adapting tracking in extremely unstable or more stable environments, respectively.

In this study, mutational robustness can be defined as the effect of a mutational step in the overall network output. Networks with larger weights will change their output the least from a mutation, while networks with lower weights will experience a greater impact from the same mutation (Visser *et al.*, 2021). If we take this definition of mutational robustness in networks, there is a difference in mutational robustness between reckless learners and fine-tuned learners. Reckless learners usually evolve to mean absolute weight values of around 4 or 5 in four of the selected populations for Experiment 2, with the lowest value in one replicate reaching values of 3. In contrast, fine-tuned learners maintain lower mean absolute weights without exceeding values of 2, except for one replicate that reached values of 3. This led me to conclude that reckless learners exhibit a higher mutational robustness in their networks when compared to fine-tuned learners.

Given these definitions of mutational and environmental robustness, what happens during bottleneck events appears to be counterintuitive. I observed that reckless learners would accumulate mutations in earlier stages of the simulation and would be overall more susceptible to degradation when compared to fine-tuned learners. If the decrease of population size and increase in genetic drift increases the chances of expressing deleterious mutations in the population, fine-tuned learners would be expected to degrade faster than reckless learners since their mutational robustness is lower. However, the opposite was observed, and this seems to hint that the robustness to population bottlenecks is driven by environmental robustness, and not mutational robustness. Not only these results suggest that the effect of population bottlenecks is more complex than suggested from models that assume simple traits in mutation-drift-selection equilibrium, but also hint that environmental and mutational robustness are not correlated in this model, as observed in other studies (Ciliberti *et al.*, 2007a). An example of complex networks that fine-tune biological processes, but are susceptible to deleterious mutations, is the transcription factor network involved in neural crest development where haploinsufficiency can lead to disease (Paratore *et al.*, 2022).

One of the expected outcomes of the simulations I ran was finding several network configurations that could perform the same learning style. This is an outcome that has been observed previously during simulations of gene regulatory networks that explore the properties of the network robustness. The set of network configurations separated by one mutation that produce the same phenotype are called a neutral network and robustness can be distributed through this neutral network (Wagner *et al.*, 2009; Ciliberti *et al.*, 2007a; Ciliberti *et al.*, 2007b; Wagner, 2008; Krishan *et al.*, 2008). These variants of the genetic configuration of the trait can be mapped as interconnected nodes in a metagraph of network configurations that can produce the same mechanism (the phenotype in question at least for this section) and are separated by a mutational step from each other in the genotypic space. Therefore, it would be

theoretically possible to traverse the neutral network through gradual mutations while maintaining the phenotype.

In my study, strong selection seems to act directly on the environmental robustness of the learning mechanism, which means that robustness differences can be key to the evolution of the population. In the case of environmental regimes where change was frequent and large, strong selection for environmental robustness was selected for, favouring the evolution of learning mechanisms. Therefore, selection will favour the most environmentally robust configuration of the network, resulting in the conservation of the genetic configuration of the most robust network across the population (Ciliberti *et al.*, 2007a). This allowed me to expect the fixation of a highly robust configuration of the neural network capable of learning, only being replaced with more robust configurations that could arise through mutation.

Cyclic bottlenecks can allow a relaxation of selection, which allows the exploration of the mutational neighbourhood of the network that is predominant in the population during the simulations. This would allow alternative network configurations to be explored by mutation. With increased drift, this could either lead to suboptimal configurations of the network that could be fixed in the population if they are unable to be eliminated by selection, or the exploration of mutations that traverse a fitness valley and reach another fitness peak that allows the evolution of alternative robust networks. My observations on the effect of cyclic bottlenecks seem to confirm that the robustness of the trait is maintained. The exploration of the mutational neighbourhood of the neutral network during temporary bottlenecks due to relaxed selection might be less likely, but the resolution of our data collection is not able to determine if this is the case. Whichever is the actual trajectory of the genetic configuration of the trait in the simulations with cyclic bottlenecks, the result was a robust learning mechanism that evolved and was maintained through the population contractions and expansions.

The parallel between the environmental regimes that allowed learning to evolve in my model and that of the environmental regimes that allow for the evolution of phenotypic plasticity support the extension of my analysis to some cases of costly phenotypic plasticity (Botero *et al.*, 2015). If this is the case, robustness to population bottlenecks could be found for complex regulatory or developmental processes, for phenotypic plasticity, or any other adaptive response that requires a complex genetic architecture, and therefore, could be expected to be robust when facing long and harsh bottlenecks. This could possibly explain the contrasting results found by studies like the ones from Auld and Relyea (2010), or Luquet *et al.* (2011), with some populations maintaining complex traits in small and isolated populations that have suffered genetic erosion, instead of degrading as generally expected.

The decay of highly robust traits might come after prolonged and harsh bottlenecks in terms of number of generations, which could increase the risk of extinction of species with typically shorter lifespans, as they might experience more generations during the same time as other organisms with longer lifespan might have experienced. However, organisms with shorter lifespans, such as some insects or other invertebrates, could be expected to be found in higher effective population sizes. In the end, the interplay of several factors in the population could be considered to measure the risk of extinction in a given population, but my results reinforce the well-observed fact that the most effective measure would be to

increase population size to reduce the influence of genetic drift and avoid the degradation of its adaptive mechanisms (Frankham *et al.*, 2017).

Further research in a theoretical framework could analyse the evolution of complex traits under relaxed selection to possibly observe alternative genetic configurations that maintain a similar learning mechanism. Increasing the SD of the environmental function from 0.1 to 0.4 was observed to allow more weight variation, which probably is the effect of relaxed selection for the environmental robustness of learning. Also, extending the model to a diploid network could reveal the behaviour of a complex trait during bottleneck events and tease apart the mechanistic underpinnings of inbreeding depression. Evolvability of the networks that have evolved in the simulations presented here should also be explored, especially in cases of shifts in environmental regimes. Stronger selection and a resulting higher robustness could lead to higher evolvability of the trait and in this context, the effect of cyclic bottlenecks might be relevant, as it could theoretically allow the exploration of a “flatter” fitness landscape in periods of increased drift. Finally, further research could produce empirical evidence of traits that involve networks through the study of gene expression profiles in plastic traits or in known gene regulatory networks in general (Chevin *et al.* 2021). Piecing together the genetic architecture of complex traits might be highly informative when analysing the viability of endangered populations, when monitoring the effectiveness of conservation measures and when considering the possibility of an evolutionary rescue.

V. Conclusions

The evolution of learning is favoured in environmental regimes with intermediate instability. This finding suggests that learning might evolve in similar environmental conditions as adaptive phenotypic plasticity, allowing this model to be extended to other plastic traits that might behave similarly.

The most robust strategy that evolved was the fine-tuned learning, which was characterized by having small learning rate and network weights and a larger environmental robustness. However, this contrasts with the expectations that the most robust strategy in bottleneck events to be the most mutationally robust, as previous findings suggest that mutational and environmental robustness are expected to be correlated (Ciliberti *et al.*, 2007a, Visser *et al.*, 2021). This points to different mechanisms that might drive environmental and mutational robustness in certain traits, suggesting that underlying mutational robustness of traits can be achieved by selection for environmental robustness, but not necessarily through their correlation. This link is shown especially by the robustness of the learning mechanisms to bottleneck events in experiments two and three.

The complex architecture of the learning mechanism that evolves in this model was shown to resist long periods of reduced population size and increased drift, suggesting that some regulatory mechanisms might not degrade as fast as other traits in small and isolated populations. But when they do, the mechanism is not degraded as a whole, but rather adopts sub-optimal configurations that decrease fitness in different degrees after long and harsh population bottlenecks. The learning mechanism that evolves with cyclic bottlenecks were not seen to degrade like those in constant population reductions, highlighting the importance of conservation interventions to increase population size and decrease the effect of drift even when further information on the genetic state and fitness of the population is insufficient.

VI. Literature Cited

- Auld, J.R. & Relyea, R.A. 2010. Inbreeding depression in adaptive plasticity under predation risk in a freshwater snail. *Biol. Lett.* **6**: 222-224.
- Baldwin, M.J. 1896. A new factor in evolution. *Am. Nat.* **30**: 441-451
- Botero, C.A., Weissing, F.J., Wright, J. & Rubenstein, D.R. 2015. Evolutionary tipping points in the capacity to adapt to environmental change. *PNAS.* **112**: 184-189.
- Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M. & Palmer, T. 2015. Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Sci. Adv.* **1**: 1:e1400253.
- Chevin, L.M., Leung, C., Le Rouzic, A. & Uller, T. 2021. Using phenotypic plasticity to understand the structure and evolution of the genotype-phenotype map. *Genetica*: s10709-021-00135-5
- Ciliberti, S., Martin, O.C. & Wagner, A. 2007a. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Comput. Biol.* **3**: e15.
- Ciliberti, S., Martin, O.C. & Wagner, A. 2007b. Innovation and robustness in complex regulatory gene networks. *PNAS.* **104**: 13591-13596.
- Ehlers, A., Worm, B., Reusch, T.B. 2008. Importance of genetic diversity in eelgrass *Zostera marina* for its resilience to global warming. *Mar. Ecol. Prog. Ser.* **355**: 1-7.
- Frankham, R., Bradshaw, C.J.A. & Brook, B.W. 2014. Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol. Conserv.* **140**: 56-63.
- Frankham, R., Ballou, J.D., Ralls, K. et al. 2017. *Genetic Management of Fragmented Animal and Plant Populations*. Oxford University Press, Oxford.
- Fredrickson, R.J., Siminski, P., Woolf, M. et al. 2007. Genetic rescue and inbreeding depression in Mexican wolves. *Proc. Biol. Sci.* **274**: 2365-2371.
- Funk, W.C., Forsman, E.D., Johnson, M. et al. 2010. Evidence of recent population bottlenecks in Northern Spotted Owls (*Strix occidentalis caurina*). *Conserv. Genet.* **11**: 1013-1021.
- Johnson, W.E., Onorato, D.P., Roelke, M.E. et al. 2010. Genetic restoration of the Florida panther. *Science.* **329**: 1641-1645.
- Krishan, A., Tomita, M. & Giuliani, A. 2008. Evolution of gene regulatory networks: Robustness as an emergent property of evolution. *Physica, A.* **387**: 2170-2186.
- Masel, J. & Siegal, M.L. 2009. Robustness: mechanisms and consequences. *Trends Genet.* **25**: 395-403.
- Matthies, D., Bräuer, I., Maibom, W. & Tschranke, T. 2004. Population size and the risk of local extinction: empirical evidence from rare plants. *Oikos.* **105**: 481-488.
- Maynard Smith, J. 1956. Acclimatization to high temperatures in inbred and outbred *Drosophila subobscura*. *J. Genet.* **54**: 497-505.
- Paratore, C., Eichenberger, C., Suter, U. & Sommer, L. 2022. Sox10 haploinsufficiency affects maintenance of progenitor cells in a mouse model of Hirschsprung disease. *Hum Mol Genet.* **11**: 3075-85.
- Pazzaglia, J., Reusch, T.B.H., Terlizzi, A., Marín-Guirao, L. & Procaccini, G. 2021. Phenotypic plasticity under rapid global changes: The intrinsic force for future seagrasses survival. *Evol. Appl.* **14**: 1181-1201.
- Román-Palacios, C. & Wiens, J.J. 2020. Recent responses to climate change reveal the drivers of species extinction and survival. *PNAS.* **117**: 4211-4217.
- Stade, I.R., Navarro, L.M. & Pereira, H.M. 2019. Range size predicts the risk of local extinction from habitat loss. *Glob. Ecol. Biogeogr.* **29**: 16-25.

- Stewart, A.J., Parsons, T.L. & Plotkin, J.B. 2012. Environmental robustness and the adaptability of populations. *Evolution*. **66**: 1598-1612.
- Tallmon, D.A., Luikart, G., Waples, R.S. 2004. The alluring simplicity and complex reality of genetic rescue. *Trends Ecol. Evol.* **19**: 489–496.
- Visser, B., Alborn, H.T., Rondeaux, S., Haillot, M., Hance, T., Rebar, D., Riederer, J.M., Tiso, S., van Eldijk, T.J.B., Weissing, F.J. & Nieberding, C.M. 2021. Phenotypic plasticity explains apparent reverse evolution of fat synthesis in parasitic wasps. *Sci. Rep.* **11**: 1-13
- Wagner, A. 2005. Robustness, evolvability, and neutrality. *FEBS Letters*. **579**: 1772-1778.
- Wagner, A. 2008. Robustness and evolvability: a paradox resolved. *Proc. R. Soc. B.* **275**: 91-100.
- Wiens, J.J. 2016. Climate-Related local extinctions are already widespread among plant and animal species. *PLoS Biol.* **14**: e2001104.
- Zheng, J., Guo, N. & Wagner, A. 2020. Selection enhances protein evolvability by increasing mutational robustness and foldability. *Science*. **370**: 10.1126/science.abb5962

VII. Appendix 1: Experiment 1

Time-course plots for different simulations

The simulations ran comprise a period of 50,000 generations where the populations evolve. These population usually experience significant fluctuations in learning episodes and learning rate, which stabilize into different strategies and outcomes. A specific case I mentioned, but do not explore in depth in the main text is the adaptive tracking strategy. In the following plots, I present the effect of environmental change in fitness of the populations evolving under a regime of environmental shifts in the mean of 0.4 every 125 generations. At the beginning of the simulation (Figure S1-A), the collapse in fitness is evident immediately after the shift of the environmental mean. Later in the simulation, however, the effect is not the same. Despite the change in dynamics, the environmental shifts still seem to influence the fitness of the populations.

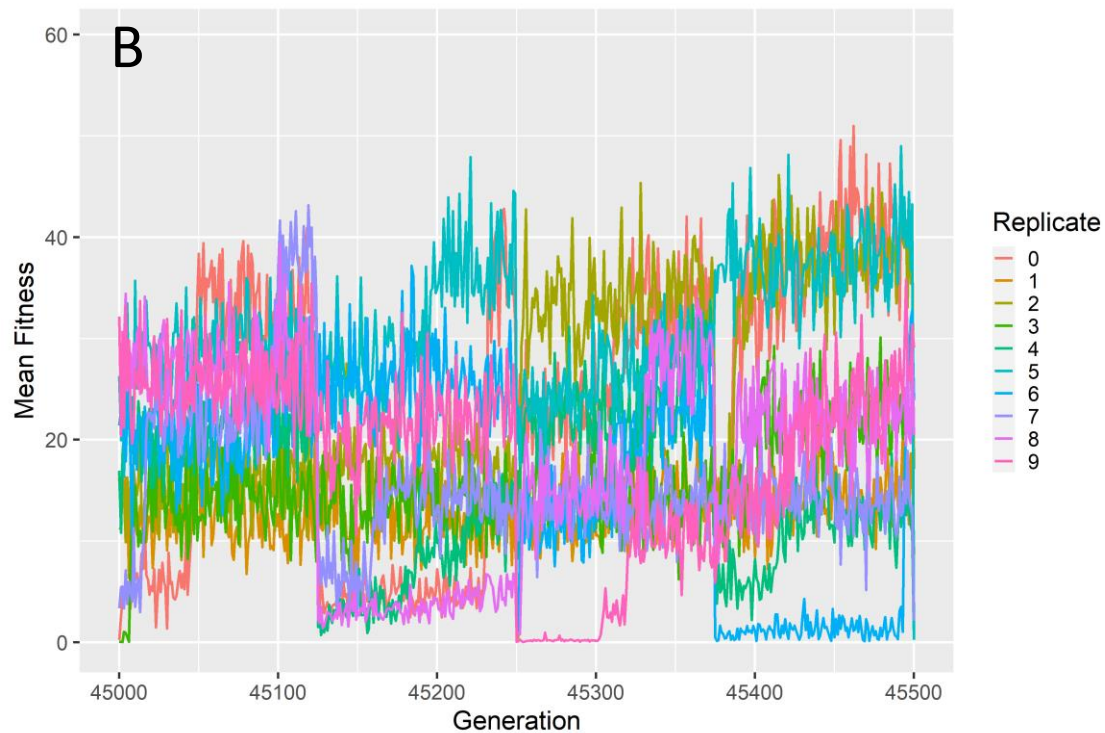
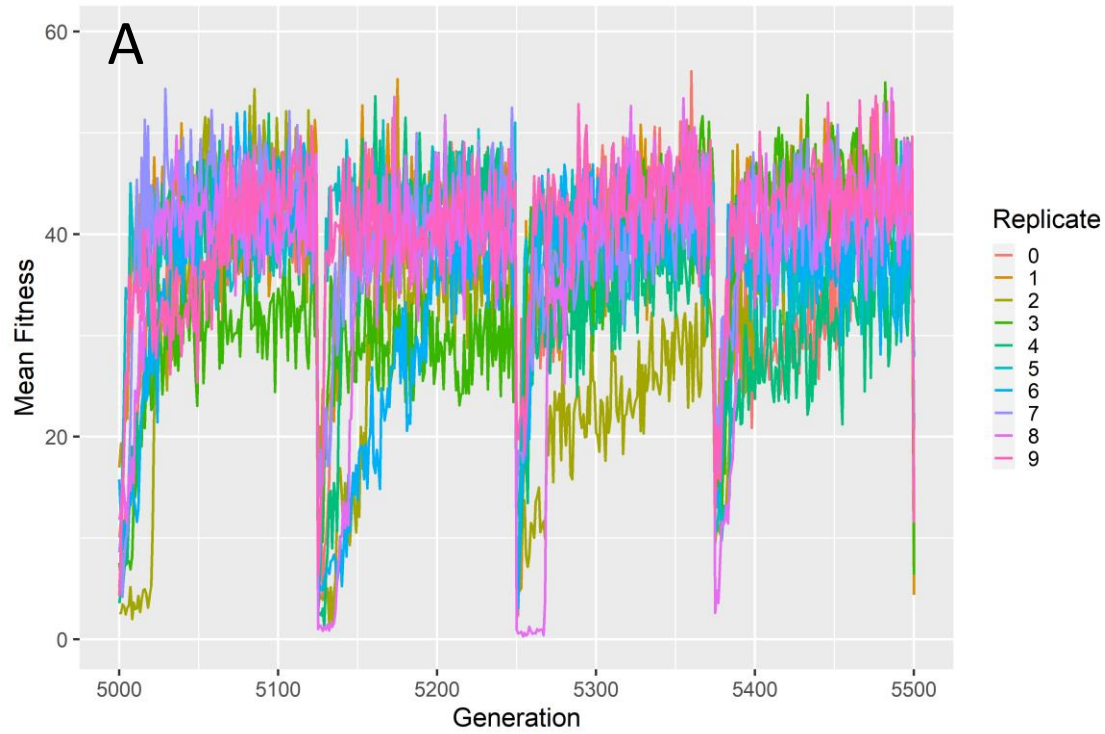


Figure S1. Fitness dynamics across simulation time in populations that resort to adaptive tracking. Fitness is strongly affected by environmental change. The environmental function shifts in this scenario by 0.4 in the environmental axis every 125 generations. All populations or replicates in this scenario react in a very similar way during the early stages of the simulation (A, 5,000-5,500 generations), but the way

they react to environmental change differs when reaching later stages of the simulation (B, 45,000-45,500 generations).

Performance of the evolved ANNs

The measure of performance in our model corresponds to the proportion of the amount of resources gained by the individual to the maximum amount it could get if it would always choose the best cue during the exploitation phase. The performance was shown to vary depending on the learning style that evolved (either fine-tuned learners, reckless learners) or other evolved strategies like adaptive trackers or random choice. The results of performance are summarized in the following plot for experiment 1.

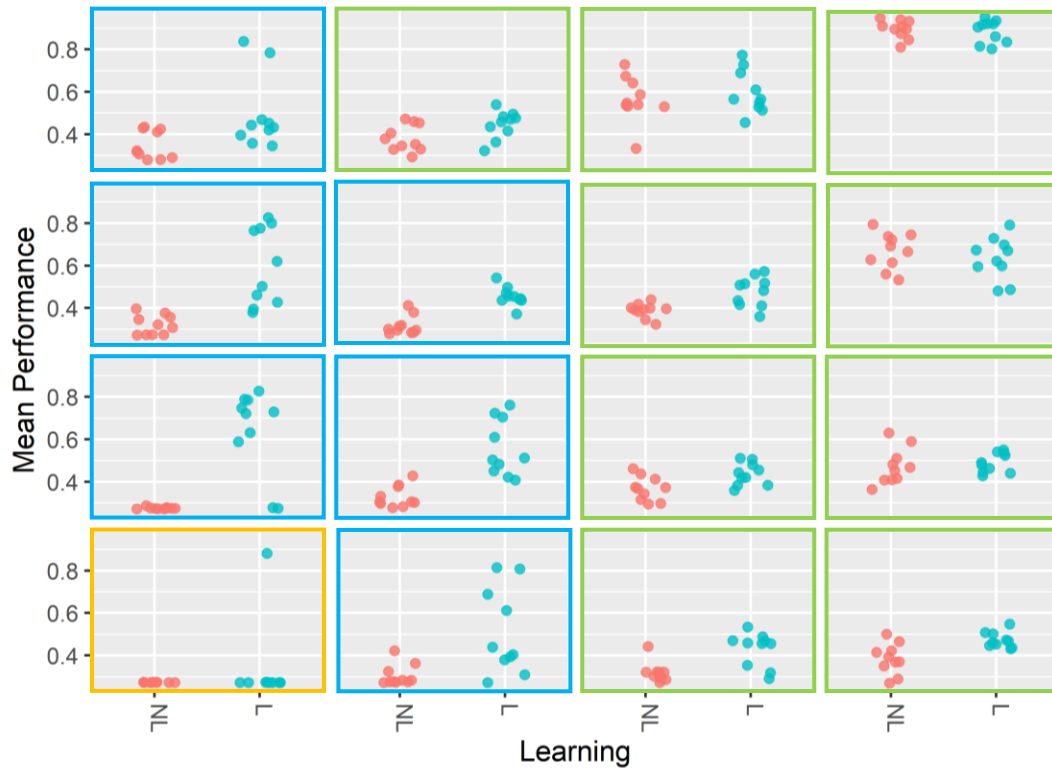


Figure S2. Performance of the evolved strategies in different environmental regimes. The rows represent the different magnitudes of the environmental shifts (0.1, 0.2, 0.3, and 0.4, from top to bottom) and columns represent the different frequencies of environmental change used in the simulations (every generation, 1/5 generations, 1/25 generations and 1/125 generations, from left to right). Mean performance at the end of the simulations shows that learning styles outperform non-learning controls in parameter combinations that favour learning. The higher levels of performance (ranging from 0.6-0.8) in environmental regimes that favour learning (highlighted in blue) are associated to fine-tuned learners, followed by reckless learners at around 0.4-0.5. The worst performance was found by populations that were employing random choice in the most unstable environmental regime, where learning was not favoured to evolve (highlighted in orange) and had a performance of around 0.26. Stable environmental regimes favoured adaptive tracking with a variety of performance levels, but reaching means close to 1 in the most stable regime (highlighted in green).

Evolved decision curves

The task that was presented to the ANNs during this simulation could be considered an optimization problem in which the networks needed to approximate the peak of its decision curve to the one from the real environmental function. I identified two different learning styles that create fundamentally different decision curves, shown below. The reckless learners are able to shift drastically their decision curve with one or two cues across the range (Figure S3), while fine-tune learners take several learning episodes to create a good approximation of the environmental function (Figure S4).

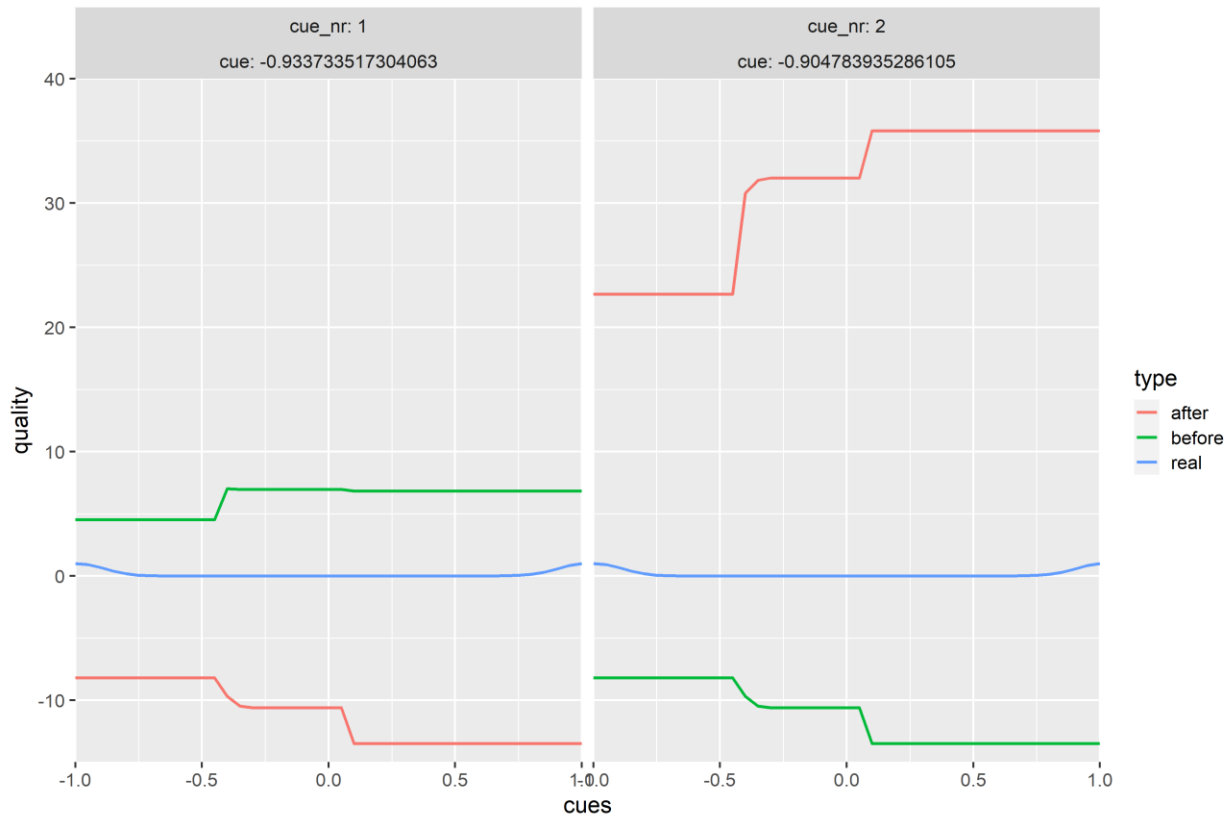


Figure S3. Effects of learning episodes on decision-making curves for a reckless learner neural network.

The inherited decision curve (green line) of a reckless learner can shift dramatically after receiving a single cue to a new approximation (red line) of the environmental function (blue line). The panels correspond to a series of cues that were presented to a RL network and the shift of its decision curve find the higher quality associated to the extremes of the environmental function, but leave a large area of the environmental range with an equal expected quality.

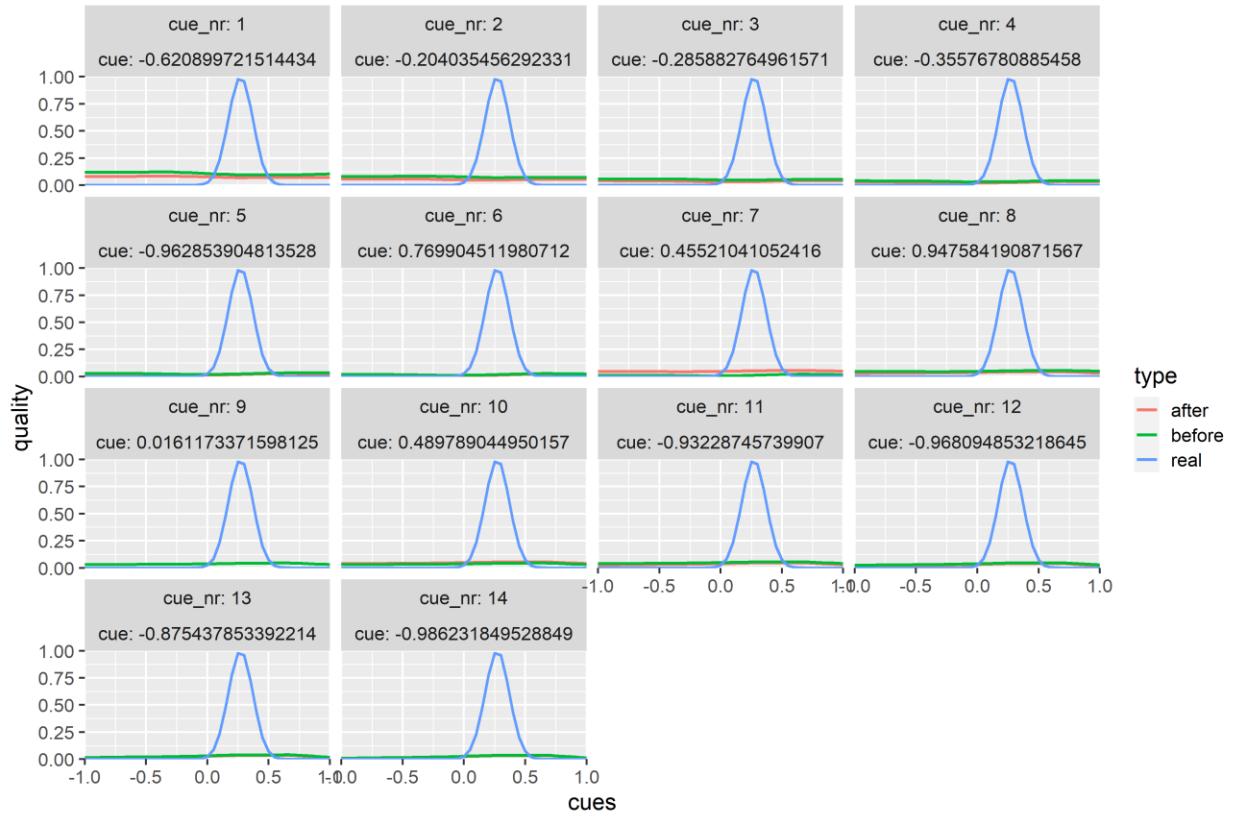


Figure S4. Effects of a sequence of learning episodes on decision-making curves for a fine-tuned learner neural network. A fine-tuned learner is able to modify its inherited decision curve (green line in the first panel) through many learning episodes that shift its approximation of expected quality (red line) to match the environmental function (blue line). Here the effect of cues can be varied, and a sequential representation of the shifts of the decision curves after each learning episode is best to show the functioning of the learning style. In the end, the decision curve is very slightly curved under the range of increased environmental quality, but small differences in relative quality allow these networks to discriminate cues and acquire performances of 0.6-0.8. It is important to note that the decision-making performed by the networks is only dependent of the relative quality associated to each cue given by their decision curve, and not depending on real quality.

VIII. Appendix 2: Experiment 2

Alternative permanent bottleneck scenarios

The permanent bottleneck experiment (Experiment 2) was run in three different scenarios. The base-line scenario is presented in the main text, but the two other present somewhat different characteristics. The first scenario shown below takes networks that evolved with weights constrained from -1 to 1. The best network from each replicate at the end of the simulation was selected to create monomorphic populations of different sizes, as presented in the main text. Their mutation rate and possible weights range was maintained, and the populations were able to evolve for another 50,000 generations. The results show a similar pattern to that observed in the scenario with relaxed constraint weights, but with a slightly higher tendency to lose fitness, especially visible for fine-tuned learners (Figure S5).

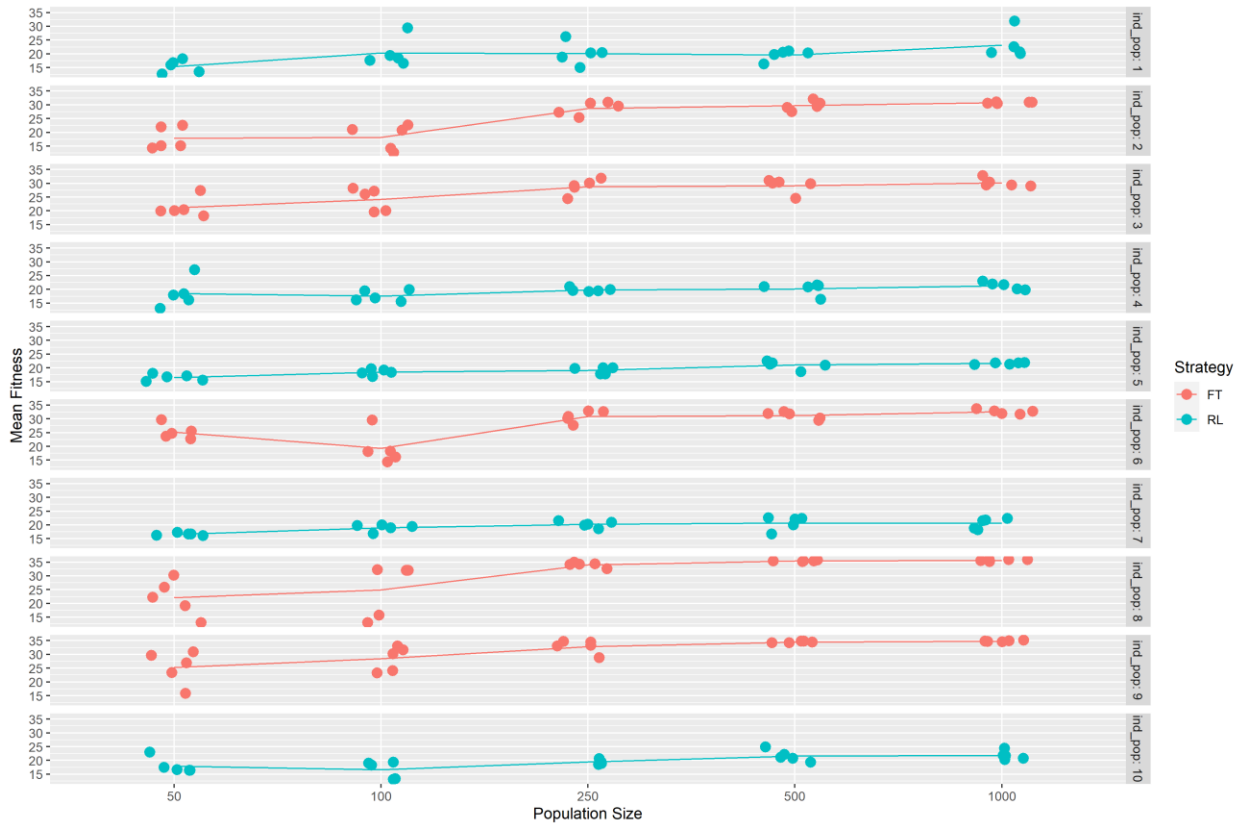


Figure S5. Mean fitness at the end of a bottleneck period for monomorphic populations of fine-tuned learners (red) and reckless learners (blue) in a scenario of evolution of the neural network under constrained weights. Populations of fine-tuned learners and reckless learners experience a decline in fitness when suffering prolonged and harsh bottlenecks. The decline of both strategies is especially evident when the population is reduced to 100 or less individuals.

The population bottlenecks of 100 and 50 individuals lead to decrease the fitness of the populations of both learning styles after 50,000 generations of evolution.

The next scenario presents an approximation to stress-induced mutations that could occur during a bottleneck event. The populations used in the size reduction phase in this scenario are the same ones

from the previous scenario, meaning they have evolved with the same constraint in their weights. The mutation rate was kept at 0.01, but the mutation step size was increased (scale parameter in Cauchy distribution was increased from 0.001 to 0.01), increasing the effect of each mutation. Also, the constraint in the weights was relaxed to -10 to 10. This scenario of increased influence of mutations caused the degradation of the learning mechanisms in all replicates after 50,000 generations under harsh reduction in population size, and even in some of the replicates evolving at sizes of 250 individuals (Figure S6).

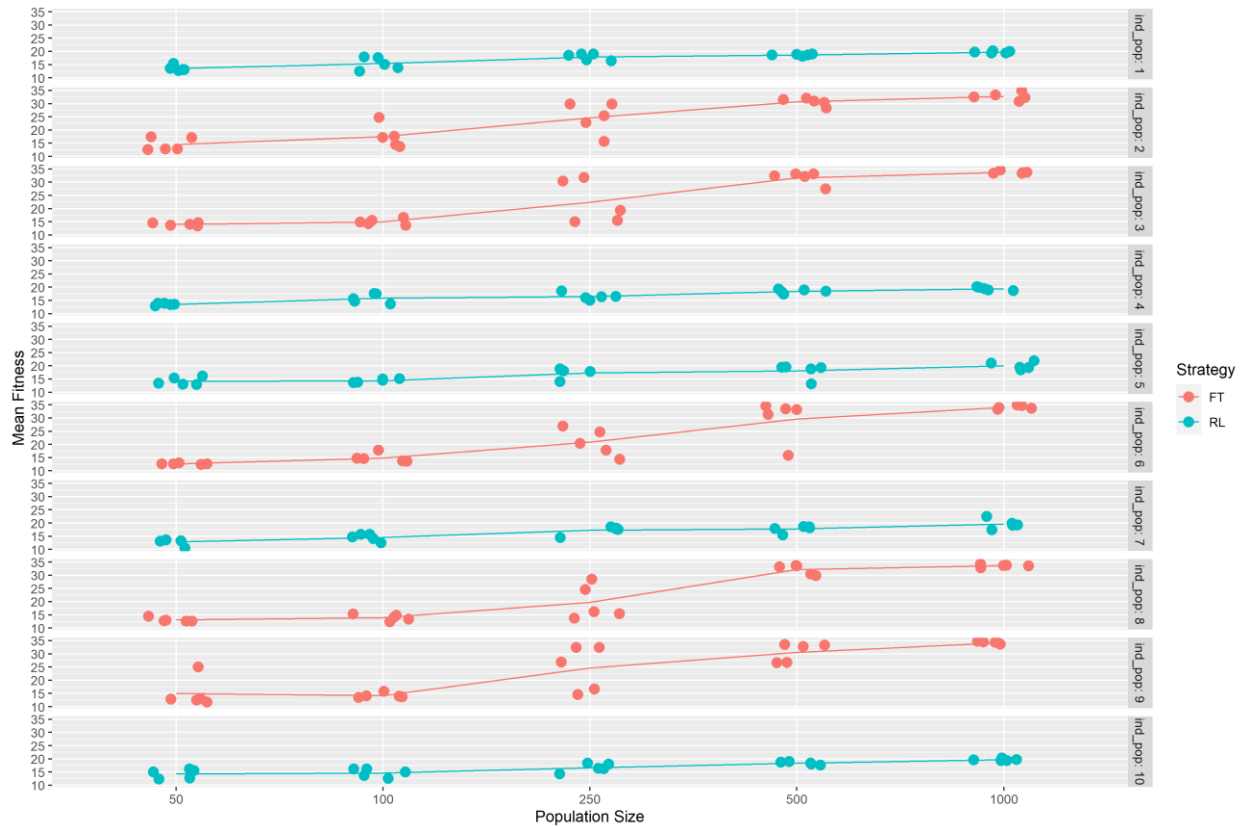


Figure S6. Mean fitness at the end of a bottleneck period for monomorphic populations of fine-tuned learners (red) and reckless learners (blue) in a scenario of evolution of the neural network under stress-induced mutagenesis. Permanent population bottlenecks under a stress-induced mutagenesis scenario caused the degradation of all learning mechanisms in population sizes of 100 and 50 individuals. Decline in fitness was considerable even at population sizes of 250.

Out of this last scenario, I picked an individual network from each replicate whose decision curve I show above (reckless learner number 7 in Figure S3, and fine-tuned learner number 9 in Figure S4) and I plotted their decision curves again at the end of the experiment. Their decision curves have been altered completely and are no longer able to track the environmental function.

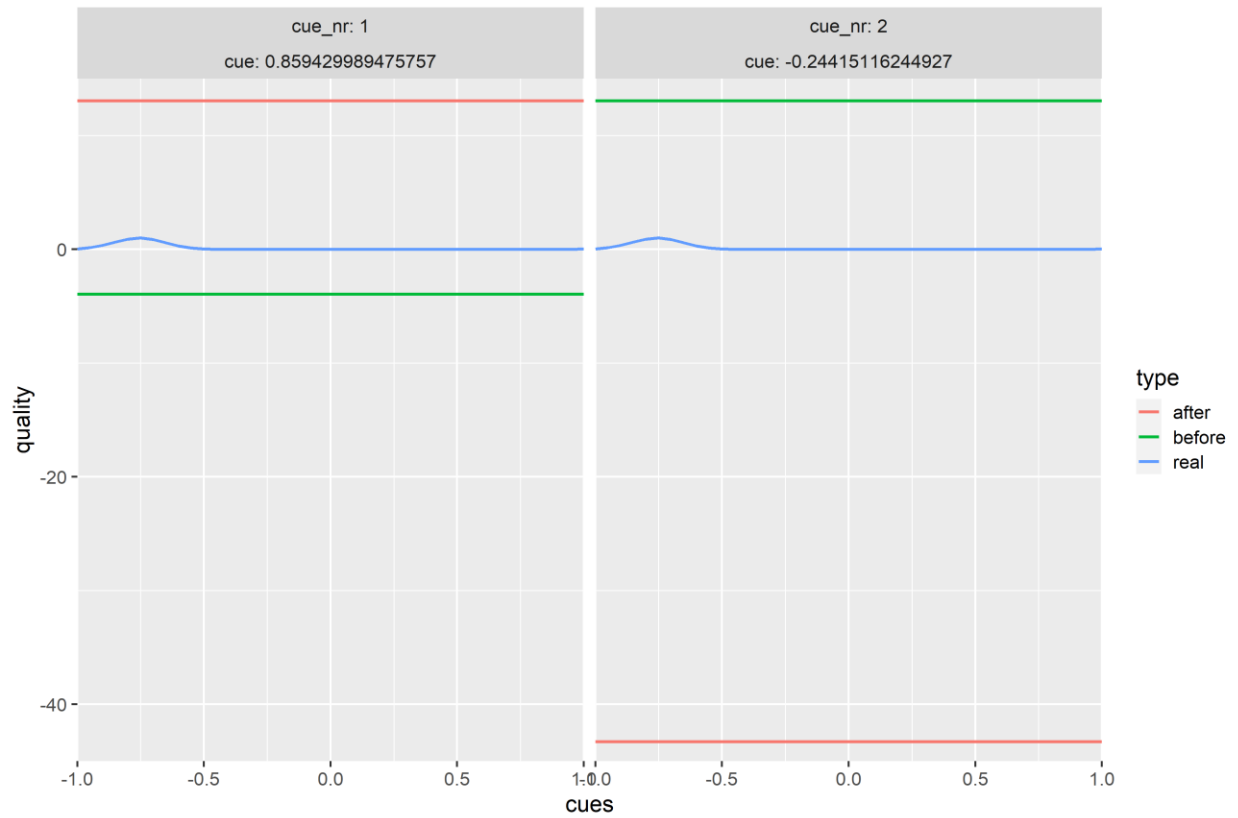


Figure S7. Response of the decision-making curve for a reckless learner neural network that has been experiencing a population bottleneck to 50 individuals under the baseline conditions scenario after 2 consecutive learning episodes. Degraded reckless learner decision curve is no longer able to track the environmental mean in an accurate way and takes large leaps in estimated quality every learning episode. Instead, this networks has fallen into random choice, as depicted by a flat decision curve.

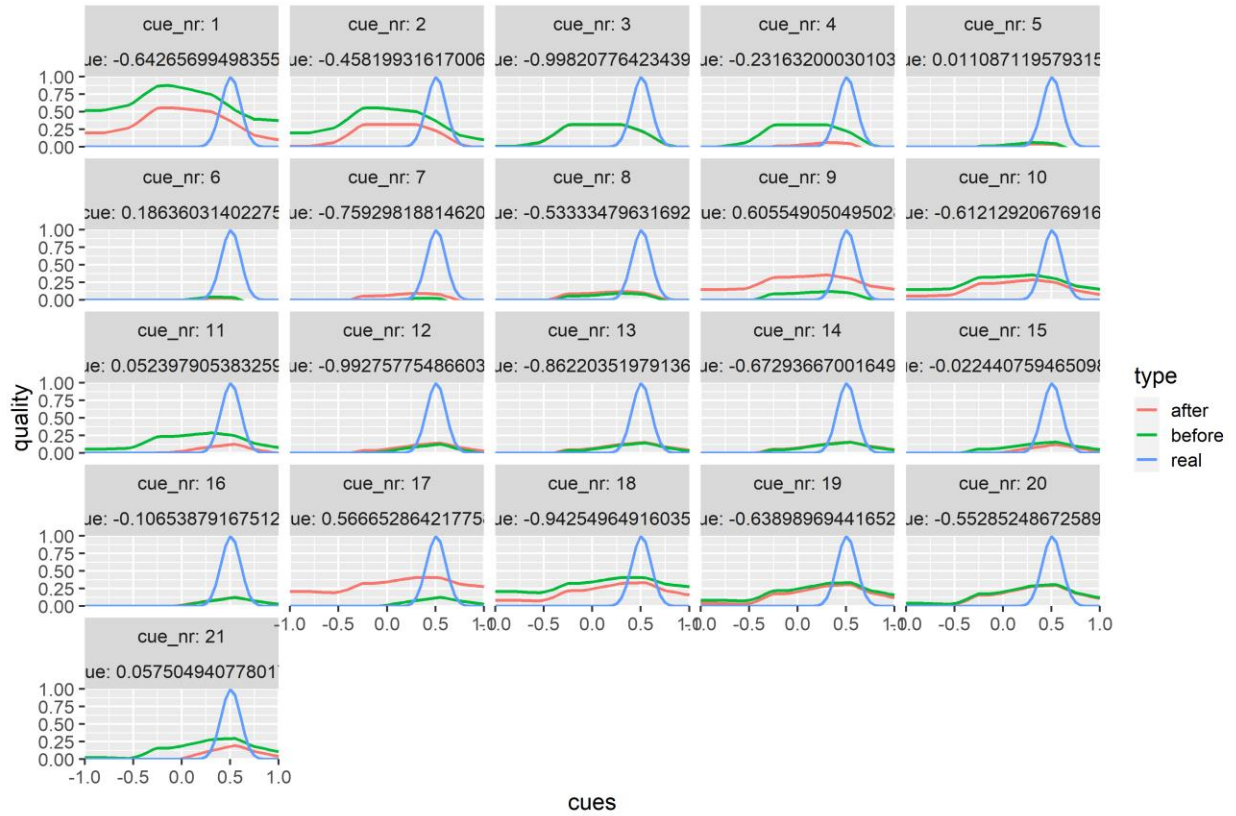


Figure S8. Response of the decision-making curve to a series of learning episodes for a fine-tuned learner neural network that has been experiencing a population bottleneck under the baseline conditions scenario. Despite having experienced a prolonged and harsh bottleneck, this individual still manages to modify its decision curve to have a fair approximation of the environmental quality peak, but experiences a slight decrease in precision. In some panels, the red line is not shown as the learning episode is unable to create a shift in the decision-making curve and both “before” and “after” curves are found superposed.

The decision curves produced by the ANNs following the population reduction experiment show a degradation in the learning mechanism that has led to slightly sub-optimal performance, but it seems to ultimately be maintained.

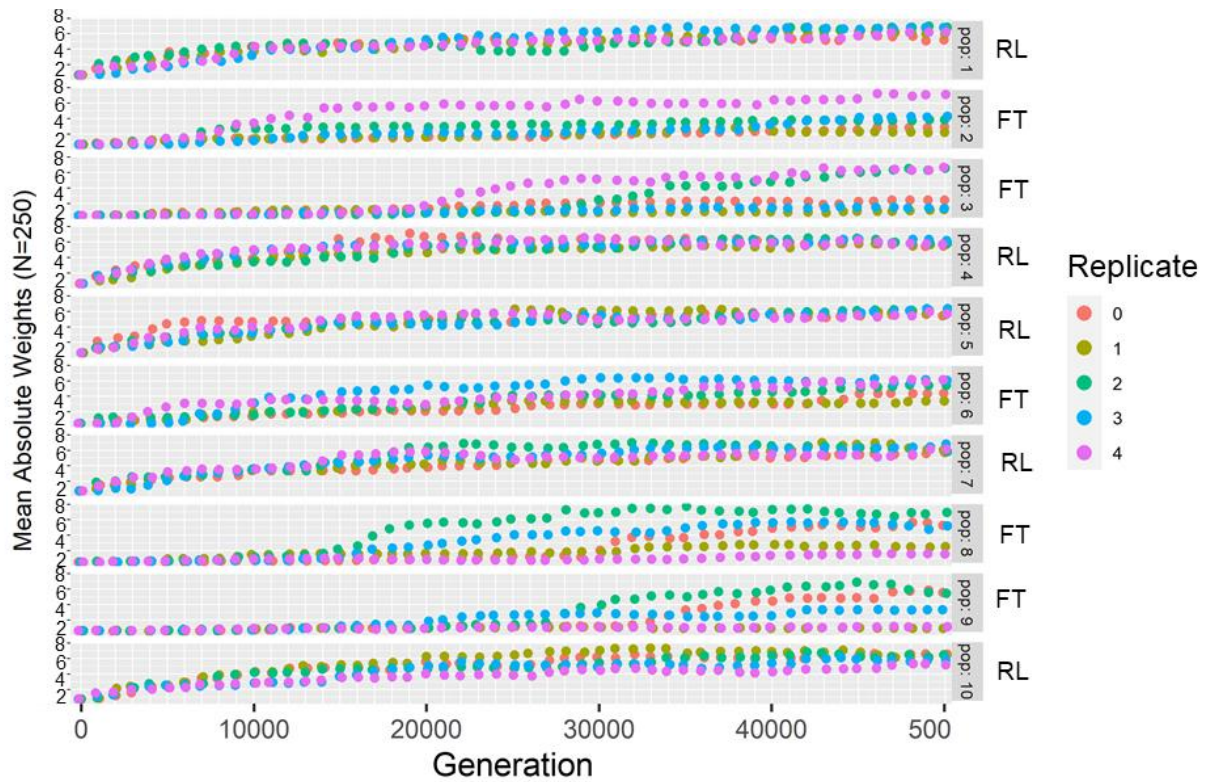


Figure S9. Learning styles experience mutational build-up at different rates in a stress-induced mutagenesis scenario during a bottleneck to 250 individuals. In this plot, the mean absolute weight values in populations of reckless learners (RL) and fine-tuned learners (FT) shows that RL accumulate mutations at a higher rate than FT. This difference is evident at a population size of 500 and 250 individuals but disappears when the bottleneck is harsher. Stress-induced mutagenesis causes RL to experience a mutational meltdown even at constant population size, while FT remains mutationally robust.

IX. Appendix 3: Experiment 3

Alternative cyclic bottleneck experiment

In addition to the cyclic bottleneck experiment that I present in the text above, I ran cyclic bottlenecks that reduced the population to 100 individuals, instead of 50. Every other parameter was kept the same. The only difference is that I did not run a constant bottleneck scenario, as my main interest was to explore the effect of cyclic bottlenecks.

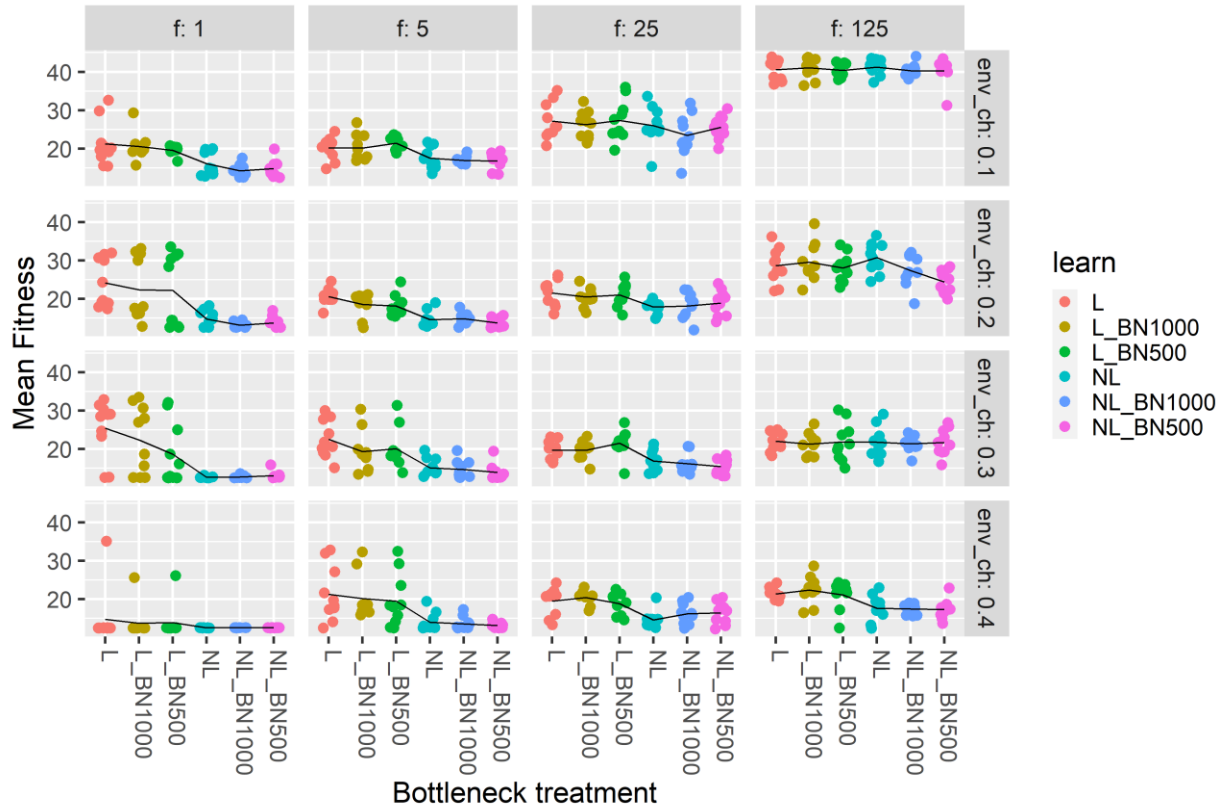


Figure S10. Mean fitness of learning and non-learning populations that evolved under constant population size or different bottleneck treatments. During the bottlenecks populations size was reduced to 100 individuals which seem to allow the maintenance of learning styles in the same way as harsher reductions to 50 individuals (see fig ... main text). The populations that evolve learning styles seem to maintain it, with a possible exception in the regime of environmental shifts of 0.3 every generation and the shorter bottleneck treatment every 500 generations.

The observed pattern for this scenario does not differ drastically to that observed with harsher cyclic bottlenecks. The only exception is for the regime of environmental shifts of 0.3 every generation and the shorter bottleneck treatment every 500 generations. In this case, several replicates were not able to evolve a learning style and maintain a lower fitness towards the end state of the simulation. It is unlikely that this is due to the degradation of a pre-existing learning mechanism, but more likely due to the difficulty to evolve a learning mechanism under cyclic bottlenecks and environmental instability.

X. Appendix 4: Technical clarifications

Example of a ReLU function

The ReLU function normalizes the activation strength of a node by limiting the values from 0 to 1 in the following way.

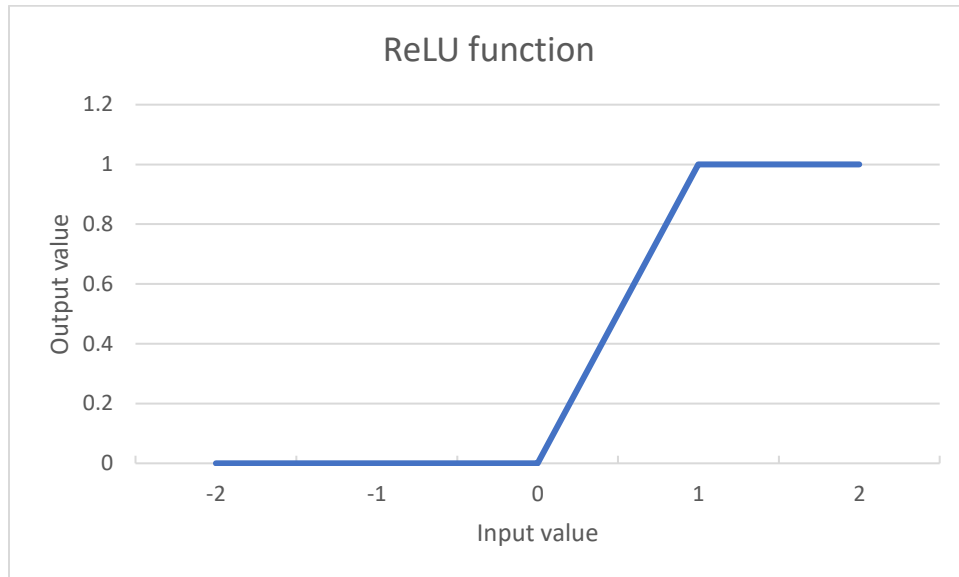


Figure S11. Visual representation of the ReLU function. In this case, the input value is the node output and the ReLU function limits the activation signal value that is transferred to the next node in the neural network.

Up next, I have prepared a table with an overview of the parameters that I used during my simulations.

Table S1. Model base parameters and examples of values they can take during the simulation.

Parameter	Example value
Initial weight range	Range within which weights were randomly assigned at the beginning of a simulation
Initial learning rate range	Range within which learning rate was randomly assigned at the beginning of a simulation
Initial learning episode mean	Mean of the randomly assigned learning episodes at the beginning of the simulation
Weight mutation rate	Mutation probability per weight in the network
Weight mutation step size	Location of the Cauchy distribution that determined the sizes of the mutation steps
Learning rate mutation rate	Mutation probability for the learning rate
Learning rate mutation step size	Mean of the normal distribution from which mutation sizes were drawn for learning rate
Learning episode mutation rate	Mutation probability for the number of learning episodes
Learning episode mutation step size	Mean of the Poisson distribution from which the size of learning episode mutations was drawn from
Population size	Number of individuals in the population
Generations	Number of generations that the simulation would run
Lifespan	Timesteps that compose the lifespan of a generation
Number of replicates	Number of replicate populations that will be simulated from a set of initial conditions
Environmental sample size	Number of environmental cues that will be presented to the individuals when exploiting the environment
Environmental range	Range of values that are spanned by the wrapped gaussian function that represent environmental quality
Initial mean environment	Initial mean of the environmental function. Shifts with environmental change parameters
Standard deviation of the environmental function	Standard deviation of the environmental function throughout the simulation
Size of environmental change	Magnitude of the shifts in the mean of the environmental function
Standard deviation of environmental change	Standard deviation applied to the environmental shift
Environmental change rate	Frequency of environmental shifts. Is interpreted as a shift every x generations during the simulation
Frequency of bottlenecks	Frequency of population bottleneck. Is interpreted as a shift every x generations during the simulation
Proportion of the population during bottlenecks	Proportion of the population that survives and reproduces during a bottleneck event.
Duration of bottlenecks	Generations that the population spends under a reduced size due to a population bottleneck event
Stable period	Number of generations before bottleneck events
Minimum weight value	Lower limit of possible weight values
Maximum weight value	Upper limit of possible weight values

XI. Acknowledgements

I would like to thank Magdalena Kozielska and Franjo Weissing for their close supervision, the time and energy that was invested in my project, and in the priceless discussions and feedback to my work. It was deeply cherished. I would like to thank my all the friends I made during my MEME journey, especially to Lianne, Juan José, Anne, Eleanor, Clem, and Juliana, for their company, their support, and the many shared meals at the 5th floor table. I would like to thank the MARM group for making me feel extremely welcome and included into a research community that challenged me and nurtured me. I would like to thank MEME for the scholarship that made all this possible. Finally, I would like to thank Ana for her love, her enthusiasm, and her support which helped me finish this project, which was as fascinating as it was confusing at times for the both of us.