

# Analysis of Social Learning Strategies when Discovering and Maintaining Behaviours Inaccessible to Incremental Genetic Evolution

Ben P. Jolley, James M. Borg, and Alastair Channon

School of Computing and Mathematics, Keele University, UK  
{b.p.jolley,j.borg,a.d.channon}@keele.ac.uk

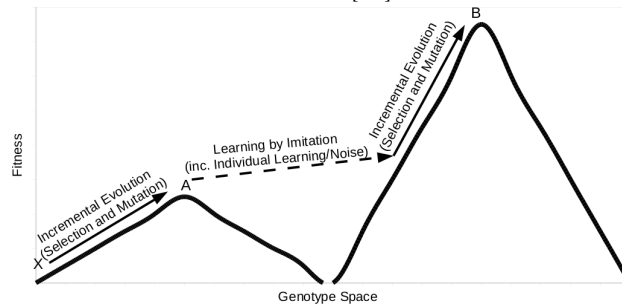
**Abstract.** It has been demonstrated that social learning can enable agents to discover and maintain behaviours that are inaccessible to incremental genetic evolution alone. However, previous models investigating the ability of social learning to provide access to these inaccessible behaviours are often limited. Here we investigate teacher-learner social learning strategies. It is often the case that teachers in teacher-learner social learning models are restricted to one type of agent, be it a parent or some fit individual; here we broaden this exploration to include a variety of teachers to investigate whether these social learning strategies are also able to demonstrate access to, and maintenance of, behaviours inaccessible to incremental genetic evolution. In this work new agents learn from either a parent, the fittest individual, the oldest individual, a random individual or another young agent. Agents are tasked with solving a river crossing task, with new agents learning from a teacher in mock evaluations. The behaviour necessary to successfully complete the most difficult version of the task has been shown to be inaccessible to incremental genetic evolution alone, but achievable using a combination of social learning and noise in the Genotype-Phenotype map. Here we show that this result is robust in all of the teacher-learner social learning strategies explored here.

## 1 Introduction

Previous research has shown that with the use of social learning, individuals are able to discover more complex behaviours that are not accessible via incremental genetic evolution alone [3]. In this work, and many other simulation models that explore social learning and culture, social learning itself is often limited. These limitations are often centred around who individuals learn from. Here we expand on previous work to explore whether behaviours inaccessible to incremental genetic evolution alone are still discovered, and maintained, when individuals are permitted to learn from a variety of different individuals. We go on to discuss why these differing teacher-learner social learning strategies solve the task used here in differing ways.

## 1.1 Incremental Genetic Evolution

Incremental genetic evolution necessarily uses converged populations, which is often referred to as the Species Adaptation Genetic Algorithm (SAGA) approach [6]. SAGA impacts on the way populations evolve: recombination will have a far smaller effect on the motion of the population than in a standard genetic algorithm, as each species is already genetically similar, leaving mutation as the primary driving force behind evolution. Mutation can be substantially effective in spaces percolated by neutral networks: pathways of level fitness through the fitness landscape; in this case genotypes can vary while still producing similar phenotypes and behaviours. When phenotypes of higher fitness are found the population converges onto them thus enabling species to discover and converge upon easily accessible solutions and behaviours. However, if there is no neutral or incremental path between the corresponding basic behaviour and fitter behaviours, the population will struggle to move away from sub-optimality. Fig 1 depicts a mock example. One approach to solving the problem of suboptimal convergence is to increase the rate at which mutation is applied, potentially allowing the population to explore more of the fitness landscape. However, there are problems with this approach: as mutation rates increase, evolutionary search begins to resemble random search making it increasingly difficult for the population to maintain solutions. The point at which mutation becomes so large that favourable structures discovered by evolution are lost more frequently than they are found is known as the error threshold [10].



**Fig. 1.** A species starting from point X on the above mock fitness landscape would achieve peak A via the hill climbing strategy adopted by incremental genetic evolution (driven primarily by mutation and selection). The inclusion of noise in the genotype to phenotype map and social learning (e.g. imitation) can enable the species to bypass areas of lower fitness.

## 1.2 Discovering and Maintaining Inaccessible Solutions

To solve the issue of sub-optimal population convergence without crossing the error threshold, noise can be added to the fitness landscape via the genotype to fitness map. However, depending on where such noise is in the phenotype to fitness section of that map, its ability to aid in the transition between peaks is

limited. By instead incorporating noise into the genotype to phenotype map, behaviours inaccessible to incremental genetic evolution may be exhibited reliably by individuals while leaving the genotype untouched. One method for introducing noise in this way is to introduce transcription errors when writing from the genotype to the phenotype in systems with equivalent genotype and phenotype encodings, such as direct artificial neural network weight encodings [3]. By introducing potentially new behaviours to the phenotype we deny the initial possibility of these behaviours being inherited by new individuals through standard Darwinian evolutionary mechanisms. Therefore in order to maintain successful behaviours in the population, some form of extra-genetic learning needs to take place. The extra-genetic learning employed in this model is a combination of the aforementioned genotype to phenotype noise and social learning through interaction between teachers and learners to facilitate the transmission of learnt behaviours [1,4]. As in Borg et al. [3], learners or pupils follow teachers in a mock evaluation on a set of environments or maps. As both teacher and pupil receive the same environmental input the teacher’s output may be used as a target pattern for error back-propagation, reducing the pupil’s output error compared to that of the teacher. By learning in this way pupils are able to partially imitate the behaviours exhibited by teachers, thus maintaining aspects of new behaviours in the population that would have been lost by a stand alone evolutionary process.

The use of teacher-learner social learning has been shown to be sufficient for discovering and maintaining behaviours inaccessible to incremental genetic evolution alone in a grounded simulation [3]. However, these simulations only allowed one form of social learning, in which offspring would learn from their fittest parent. Though a valid approach that has been used in previous work [4], there are other theoretical and empirical models that can be adapted to this work to evaluate whether or not other social learning strategies are still capable of achieving these complex behaviours.

Social learning is seen widely in nature [11] and in a range of species as diverse as humans and nine-spined stickleback fish [7]. The mechanisms and processes that underpin social learning are themselves broad, ranging from teaching, imitation and emulation to stimulus enhancement and exposure [5], with any of these mechanisms potentially leading to formation of traditions and cultures [15,16]. However, within each social learning category there is some dependence on who information is obtained from, be it a teacher or which agent is unintentionally (or intentionally) exposing an individual to something new. As social learning is necessarily conformist, a poor learning model may result in the discovery and propagation of sub-optimal behaviours. In this work we assess whether who you are learning from, otherwise known as ‘who’ social learning strategies [8], can hinder social learning’s ability to discover and maintain behaviours inaccessible to incremental genetic evolution alone, thus undermining social learning’s adaptive advantage over incremental genetic evolution in complex environments.

Laland [8] assess both ‘who’ and ‘when’ social learning strategies, alongside the complexity of social learning in animals, providing evidence to show its adaptive advantages. Laland [8] has a particular focus on conformity: a population’s

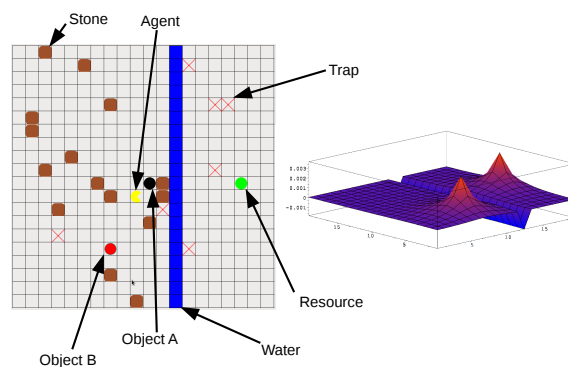
ability to share popular behaviours amongst each other while minimising exploration for new behaviours; the use of conformist social learning can be beneficial or detrimental depending on the environment or task [8,2,9]. It has also been suggested that conformist social learning that is not supplemented with non-social exploration can lead to population collapse in temporally varying environments [2], though recent work suggest that conformist learning may be of benefit in spatially varying environments [9]. The ‘who’ social learning strategies (concerned with who an agent should learn from rather than when learning should take place) inspired by Laland [8] are modelled here as three core social learning strategies: ‘Best Parent’, ‘Oldest’ and ‘Fittest’. The ‘Fittest’ strategy selects the fittest individual from the population to be the teacher. The theoretical basis behind this strategy falls partially into the ‘Learning from majority’ category discussed by Laland [8], but also has a wider basis in nature with many animals being shown to learn from more successful individuals. Learning from older individuals derives from the rationale that older individuals must have exhibited successful behaviours to survive, however this does not have to mean the older individual in question is in fact the fittest individual, due to this the ‘Oldest’ strategy is likely to provide a broader range of behaviours than the ‘Fittest’ strategy. The ‘Best Parent’ strategy (as seen in Borg et al. [3]) sets the teacher to be the parent who has won the right to reproduce in a tournament. This is the least conformist strategy of the three as it allows unfit individuals, relative to the rest of the population, to be parents as tournaments only involve a small number of individuals. Additional to these three core strategies we also introduce social learning strategies for learning from random and young individuals. Though not widely evident in nature, the theoretical benefits of learning from a random individual (sometimes described as unbiased social learning) have been discussed in numerous works [12,9]. The theorised benefits of unbiased social learning arise in temporally varying environments, where learning from a broader set of individuals enables increased access to new behaviours that may be relevant in the specific environmental state being experienced. A ‘Youngest’ strategy, despite no theoretical basis, is being evaluated as a contrast to the ‘Oldest’ strategy.

### 1.3 Neuroevolution of Deliberative Behaviours for an Advanced River Crossing Task

This work uses populations of hybrid neural networks embodied in agents (often referred to as animats). The hybrid networks are comprised of two different neural networks: the first controlling the high level deliberative behaviours of the animat, and the second controlling the animat’s reactive capabilities. Hybrid neural network architectures of this sort have demonstrated the ability to seek long term goals whilst also reacting to unforeseen events ultimately enabling the evolution of complex problem solving abilities [13,3,14]. To demonstrate these problem solving abilities Robinson et al. [13] developed a complex problem called the ‘river crossing’ or RC task. The RC task required animats

to find a reward-giving Resource in a 2D grid-world environment containing a number of obstacles, including Traps, Water (connected to form an impassable river), and Stones. In order to cross the river animats were required to pick up Stone objects, which could be carried at no cost to the animat, and place them in the same cells as Water thus negating the cell's lethality. Once a continuous bridge of Stones over the river had been built, animats could access the Resource. Despite the RC task being reasonably complex, it has been demonstrated that it could be solved by initially converged populations of animats using only incremental genetic evolution [13]. To test whether social learning could discover and maintain behaviours inaccessible to incremental genetic evolution alone, a more complex version of the RC task, known as the RC+ task, was developed by Borg et al. [3]. A snapshot of the RC+ task can be seen in Fig 2.

An important aspect of the RC task was that individuals were evaluated on increasingly difficult environments. The RC+ task maintains this principle whilst making the RC task more difficult in regard to both river width and exposure to Stone objects. The number of environments an animat was evaluated on increased from three to five, with environments becoming increasingly difficult to solve due to river width increasing from zero to four cells. To add additional difficulty, the number of Stone objects gradually decreases from twenty in the first environment to zero in the final environment, rendering the bridge building behaviour useless for solving the final environment. In order to make the final environment solvable, two extra objects, Object A and Object B, were introduced into the environment. Object A and Object B were rare objects, with only one instance of each found in each environment. Object A and Object B are moveable at no cost to the animat and may be placed upon any cell or object. If an animat places both Object A and Object B on a cell containing Water, a reward equal to that of the Resource is received and the animat is considered to have successfully solved the map. The RC+ task has been shown to be impossible to solve with incremental genetic evolution alone [3].



**Fig. 2.** RC+ environment with accompanying activity (shunting) landscapes.

## 2 The Model

The hybrid neural network used by Borg et al. [3] and others [13,14], and therefore used here, may be broken down into two network models: a shunting network and a decision network, with the decision network passing information on to the shunting network which in turn controls animat movement. The shunting network is not directly exposed to any evolution or learning. The deliberative network is exposed to both evolution and learning, enabling the evolution and inheritance of behaviour.

### 2.1 The Shunting Network

The shunting network is a locally-connected, topologically-organised network of neurons that was originally used for collision free motion planning and has already been applied to the river crossing task [13,14] and RC+ task [3]. This network is advantageous as it exhibits computational efficiency by not explicitly searching over all possible paths. The shunting model is used here by mapping the topologically-organised neurons as cells in the RC+ environment’s 20 by 20 grid. Using the shunting equation (see equation 1) values are to propagate across the neurons/cells using the outputs from the decision network, producing an activity landscape with peaks and valleys representing desirable and undesirable features in the environment. The result is a landscape which allows the animat to follow a route with the higher (iota) values while avoiding undesirable valleys. An example of an activity landscape with a snapshot of the environment it represents can be seen in Fig 2.

$$\frac{dx_i}{dt} = -Ax_i + \sum_{j \in N_i} w_{ij} [x_j]^+ + I_i \quad (1)$$

In equation 1 each neuron/node in the shunting network corresponds to one cell in the RC+ environment;  $x_i$  is the activation of neuron  $i$ ;  $A$  is the passive decay rate;  $N_i$  in the receptive field of  $i$ ;  $w_{ij}$  is the connection strength from neuron  $j$  to  $i$ , specified to be set by a monotonically decreasing function of the Euclidean distance between cells  $i$  and  $j$ ; the function  $[x]^+$  is  $\max(0, x)$ ; and  $I_i$  is the external input to neuron  $i$  (known as the Iota output).

### 2.2 The Decision Network, Evolution and Learning

The decision network can inform agents of desirable and undesirable objects in the environment based on the agent’s current environmental position. The decision network is a feed-forward neural network, with a single four neuron hidden layer, that inputs information from the animats current location in the RC+ task world to gain an iota value for each possible environmental state, with the exclusion of grass whose value is always set to 0 (to ensure neutral space

for activity to propagate through when producing an activity landscape using the shunting model). A hyperbolic tangent activation function is applied at each output node with a boundary of -0.3 to 0.3; values below -0.3 round to -1, above 0.3 rounds to 1 and in between rounds to 0. The output layer contains sixty-seven neurons representing the Iota values of all sixty-four possible environmental states, including inaccessible and redundant combinations, with an additional output neuron for each pick-up/put-down operation on each non-static objects (Stones, Object A, Object B). A standard hyperbolic tangent activation function is applied at each hidden node.

Here error back-propagation is used to simulate learning. The use of error back-propagation to simulate learning has been previously used by Curran et al. [4] to enable pupil outputs to be corrected to more closely resemble the teacher outputs. In Curran et al. [4] multiple learning cycles are conducted, until the error between learner and teacher outputs is minimised to an acceptable level. Here, as in Borg et al. [3], a similar approach is taken, with learning cycles represented by each move in a mock evaluation of the environment by the teacher. However, unlike Curran et al. [4], learning only continues until either the demonstrator completes all five maps or fails, no direct attempt is made to ensure learner outputs were minimised to some arbitrary level. A novel approach is taken in this experiment, that builds on previous work by having different simulations with different teacher-learner social learning strategies.

### 3 Experimentation

The model used here is fundamentally the same as introduced by Borg et al. [3]. Each iteration/generation has a tournament event in which two individuals from the population of 100 are ran through the RC+ task, with each individual’s fitness being determined by the number of maps successfully completed. Each map gets increasingly more complex therefore if an individual is not able to complete a map they are prevented from continuing on to further maps. Each map has seven Trap objects and  $20 - (5 \times \text{riverwidth})$  Stone objects, both of which are randomly placed, though never on the same space, one reward-giving Resource on the opposite side of the map to the agent starting position, and one instance each of Object A and Object B. The river width varies from an initial width of zero, increasing by one cell per map. Each individual is evaluated on their ability to reach the resource or place Object A and Object B on to a cell containing Water. Agents fail when they come into contact with an uncovered Water or Trap element. Failing to complete a map within 100 steps is also evaluated as a failed attempt. The two tournament individuals are compared, with the fitter agent reproducing with a randomly selected agent from the population, with the child replacing the weaker of the tournament agents. Each loci in an agent’s genotype directly writes to a locus in the agent’s phenotype, which itself directly encodes a weight in the decision network, with all genotypes and phenotypes being of length  $L = 308$ . To ensure network structures from parents are maintained during reproduction, a single point recombination mechanism is

applied. Mutation follows recombination; each loci has a probability  $P_{mut} = 1/L$  of having a random value from  $N(0, 0.4)$  added to it, with the resulting values being bounded within the range  $[-1, 1]$ . Once the child genotype has been constructed it is written to the child agent’s phenotype; this process is referred to as transcription. During transcription two randomly selected connection weights are overwritten with a new random value selected from a discrete uniform distribution  $U(-1, 1)$ . Directly following reproduction the learning strategy is enforced via back-propagation. A mock evaluation of the RC+ task takes place between the teacher and child (now thought of as the learner), with the learner’s inputs being set to those of the teacher. Learning takes place until the teacher either fails or completes all five maps. At each step through the evaluation the learner attempts, via error-back propagation with a learning rate of  $\delta = 1$ , to imitate the teacher’s output for the current inputs.

The model in this work utilises five learning strategies, each with a different way of determining teacher selection. The winner of the reproduction tournament being set as the teacher in the ‘Best Parent’ strategy, the fittest individual in the population for the ‘Fittest’ strategy, the individual who has registered the most tournament wins for the ‘Oldest’ strategy, the last animat to be created before the current reproduction event in the ‘Youngest’ strategy, and a random individual for the ‘Random’ strategy. In any case where more than one individual met the criteria to be assigned the role of teacher, an individual from the valid sub-set was chosen at random, this situation only every arose when using the ‘Fittest’ or ‘Oldest’ strategies. One hundred populations for each learning strategy were evaluated so the results can be aggregated for an overview of each strategy’s performance. Simulations were run for 2,000,000 tournaments, with each simulation recording the fitness of the fittest individual and the mean fitness of the population at every 500th tournament. The highest fitness is five, which indicates an agent completed map five. To indicate the behaviour has not only been achieved but also maintained the fitness of five has to have been recorded a further ten times, without a suboptimal result. Each learning strategy is comprised of 100 populations of agents.

## 4 Results

Table 1 (top) shows the proportion of populations that were successful in solving each map. The most notable result was that all strategies were able to complete map five, the map which required exhibiting and maintaining a behaviour that in previous work was not obtainable by incremental genetic evolution alone [3], thus demonstrating that discovering and maintaining behaviours inaccessible to genetic evolution alone is possible using various teacher-learner social learning strategies, even those strategies that are either non-conformist (the ‘Random’ strategy) or contrary to strategies observed in nature (the ‘Youngest’ strategy). It should be noted that to complete one map, all preceding maps must have also been completed, therefore the ability to solve map five indicates that a population also managed to successfully complete maps 1-4. In Table 1 (top) we



Map	BP	Fittest	Oldest	Random	Youngest
<b>1</b>	99%	99%	99%	99%	99%
<b>2</b>	71%	68%	54%	74%	63%
<b>3</b>	47%	47%	37%	54%	47%
<b>4</b>	39%	46%	34%	49%	38%
<b>5</b>	8%	15%	5%	10%	7%

Map	BP	Fittest	Oldest	Random	Youngest
<b>None</b>	1%	1%	1%	1%	1%
<b>1</b>	28%	31%	45%	25%	36%
<b>2</b>	24%	21%	17%	20%	16%
<b>3</b>	8%	1%	3%	5%	9%
<b>4</b>	31%	31%	29%	39%	31%
<b>5</b>	8%	15%	5%	10%	7%

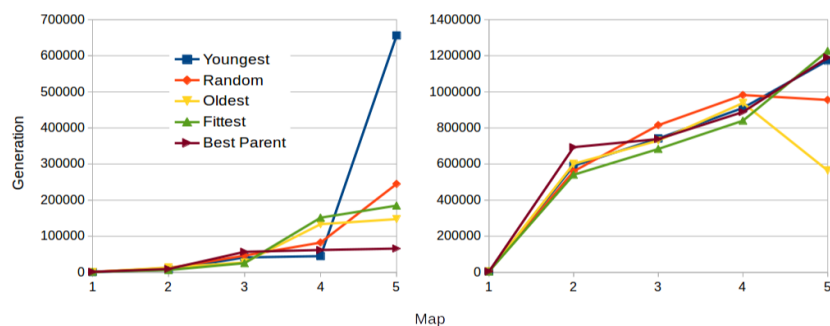
**Table 1.** (Top) % of populations completing each map for each social learning strategy. (Bottom) % of populations achieving each map as their maximum achievement for each social learning strategy. (BP = Best Parent)

do see many instances of learning strategies failing to complete simpler maps; we also see this in Table 1 (bottom), which shows how many populations were successful at completing each map as their maximum achievement, that is to say completed map one or two, ... without going on to complete any later maps. Maps 2-4 were all solvable using either a ‘bridge building’ strategy or the more advanced Object A + Object B strategy, the suggestion here is that some learning strategies sometimes failed to find the sub-optimal, but more incrementally accessible, ‘bridge building’ strategy. We would also expect to see populations that were able to complete map two also completing map four as the behaviour required is the same, the only difference being a wider river, however Table 1 (bottom) suggests that all strategies had populations that exhibited flawed behaviours which were not as generally applicable as they should have been. In comparable tests by Borg et al. [3], non-learning populations were shown to achieve above 90% success on maps three and four, with 100% success for maps one and two, the failure of the social learning strategies explored here to achieve this rate of success for maps three and four (as indicated by Table 1 (top)) indicates that whilst social learning can enable access to, and maintenance of, behaviours inaccessible to incremental genetic evolution, they are less effective at solving simpler, incrementally accessible, tasks. One explanation for this result is that social learning is necessarily conformist, even when unbiased or random, thus running the risk of sub-optimal behaviours being maintained and dispersed within the population.

The results also offer no definitive best strategy for the solving the RC+ task, as all are able to achieve the final map. However both Table 1 (bottom) and Fig 3 do allow us to begin seeing the differences between strategies. Performance may be viewed from three differing perspectives: (1) the number of populations achieving map five, (2) the distribution of maps achieved by populations, (3) the speed at which populations were capable of completing maps. Both measure

(1) and (2) may be considered using the data from Table 1 (bottom): from this data we can see that ‘Fittest’ strategy achieves the highest proportion of populations completing map five, however if we conduct a Chi squared test to find whether the proportion of populations achieving map five is dependent on the social learning strategy applied or not we come out with a p-value of 0.1316, thus indicating that the proportion of populations achieving map five is in fact independent of the strategy applied, therefore we cannot say with any certainty that the ability of the ‘Fittest’ strategy to achieve map five is significantly better than any other strategy (we do find that a Chi squared test that only considers the ‘Fittest’ and ‘Oldest’ strategies does provide a p-value below a significance level of 0.05, but no other pairings do). If we take Table 1 (bottom) to be a contingency table on which a Chi Squared test may be conducted we may be able to derive whether the distribution of maps achieved by populations (measure (2)) is dependent or independent of the social learning strategy used. When such a test is conducted a p-value of 0.04739 is produced, suggesting that the distribution of maps achieved by populations is dependent on the strategy used. This result requires further investigation of the data for each population, for each strategy, in order to provide a robust overview of the dynamics each strategy employs to solve the task. Measure (3) may be considered using the graphs seen in Fig 3. From Fig 3 it seems that populations employing the ‘Best Parent’ strategy are able to achieve map five quicker than other strategies, with the ‘Youngest’ strategy seeming to struggle to achieve map five in any haste. However if we only consider the average number of generations to complete each map both ‘Best Parent’ and ‘Youngest’ seem to give an average performance, with ‘Random’ and ‘Oldest’ giving the best general performance. It is interesting to note that those populations employing the ‘Oldest’ strategy who are able to complete map five, do so quicker on average than ‘Oldest’ strategy populations that complete map two, three or four this result suggests that when individuals in ‘Oldest’ strategy populations do discover the behaviour required to solve map five, it spreads rapidly through the population. As the ‘Oldest’ strategy acts somewhat like a ‘Dominance’ strategy, with only the dominant tournament winning agent acting as the teacher, it is maybe unsurprising that behaviours can spread rapidly, however the random nature of tournament selection can somewhat undermine this strategy’s ability to guarantee fit behaviours or a consistent teacher. The best performing populations for the ‘Best Parent’, ‘Oldest’ and ‘Fittest’ strategies (as seen on the left of Fig 3) also seems to indicate that once a favourable behaviour is discovered using these strategies it is able to spread reasonably quickly. This is unsurprising as each of these strategies can be highly conformist, with successful individuals potentially having a monopoly on being the teacher for new agents. With the ‘Youngest’ strategy, the high turnover of teachers provides little opportunity for beneficial behaviours to take hold, though these teaching agents are the progeny of tournament winning parents, so can be expected to be reasonably fit. The most surprising result is the general performance of the ‘Random’ strategy, given that unlike the other strategies there is no guarantee of the teacher being either consistent nor particularly fit. One reason for the ‘Random’ strat-

egy performing at least as well as the other strategies is the nature of the RC+ task itself. If a population only discovers the ‘bridge-building’ behaviour needed for maps 2-4, whilst forming a dislike for Object A and/or Object B, any conformist strategy will struggle to discover the behaviour required for map five, as the population will tend to conform to the sub-optimal behaviour. However, the very nature of the ‘Random’ strategy allows for a variety of individuals to fulfil the role of teacher, regardless of fitness, thus enabling newer ideas to potentially establish themselves and sub-optimal behaviours to be lost. However, maintaining these newly found optimal behaviours may be difficult in such a strategy. This does suggest that a hybrid approach may be beneficially, whereby numerous conformist and non-conformists strategies may exist within a population thus enabling both innovation and rapid behavioural convergence to occur.



**Fig. 3.** (Left) Graph showing the first time any population achieved each map. (Right) Graph showing the average generation populations achieved each map.

## 5 Conclusions and Further Work

The aim here was to demonstrate that multiple, varied, social learning strategies would be capable of discovering and maintaining behaviours that are inaccessible to hill-climbing strategies such as incremental genetic evolution. The results presented here echo previous work [3], while extending the research to show that various social learning strategies are capable of both discovering and maintaining inaccessible behaviours. Due to each strategy applied being highly abstracted from behaviours seen in nature, along with the task being highly artificial, this work is unable to draw strong parallels to observed social behaviours in nature. Achieving a comparable status will require a more complex use of social learning, a sensible progression would be the inclusion of synchronous, distinct learning styles into a single population. A model that allows for multiple social learning strategies to be employed along side genetic evolution has compelling

implications for agents, i.e. choosing optimal learning styles for the appropriate task. Going forward, additional tests will explore the dynamics between multiple learning strategies and incremental genetic evolution when included in the same population.

## References

1. Acerbi, A., Parisi, D.: Cultural transmission between and within generations. *Journal of Artificial Societies and Social Simulation* 9(1) (2006)
2. Borg, J.M., Channon, A.: Testing the variability selection hypothesis - the adoption of social learning in increasingly variable environments. In: *ALIFE 13: The Thirteenth Conference on the Synthesis and Simulation of Living Systems*. pp. 317–324. MIT Press (2012)
3. Borg, J.M., Channon, A., Day, C.: Discovering and maintaining behaviours inaccessible to incremental genetic evolution through transcription errors and cultural transmission. In: *ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*. pp. 101–108. MIT Press (2011)
4. Curran, D., O’Riordan, C.: The effects of cultural learning in populations of neural networks. *Artificial Life* 13(1), 45–67 (2007)
5. Galef, B.G.: Social learning and traditions in animals: evidence, definitions, and relationship to human culture. *Wiley Interdisciplinary Reviews: Cognitive Science* 3(6), 581–592 (2012)
6. Harvey, I.: Artificial evolution: A continuing saga. In: *Evolutionary robotics. From intelligent robotics to artificial life*, pp. 94–109. Springer (2001)
7. Kendal, J.R., Rendell, L., Pike, T.W., Laland, K.N.: Nine-spined sticklebacks deploy a hill-climbing social learning strategy. *Behavioral Ecology* 20(2), 238–244 (2009)
8. Laland, K.N.: Social learning strategies. *Animal Learning & Behavior* 32(1), 4–14 (2004)
9. Nakahasi, W., Wakano, J.Y., Henrich, J.: Apative social learning strategies in temporally and spatially varying environments. *Human Nature* 23(4), 368–418 (2013)
10. Ochoa, G., Harvey, I., Buxton, H.: Error thresholds and their relation to optimal mutation rates. In: *Advances in Artificial Life*, pp. 54–63. Springer (1999)
11. Reader, S.M., Biro, D.: Experimental identification of social learning in wild animals. *Learning & Behavior* 38(3), 265–283 (2010)
12. Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M.W., Fogarty, L., Ghirlanda, S., Lillicrap, T., Laland, K.N.: Why copy others? insights from the social learning strategies tournament. *Science* 328(5975), 208–213 (2010)
13. Robinson, E., Ellis, T., Channon, A.: Neuroevolution of agents capable of reactive and deliberative behaviours in novel and dynamic environments. In: *Advances in Artificial Life*, pp. 345–354. Springer (2007)
14. Stanton, A., Channon, A.: Incremental neuroevolution of reactive and deliberative 3d agents. In: *ECAL 2015: Proceedings of the Thirteenth European Conference on the Synthesis and Simulation of Living Systems*. pp. 341–348. MIT Press (2015)
15. Whiten, A., Van Schaik, C.P.: The evolution of animal cultures and social intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1480), 603–620 (2007)
16. Zwirner, E., Thornton, A.: Cognitive requirements of cumulative culture: teaching is useful but not essential. *Scientific Reports* 5 (2015)