Article

Social Learning and the Exploration-Exploitation Tradeoff

Brian Mintz * and Feng Fu

Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA; feng.fu@dartmouth.edu
* Correspondence: brian.a.mintz.gr@dartmouth.edu

Abstract: Cultures around the world show varying levels of conservatism. While maintaining traditional ideas prevents wrong ones from being embraced, it also slows or prevents adaptation to new times. Without exploration there can be no improvement, but often this effort is wasted as it fails to produce better results, making it better to exploit the best known option. This tension is known as the exploration/exploitation issue, and it occurs at the individual and group levels, whenever decisions are made. As such, it is has been investigated across many disciplines. We extend previous work by approximating a continuum of traits under local exploration, employing the method of adaptive dynamics, and studying multiple fitness functions. In this work, we ask how nature would solve the exploration/exploitation issue, by allowing natural selection to operate on an exploration parameter in a variety of contexts, thinking of exploration as mutation in a trait space with a varying fitness function. Specifically, we study how exploration rates evolve by applying adaptive dynamics to the replicator-mutator equation, under two types of fitness functions. For the first, payoffs are accrued from playing a two-player, two-action symmetric game, we consider representatives of all games in this class, including the Prisoner's Dilemma, Hawk-Dove, and Stag Hunt games, finding exploration rates often evolve downwards, but can also undergo neutral selection as well depending on the games parameters or initial conditions. Second, we study time dependent fitness with a function having a single oscillating peak. By increasing the period, we see a jump in the optimal exploration rate, which then decreases towards zero as the frequency of environmental change increases. These results establish several possible evolutionary scenarios for exploration rates, providing insight into many applications, including why we can see such diversity in rates of cultural change.

Keywords: adaptive dynamics; exploration/exploitation; social learning; evolutionary dynamics; mutation rate evolution

1. Introduction

In any learning process, individuals leverage past information along with the opinions of others to decide their best action. Broadly speaking, one can either continue using a strategy that has worked, or try a new approach. While exploration is necessary to discover better strategies, it often results in wasted effort, so it is usually better to exploit the best known strategy. These opposing approaches are very general, applying any time a decision must be made. As such, this concept is relevant across scales, both at the individual such as animal or cells, and group levels, in a wide range of areas from biology to economics [1]. Much work has gone into studying this issue from a variety of perspectives.

One can think of mutation as exploration in the space of genomes. Since all lifeforms replicate their genetic information, the study of mutation rates has been a longstanding area in biology with significant implications. One theory, the drift-barrier hypothesis, posits that natural selection favors arbitrarily small mutation rates, and is evidenced by relative measures of mutation [2]. Other studies have investigated the mechanisms for viral RNA repair mechanisms, which allow for the mutation rate to evolve up or down depending on which errors get corrected, and phenotypic switching in bacteria, finding that recombination reduces or even eliminates stable non-zero switching rates [3,4]. There has also been considerable theoretical work on mutation rates in sexually reproducing organisms, finding

higher mutation rates can be selected for or against depending on model specifics, such as the type of fitness, whether individuals are haploid or diploid, reproduce asexually or sexually, and if so with or without recombination [5,6]. Beyond the level of individual cells, decision making in humans has been studied in the exploration/exploitation framework, including its neuroscientific underpinnings [7,8]. Additionally, this approach has been employed in several areas of ecology, including foraging and analyzing host-parasite or predator-prey systems [9–11]. Through simulation and analysis, these studies found that exploration rates generally decrease to or stabilize around zero, though factors like limited lifespans or recombination can make exploration less valuable.

Computer scientists have also investigated the balance between exploration and exploitation through evolutionary algorithms, which feature a mutation parameter [12,13]. This value is critical to the success of the algorithm, however few general techniques guide its tuning. For example, particle swarm optimization is a technique that uses a collection of agents to discover optimal values in a complex space [14]. One approach, known as simulated annealing, decreases the exploration rate over time to concentrate the population around the global optimum. Yet another technique called reinforcement learning has individuals track the performance of a set of possible actions over time to determine the optimal choice [15]. In this framework, one makes an explicit policy for whether new actions are chosen to update these values, exploration, or the current best value is used, exploitation. This area has seen increasing interest from its application to artificial intelligence.

Lastly, there is a significant history of studying exploration/exploitation in economics [16]. Applications include theories of firm's flexibility and understanding product development and innovation [17–19]. By analyzing economic data, these studies characterize the optimal balance between exploiting present capabilities and exploring new ones. Additionally, March's seminal model of mutual learning in organization, where an individuals and the firm learn from each other dynamically, has been extensively studied and generalized over the last few decades in management science [20–23]. These dynamics tend to lead to low rates of exploration, which is often beneficial in the short term but detrimental in the long term.

While previous studies have applied a variety of tools from different disciplines, few have asked how the exploration rate changes over time. Our study builds on recent work that applies the technique of adaptive dynamics to answer this, determining the evolution of exploration rate [24,25]. We extend previous research by approximating a continuous trait space, considering a local model of exploration, and investigating several realistic fitness functions. Specifically, we investigate the evolutionary forces on exploration rates in dynamic environments. Earlier work has considered a finite set of traits, with fitness a function of the current environment that cycles through a finite set of possibilities, finding the optimal exploration rate was near zero, and zero was a local optima [26,27]. In contrast, our work investigates local exploration and considers a variety of contexts to determine fitness, broadly divided into two classes. The first uses a feedback mechanism between the strategies in a population and the fitness of those strategies. Specifically, we encode this as the average payoff of an individual when interacting with other players uniformly at random in a population playing a two-player two-action symmetric game. This approach is grounded in the tradition of evolutionary game theory. The other scenario we consider in this work is explicitly representing the fitness of each strategy by a time dependent function. In particular, we consider the fitness landscape with to have a single peak of some width, and whose location oscillates in time in some regular manner. Earlier work has found that in the absence of recombination, the optimal mutation rate maximizes the geometric mean of the fitness of a population [28]. Other studies have also investigated the population dynamics where a game governs fitness as above, but where the game changes over time [29,30]. This case represents the fact that few environments are static in time, and often undergo periodic changes. For example, if we think of traits as preferred nesting sites in space, then the changing fitness could apply to the study of migration or dispersal.

We will interpret some of our results through the concept of an Evolutionarily Stable Strategy (ESS), introduced by Maynard Smith and Price to describe traits that were stable under evolutionary change, which has been an influential idea throughout biology [31,32]. By using a combination of analysis and simulation, we determine the evolutionary trajectory of the exploration rate in a variety of realistic yet understudied contexts.

2. Methods

We think of the set of actions an individual could take as a bounded continuous set, specifically real numbers in the unit interval, and the best action as a trait. This may seem restrictive, but up to linear transformations it can capture any bounded trait one can reasonably assign a number to, for example an organism's height or weight. By putting traits in a space, we can ensure exploration is local, with an exploration kernel to describe the probability distribution of an individual's trait in the near future given its current trait and some exploration rate u. In this work we consider a normal distribution with variation equal to u. Specifically, the model we will use for population dynamics is the replicator mutator equation:

$$\frac{d}{dt}\vec{x} = (Q(u)F(\vec{x}, t) - \phi I)\vec{x} \qquad (1)$$

where $\vec{x}$ is the trait distribution, $Q(u)$ gives the probability of exploration from one trait to another based on a exploration rate parameter u, $F(\vec{x}, t)$ is a diagonal matrix with iith entry the fitness of trait i given the population distribution $\vec{x}$ and time t, and ϕ is the average fitness in the population (introduced as in the classical replicator equation to ensure the vector $\vec{x}$ sums to one). Essentially this equation makes individuals reproduce according to their fitness, for example by being imitated through social learning, and explore by the matrix $Q(u)$, depending on their exploration rate u. This framework is built on vectors, so the trait space is necessarily finite. Consequently, exploration cannot occur outside of the boundary, which we handle by truncating these values. Alternatively, one could accumulate them at the endpoints, or shift them to the other endpoint, making the trait space circular [33].

Equation (1) gives the short term population dynamics, and to represent the long term dynamics on exploration rate, we use the approach of adaptive dynamics. In this method, one considers the invasion fitness $f_x(y)$ of a mutant with trait y in a population of individuals all with trait x, defined as their reproduction rate. It assumes mutations on the exploration rate is rare, so it suffices to consider monomorphic populations, as one trait will fixate before the next mutant arises. Further, if we assume these mutations are small, we can use a linear approximation $f_x(y) \approx f_x(x) + (y - x)\partial_y f_x(y)|_{y=x}$. If the term $\partial_y f_x(y)|_{y=x}$ is positive, $y - x$ must be as well for the trait to fixate, so the trait will evolve upwards. The same thing happens if this term is negative, so we can think of it as the rate of change of our trait. This framework was connected to the replicator-mutator equation in [24], which determined $f_x(y)$ was the difference between the maximum eigenvalue of $Q(u)F(\tilde{x}, t^*)$ and the current average fitness, where $\tilde{x}$ is the stable distribution reached by the replicator-mutator equation under the exploration rate x, and t^* is the time the mutant emerges. Using this, we can investigate the evolution of exploration rate if we specify the fitness function $F(\vec{x}, t)$ and use the model of exploration specified above to define $Q(u)$. The code that implements this approach is available in the Github repository https://github.com/bmDart/exploration-rate-evolution.

To make fitness frequency-dependent, we consider a population playing a game. Each individual will then receive fitness that is the average payoff received over all possible interactions. Specifically, we will consider two-player, two-action, symmetric games, as these have a large degree of richness in their behavior. In these games, two players interact by each choosing one of two actions, *A* or *B*, then receive a payoff dependent on the pair of actions chosen. There are four pairs, so one write can the payoffs to a player in the matrix

	A	B
A	*a*	*b*
B	*c*	*d*

where the rows correspond to a player's choice of *A* or *B*, and the columns indicate the other player's choice of action. This class is called symmetric, as both players use the same matrix to determine their payoffs. It includes well known examples like the Prisoner's Dilemma (PD), if $b < d < a < c$, where individuals always do better by choosing the second action, even though the best outcome is both players choosing the first. Also included is the less intense version called the Hawk-Dove (HD), also called the Snowdrift, game, if $d < b < a < c$. In this game, the optimal action is the opposite of the other player's action, making this an anti-coordination game. Another game this framework encompasses is known as the Stag-Hunt (SH) game, where $b < d < c < a$. Here the optimal action is the same choice the other player makes, so this is a coordination game.

Strategies in these games can be complex, but if the game only consists of one round, and players have no information about each other, any strategy can be completely described by a probability distribution over the actions, a mixed strategy. Since there are only two possible actions, any strategy is a single number x, the probability of choosing the first action. Then the average payoff to a player with strategy y interacting with a player of strategy x is

$$R(y, x) = ayx + by(1 - x) + c(1 - y)x + d(1 - y)(1 - x)$$

Since this is linear in x, the average payoff of an y player interacting with a population with mean strategy $\bar{x}$ is just $R(y, \bar{x})$. Since this is also linear in y, so the average payoff over this population is $R(\bar{x}, \bar{x})$. This function, $R(x, x)$, can give some insight, so we will refer to it as the population fitness of a strategy x. Interestingly, this can look differently within a class of games, for example $(a, b, c, d) = (2, 0, 2.5, 1.5)$ and $(2, 0, 5, 1.5)$ are two prisoner's dilemmas, yet the first makes $R(x, x)$ concave up while the other is concave down. We will see that increasing exploration rate often leads to a more spread out stable distribution, which moves the average strategy closer to 0.5, and see the effect this will have on a population's fitness. We can also apply Adaptive Dynamics, thinking of $R(y, x)$ as $f_x(y)$ the fitness of an invading strategy y into a resident population of all x-players. Here we see the strategy should evolve according to

$$\partial_y R(y, x)|_{y=x} = ax + b(1 - x) - cx - d(1 - x)$$

This is a line connecting $b - d$ at $x = 0$ to $a - c$ at $x = 1$, so there are essentially four cases depending on the relative signs of these terms, as shown in Figure 1 with representative games, and arrows indicating the dynamics of the invading strategies. Since the diagonal cases are essentially mirrors, we consider just the three games mentioned above.
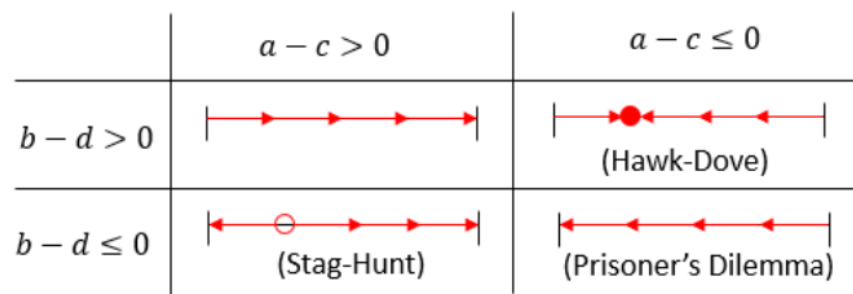
Figure 1. The four possible cases for the evolution of strategies in two-player two-action symmetric games. Depending on the entries of the payoff matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, either one endpoint will be attracting, or there will be an interior equilibrium that is either stable or unstable. The corresponding cases are labeled with their archetypal game.

Lastly, we consider time-dependent fitness functions where the fitness landscape has a single peak, of some width, whose location oscillates at some frequency. In particular, we take the fitness at time t to be the normal distribution with variance 0.1 and mean $(1 + \sin(\omega t))/2$, as this oscillates between the endpoints zero and one with period ω. We investigate the effect of changing this period and also the variance of this distribution.

3. Results

First, we used the payoff matrix

$$\begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$$

for the Prisoner's Dilemma, finding that the replicator-mutator equation stabilized at the distributions given in Figure 2. These show that lower exploration rates more closely exploit the optimal strategy of defection, as expected. However, those populations have lower fitness, as higher rates of choosing the second action, defection, are worse for the population overall, since the population fitness $R(x,x)$ is increasing for this game. Despite higher exploration rates leading to a population with greater fitness, we see the invasion fitness is only positive for lower exploration rates, so it can only evolve downward. This is a dilemma, as it is better to have a large exploration rate and flat trait distribution, but this will be selected against.
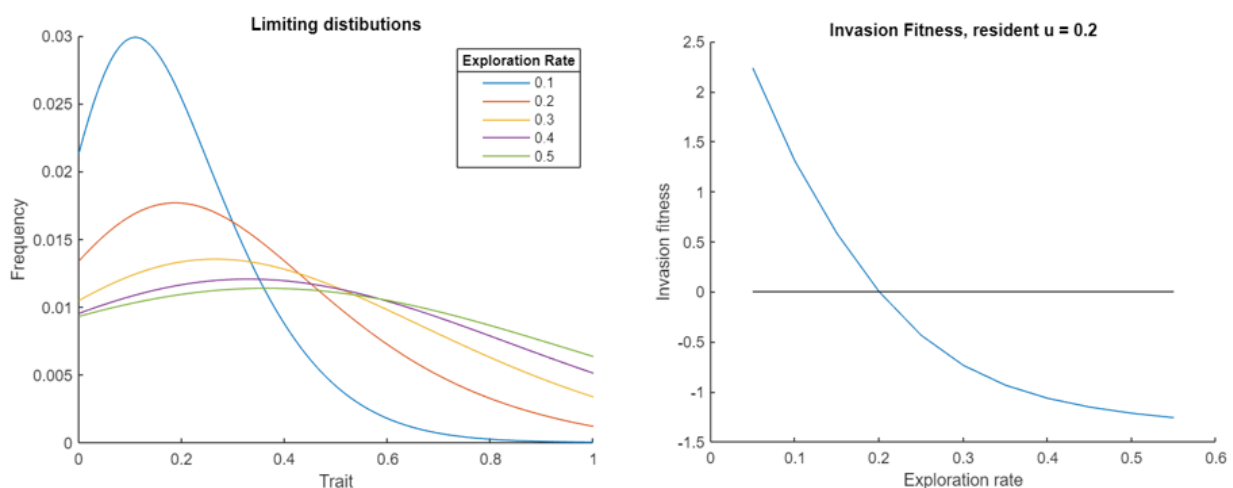

Figure 2. On the left we see the stable distributions of the replicator mutator-equation in the PD game for various exploration rates. The right plot shows the invasion fitness as a function of the invading exploration rate, for a resident value of 0.3. Since this is only positive to the left of the resident value, exploration rates can only evolve downwards. This is representative of all resident values.

The next game we considered is the Hawk-Dove game, with payoff matrix

$$\begin{bmatrix} 1 - c & 2 \\ 0 & 1 \end{bmatrix}$$

where c is a parameter representing the cost of competing over a contested resource of value one. In this case, we see populations approach the equalizer strategy $cH + (1 - c)D$, which makes all strategies have the same fitness, so there is no selection. Consequently, there is no selection on exploration rate, so it will be subject to neutral selection. This is consistent with the results of [24], which found multiple mutation rates could coexist in this game. Like in the previous game, different stable distributions are reached for different exploration rates. Here we see increasingly uniform distributions as the exploration rate increases, in Figure 3, which is expected, as this represents larger exploration. Surprisingly, we see

a dependence on the game parameter c. When this is not 0.5, the population does reach the equalizer strategy c, as seen by plotting the average strategy over time, also in Figure 3. This results in downward selection on exploration rate. Despite all Hawk-Dove games have the same strategy dynamics from the perspective of a single player, this population model demonstrates different effects depending on a parameter's value. Interestingly, despite exploration rates evolving downwards for $c \neq 0$, the population's fitness can either increase or decrease with exploration rate depending on if c is above or below one half, as seen by considering the population fitness $R(x, x)$. Thus as in the PD, it is possible the exploration rate will evolve towards value that are worse for the population overall.
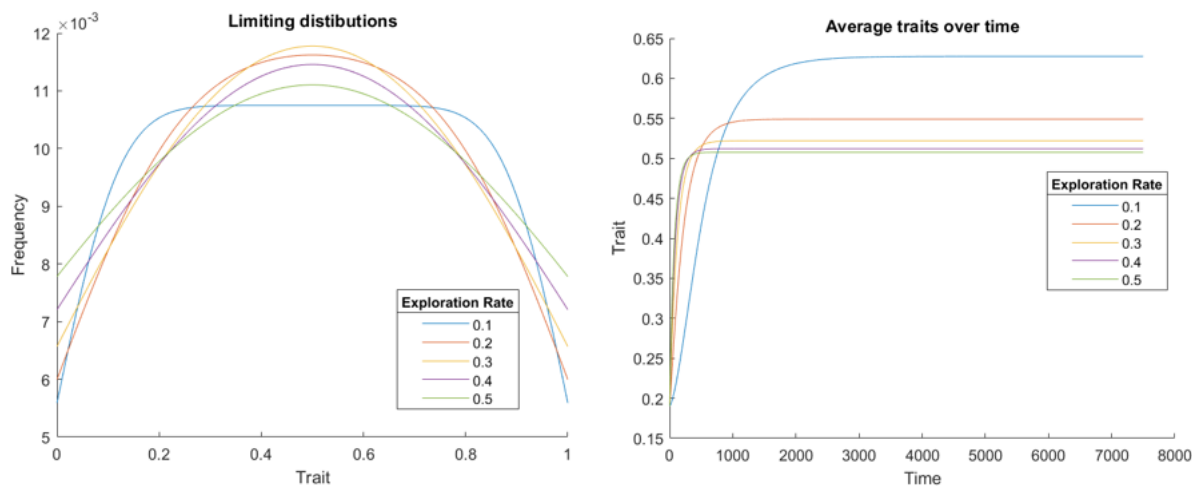
Figure 3. In the Hawk-Dove Game, we see flatter stable distributions, the first plot, for large exploration rates for $c = 0.5$. However, for $c \neq 0.5$, the average strategy does not reach the equalizer strategy c, shown in the second plot.

The final game we considered was the Stag-Hunt, with payoff matrix

$$\begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}$$

In this case, the population reaches unimodal distributions as in the Prisoner's dilemma, with individuals favoring one option more than the other. This is because the optimal action is to choose the same action as the other player, so the population becomes increasingly concentrated towards whichever pure strategy the initial mean was closer to. As such, the population will evolve away from the unstable equilibrium of 0.5, as in the single player dynamics. Depending on whether the initial mean is above or below 0.5, the population fitness is either increasing or decreasing with exploration rate, since this moves the mean strategy closer to a half, which is good if the population is concentrated around one but bad if it is concentrated around zero. Despite this, exploration rates can only evolve downwards in both case, so in this game too, exploration rates can evolve to less desirable levels. However in this case, selection becomes neutral for sufficiently large initial exploration rates, since the population becomes centered around 0.5. Thus, exploration rates that start large will drift up and down, but eventually become caught around zero.

The other type of fitness we considered in this work had fitness an explicit time-dependent function, with no dependence on the distribution of strategies in the population. Specifically, we took $f(x, t) = \exp(-(x - (1 + \sin(\omega t))/2)^2)$ where ω is a parameter for the period of the oscillations. Here, one may also use the replicator-mutator equation to simulate the population dynamics, but now populations need not reach stable distributions. For example, a periodic fitness will lead to periodic changes in the population. Nonetheless, we can adapt the results of the model by considering a time averaged fitness. Since the

fitness function does not depend on the frequency of strategies, an invading subpopulation with a novel exploration rate will grow independently of the resident population. Thus, the exploration rate leading to a higher average fitness will eventually fixate. Here one must use the geometric mean of fitness, as populations grow geometrically. This is because fitness is essentially a reproduction rate, which are multiplied, not added, together to aggregate over time periods, as is done in the geometric mean. Indeed, the order of the geometric and arithmetic mean might swap between two sets, for example $\{50, 50\}$ and $\{100, 1\}$.

In Figure 4, we plot the time averaged fitness of each exploration rate, using fitness functions of various periods. For small periods, we see fitness is maximized around zero, decreasing with larger exploration rates until it reaches a local minimum then starts to increase. This means that for rapidly changing environments, it is best to have minimal exploration rate. However, if it starts above this minimum, exploration rates will increase arbitrarily high. This suggests some rates result in the population lagging behind the optimal strategy, to the extent that a uniform distribution is more effective. We see the opposite curve for sufficiently large periods, where environmental change is slow. Here, there is a local maximum at some nonzero exploration rate, indicating an intermediate level of exploration is optimal. Interestingly, as the period changes, the optimal exploration rate makes a jump from zero to an intermediate value. The exact period where this occurs and value the optimal rate jumps to would depend on specifics of the model, namely the type of curve defining the fitness. Further, optimal rate decreases with increasing period, that is, slower changing environments. This makes sense, as a sufficiently slow changing environment is effectively stable, for which arbitrarily small exploration rates are usually optimal. Comparable effects in the evolution of exploration rate are observed when the normal distribution has wider variance, indicating the generality of these results. Theoretically, one could also compute the time averaged fitness of the limiting exploration rates. When exploration rate is zero, the population will likely be entirely at the strategy that maximizes the time averaged fitness function, and when it is infinite, the strategy distribution will be uniform, so the time averaged fitness will simply be the average value of the function (which is constant in time, so equals its time average).
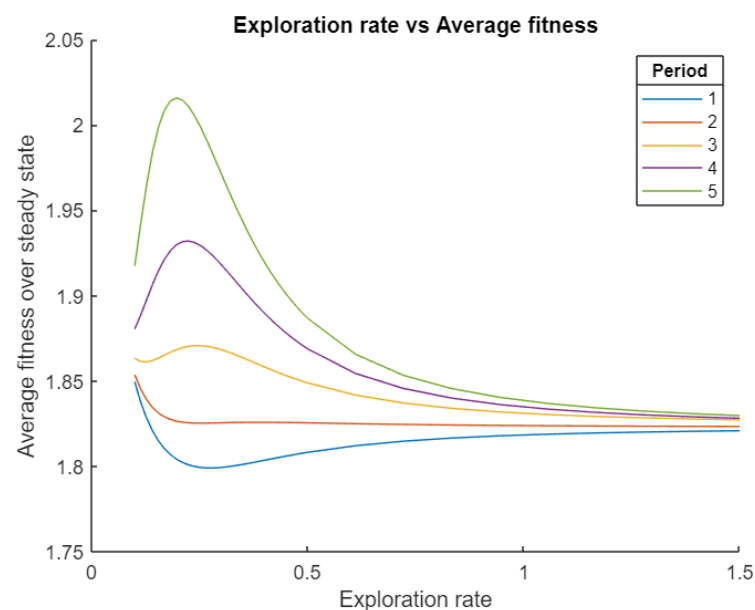
Figure 4. For fitness that explicitly depends on time, we see both unstable and stable equilibria in the evolution of exploration rate. When environmental change is slow, corresponding to a long period, an intermediate level is optimal. Whereas for fast environmental change arbitrarily small exploration rates are optimal, though if the initial value is large enough, they will become arbitrarily large.

4. Discussion

In this work, we investigated the evolution of exploration rate under variable selection, employing adaptive dynamics and the replicator-mutator equation. For frequency dependent fitness encoded by two-player/two-action, symmetric games, we found that exploration rates often evolve downward, but neutral selection is also possible. This means in most cases the exploration rate of zero (approaching from above) constitutes an ESS. Despite this, it is possible in all games we considered for larger exploration rates to be more beneficial to the population. The precise form of the exploration rate's evolution can also depend on the game's parameters or the initial conditions, as in the HD and SH games respectively. This suggests that while the cases we studied are representative of the possible dynamics in this class of games, further richness could be observed in future study. However, we conjecture that this class of games is incapable of selecting for large exploration rates, as opposed to the more complicated class of two-player/three-action symmetric games, where it was found the Rock-Paper-Scissors game led to selection for an intermediate level of mutation [25]. This is because cyclic dynamics in this game allow for a sub-population with multiple traits to remain resilient as the composition of the population changes. Since cyclic dynamics cannot be observed in the smaller class of games, it is likely that larger exploration rates cannot be selected for with these types of fitness function. Then with fitness function a single peak that oscillated according to some frequency, we found both attracting and repelling equilibria depending on the period. For fast changing environments, arbitrarily small exploration rates are optimal, though a sufficiently large initial exploration rate will evolve upwards. In contract, slow changing environments have intermediate optimal exploration rates, and evolution proceed towards this. As the rate of oscillation decreases further, this optimal value approaches zero [34].

Previous work on exploration/exploitation has applied a wide range of techniques to determine the optimal exploration rate, but not how it evolves. Our research answers this question in a variety of environments, finding lower rates are usually selected for, but time dependent fitness can make nonzero or even arbitrarily large exploration evolve. We expand on previous research by using a local model of exploration, continuous approximation of traits, and the approach of adaptive dynamics. The results we find complement experimental data from biology, providing additional theoretical evidence for the drift-barrier hypothesis, as seen in our consideration of frequency dependent fitness. These findings could also help explain how behaviors like foraging evolve, or the benefits and harms of different levels of conformity in social groups [35] and the dynamics that select for or against exploration. Our results suggest that genetic algorithms [36] might be improved by generalizing the mutation to include the exploration rate itself, rather than tuning this value manually. Lastly, the findings of this study could imply that the optimal tradeoff between exploration and exploitation found in companies can be reached through market forces [37], as these are analogous to natural selection.

5. Conclusions

Exploration and exploitation are two competing effects that combine to determine an optimal strategy. Since one can never know if their current approach is best, exploration is necessary to some degree to find better strategies. However, this often produces worse results, especially when the current method performs well, so it is also important to balance this exploration of new ideas with exploitation of the best known ones. This general dilemma is found in many areas of research, including biology, ecology, computer science, and economics.

The goal of our research was to see what resolution between these two factors is reached through natural selection, by studying the evolution of an exploration rate parameter. To answer this, we applied adaptive dynamics to the replicator-mutator equations, which gave the invasion fitness of exploration rates in various contexts. We consider fitness that depends on the relative frequencies of other traits, and fitness that varies explicitly in time. In the first case we found exploration rates often evolved towards zero, which

was then an ESS, though trajectories could vary based on game parameters or initial conditions. In the second case, we observed a discontinuous transition in optimal exploration rate, from zero to an intermediate value which decreased back to zero as environmental change slowed.

The generality of this framework makes it ideal to study several related question about exploration versus exploitation. Future study could consider other trait topologies, by adapting the matrix $Q(u)$. For example, ref. [33] explores how mutation rates can evolve upwards even in a fixed environment. This is found not just for traits in some interval, but also in a circular space or finite strings on a finite alphabet. Preliminary results showed the HD game with $c = 0.5$ led to increasingly polarized distributions are exploration rates decreased, if exploration outside of the interval was accumulated at the endpoints. In addition, making the trait space circular caused the time averaged fitness to strictly decrease with exploration rate, indicating that in the absence of asymmetry, there is no benefit to an intermediate level of exploration. Multidimensional trait spaces could also be considered, but without dependencies between the axes, this may reduce to several copies of a one dimensional trait. The fitness functions could also be changed within this framework. For example, one could consider fitness that comes from nonlinear or multiplayer games, like the Public Goods Game, or stochastic fitness functions, such as jumping to a random position at some constant frequency, or some constant positions with some random frequency. Recent studies suggest that interesting results could been found in this area [38,39] One could even make the exploration rate itself non-constant, possibly modeling it as a decreasing function of time, such as a linear or exponential function, and study the evolution of the parameters of these functions. Lastly, one could experiment with more intelligent agents. Whereas agents in this model explored randomly, one could use a reinforcement learning framework like Q-learning to model agents who explore based on previous knowledge. Such modifications would certainly change the balance of importance between exploration and exploitation, likely leading to different results.

Author Contributions: Conceptualization, B.M. and F.F.; methodology, B.M. and F.F.; software, B.M.; validation, B.M.; formal analysis, B.M. and F.F.; investigation, B.M. and F.F.; resources, B.M. and F.F.; data curation, B.M.; writing—original draft preparation, B.M.; writing—review and editing, B.M. and F.F.; visualization, B.M. and F.F.; supervision, F.F.; project administration, B.M. and F.F.; funding acquisition, F.F. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Code and data are available at https://github.com/bmDart/exploration-rate-evolution.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Berger-Tal, O.; Nathan, J.; Meron, E.; Saltz, D. The exploration-exploitation dilemma: A multidisciplinary framework. *PLoS ONE* **2014**, *9*, e95693. [CrossRef] [PubMed]
2. Lynch, M.; Ackerman, M.S.; Gout, J.F.; Long, H.; Sung, W.; Thomas, W.K.; Foster, P.L. Genetic drift, selection and the evolution of the mutation rate. *Nat. Rev. Genet.* **2016**, *17*, 704–714. [CrossRef] [PubMed]
3. Domingo, E.; García-Crespo, C.; Lobo-Vega, R.; Perales, C. Mutation rates, mutation frequencies, and proofreading-repair activities in RNA virus genetics. *Viruses* **2021**, *13*, 1882. [CrossRef]
4. Liberman, U.; Van Cleve, J.; Feldman, M.W. On the evolution of mutation in changing environments: Recombination and phenotypic switching. *Genetics* **2011**, *187*, 837–851. [CrossRef]
5. Bürger, R. Evolution of genetic variability and the advantage of sex and recombination in changing environments. *Genetics* **1999**, *153*, 1055–1069. [CrossRef]

6.　Romero-Mujalli, D.; Jeltsch, F.; Tiedemann, R. Elevated mutation rates are unlikely to evolve in sexual species, not even under rapid environmental change. *BMC Evol. Biol.* **2019**, *19*, 1–9. [CrossRef] [PubMed]

7.　Navarro, D.J.; Newell, B.R.; Schulze, C. Learning and choosing in an uncertain world: An investigation of the explore–exploit dilemma in static and dynamic environments. *Cogn. Psychol.* **2016**, *85*, 43–77. [CrossRef]

8.　Laureiro-Martinez, D.; Brusoni, S.; Zollo, M. The neuroscientific foundations of the exploration- exploitation dilemma. *J. Neurosci. Psychol. Econ.* **2010**, *3*, 95. [CrossRef]

9.　Eliassen, S.; Jørgensen, C.; Mangel, M.; Giske, J. Exploration or exploitation: Life expectancy changes the value of learning in foraging strategies. *Oikos* **2007**, *116*, 513–523. [CrossRef]

10.　M'Gonigle, L.; Shen, J.; Otto, S. Mutating away from your enemies: The evolution of mutation rate in a host–parasite system. *Theor. Popul. Biol.* **2009**, *75*, 301–311. [CrossRef]

11.　Monk, C.T.; Barbier, M.; Romanczuk, P.; Watson, J.R.; Alós, J.; Nakayama, S.; Rubenstein, D.I.; Levin, S.A.; Arlinghaus, R. How ecology shapes exploitation: A framework to predict the behavioural response of human and animal foragers along exploration–exploitation trade-offs. *Ecol. Lett.* **2018**, *21*, 779–793. [CrossRef] [PubMed]

12.　Črepinšek, M.; Liu, S.H.; Mernik, M. Exploration and exploitation in evolutionary algorithms: A survey. *ACM Comput. Surv. (CSUR)* **2013**, *45*, 1–33. [CrossRef]

13.　Eiben, A.E.; Schippers, C.A. On evolutionary exploration and exploitation. *Fundam. Informaticae* **1998**, *35*, 35–50. [CrossRef]

14.　Kumar, K.P.; Singarapu, S.; Singarapu, M.; Karra, S.R. Balancing Exploration and Exploitation in Nature Inspired Computing Algorithm. In *Intelligent Cyber Physical Systems and Internet of Things: ICoICI 2022*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 163–172.

15.　Yen, G.; Yang, F.; Hickey, T. Coordination of exploration and exploitation in a dynamic environment. *Int. J. Smart Eng. Syst. Des.* **2002**, *4*, 177–182. [CrossRef]

16.　Almahendra, R.; Ambos, B. Exploration and exploitation: A 20-year review of evolution and reconceptualisation. *Int. J. Innov. Manag.* **2015**, *19*, 1550008. [CrossRef]

17.　Mathias, B.D.; Mckenny, A.F.; Crook, T.R. Managing the tensions between exploration and exploitation: The role of time. *Strateg. Entrep. J.* **2018**, *12*, 316–334. [CrossRef]

18.　Greve, H.R. Exploration and exploitation in product innovation. *Ind. Corp. Chang.* **2007**, *16*, 945–975. [CrossRef]

19.　Gilsing, V.; Nooteboom, B. Exploration and exploitation in innovation systems: The case of pharmaceutical biotechnology. *Res. Policy* **2006**, *35*, 1–23. [CrossRef]

20.　March, J. Exploration and exploitation in organizational learning. *Organ. Sci.* **1991**, *2*, 71–87. [CrossRef]

21.　Bocanet, A.; Ponsiglione, C. Balancing exploration and exploitation in complex environments. *Vine* **2012**, *42*, 15–35. [CrossRef]

22.　Lazer, D.; Friedman, A. The network structure of exploration and exploitation. *Adm. Sci. Q.* **2007**, *52*, 667–694. [CrossRef]

23.　Posen, H.E.; Levinthal, D.A. Chasing a moving target: Exploitation and exploration in dynamic environments. *Manag. Sci.* **2012**, *58*, 587–601. [CrossRef]

24.　Allen, B.; Rosenbloom, D.I.S. Mutation Rate Evolution in Replicator Dynamics. *Bull. Math. Biol.* **2012**, *74*, 2650–2675.

25.　Rosenbloom, D.I.; Allen, B. Frequency-dependent selection can lead to evolution of high mutation rates. *Am. Nat.* **2014**, *183*, E131–E153. [CrossRef] [PubMed]

26.　Nilsson, M.; Snoad, N. Optimal mutation rates in dynamic environments. *Bull. Math. Biol.* **2002**, *64*, 1033–1043. [CrossRef]

27.　Ben-Porath, E.; Dekel, E.; Rustichini, A. On the relationship between mutation rates and growth rates in a changing environment. *Games Econ. Behav.* **1993**, *5*, 576–603. [CrossRef]

28.　Ishii, K.; Matsuda, H.; Iwasa, Y.; Sasaki, A. Evolutionarily stable mutation rate in a periodically changing environment. *Genetics* **1989**, *121*, 163–174. [CrossRef]

29.　Shu, L.; Fu, F. Eco-evolutionary dynamics of bimatrix games. *Proc. R. Soc. A* **2022**, *478*, 20220567. [CrossRef]

30.　Wang, X.; Fu, F. Eco-evolutionary dynamics with environmental feedback: Cooperation in a changing world. *Europhys. Lett.* **2020**, *132*, 10001. [CrossRef]

31.　Hines, W. Evolutionary stable strategies: A review of basic theory. *Theor. Popul. Biol.* **1987**, *31*, 195–272. [CrossRef]

32.　Taylor, P.D.; Jonker, L.B. Evolutionary stable strategies and game dynamics. *Math. Biosci.* **1978**, *40*, 145–156. [CrossRef]

33.　Mintz, B.; Fu, F. The Point of No Return: Evolution of Excess Mutation Rate Is Possible Even for Simple Mutation Models. *Mathematics* **2022**, *10*, 4818. [CrossRef]

34.　Yang, V.C.; Galesic, M.; McGuinness, H.; Harutyunyan, A. Dynamical system model predicts when social learners impair collective performance. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2106292118. [CrossRef] [PubMed]

35.　Evans, T.; Fu, F. Opinion formation on dynamic networks: Identifying conditions for the emergence of partisan echo chambers. *R. Soc. Open Sci.* **2018**, *5*, 181122. [CrossRef]

36.　Coley, D.A. *An Introduction to Genetic Algorithms for Scientists and Engineers*; World Scientific Publishing Company: River Edge, NJ, USA, 1999.

37.　Lavie, D.; Rosenkopf, L. Balancing exploration and exploitation in alliance formation. *Acad. Manag. J.* **2006**, *49*, 797–818. [CrossRef]

38. Carja, O.; Liberman, U.; Feldman, M.W. Evolution in changing environments: Modifiers of mutation, recombination, and migration. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 17935–17940. [CrossRef]
39. Matic, I. Mutation rate heterogeneity increases odds of survival in unpredictable environments. *Mol. Cell* **2019**, *75*, 421–425. [CrossRef]